

Eléonore BESNEHARD

Master 2 Information-Documentation / Parcours GIDE

Mémoire de stage

VOLUME D'ANNEXES

**Élaboration d'une enquête pour l'évaluation orientée usagers de
la recherche dans Gallica : entre usages actuels et perspectives**

Sous la direction de :

Mme Joana Casenave (tutrice pédagogique, laboratoire Geriico, Université de Lille)

M. Arnaud LABORDERIE (tuteur professionnel, service de la Coopération numérique et de Gallica, BnF)

M. Eric KERGOSIEN (tuteur professionnel, laboratoire Gérico, Université de Lille)

Soutenu le 19 septembre 2024

Université de Lille (campus Pont-de-Bois)

BP 60 149, 59 653 Villeneuve d'Ascq Cedex

Année universitaire 2023/2024

Table des matières

Figure 1 : Organigramme de la BnF (mis à jour le 18 juillet 2024).....	4
Transcription de l’atelier métier 2 : 16 avril 2024	5
Partie 1 : Les témoignages	5
Partie 3 : Les améliorations	30
Transcription de l’atelier métier 3 : 30 avril 2024	50
Temps 1 : Témoignages	50
Temps 2 : Explications moteur.....	82
Temps 3 : Améliorations	84
Transcription de l’atelier métier 4 : 16 mai 2024.....	97
Temps 1 : Témoignages	97
Temps 2 : Explications moteur.....	129
Temps 3 : Améliorations	129
Entretiens individuels.....	130
Protocole d’entretien	130
Transcription atelier GMB : 27 juin 2024.....	132
Temps 1 : Témoignages	132
Temps 2 : Explications du moteur.....	157
Temps 3 : Améliorations	157
Atelier Jeunes chercheurs : 12 juin 2024.....	167
Introduction	167
Figure 2 : Support de présentation atelier Jeunes Chercheurs	167
Figure 3 : Introduction, support de présentation atelier Jeunes Chercheurs	167
Figure 4 : Présentation de l’enquête, support de présentation atelier Jeunes Chercheurs	168
Figure 5 : Présentation de l’atelier, support de présentation atelier Jeunes Chercheurs	168
Figure 6 : Temps 1, support de présentation atelier Jeunes Chercheurs	169

Figure 7 : Temps 2, support de présentation atelier Jeunes Chercheurs	169
Figure 8 : Pause, support de présentation atelier Jeunes Chercheurs	170
Figure 9 : Temps 3, support de présentation atelier Jeunes Chercheurs	170
Figure 10 : Conclusion, support de présentation atelier Jeunes Chercheurs	171
Figure 11 : Remerciements, support de présentation atelier Jeunes Chercheurs	171
Partie 1 : Exploration.....	172
Figure 12 : Guide d'exploration, atelier Jeunes Chercheurs	178
Figure 13 : Guide d'exploration p.1, atelier Jeunes Chercheurs	179
Figure 14 : Guide d'exploration p.2, atelier Jeunes Chercheurs	180
Figure 15 : Guide d'exploration p.3, atelier Jeunes Chercheurs	181
Figure 16 : Guide d'exploration p.4, atelier Jeunes Chercheurs	182
Partie 2 : Témoignages	183
Partie 3 : Améliorations.....	203
Le questionnaire biblionautes.....	213
Figure 17 : Cumul des répondants au cours de la période de diffusion de l'enquête.....	213
Extraction des réponses au questionnaire biblionautes	213
Figure 18 : Nombre de répondants par direction/ délégation.....	213
Figure 19 : Nombre de répondants au sein de la DCO.....	214
Figure 20 : Ancienneté à la BnF.....	214
Figure 21 : Ancienneté d'utilisation de Gallica.....	215
Figure 22 : Fréquence d'utilisation de Gallica	216
Figure 23 : Cadre de consultation de Gallica	216
Figure 24 : La plupart du temps, les collègues cherchent... ..	217
Figure 25 : La part des recherches mono-motifs par type de document	217
Figure 26 : Pour rechercher un document ciblé, les collègues passent par... ..	218
Figure 27 : Pour une recherche exploratoire, les collègues passent par.....	218
Figure 28 : L'utilisation des fonctionnalités de recherche avancée	219

Figure 29 : L'efficacité des modalités d'apprentissage de la recherche dans Gallica....	219
Figure 30 : Les formations à la recherche dans Gallica demandées par les répondants	221
Figure 31 : Satisfaction générale des collègues par rapport à la recherche dans Gallica	222
Figure 32 : Calculs relatifs à la moyenne obtenue	222
Figure 33 : Une bonne séance sur Gallica c'est	223
Figure 34 : Échelle de Likert sur plusieurs affirmations	223
Figure 35 : Fonctionnalités de la recherche avancée préférées par les répondants	227
Figure 36 : Fonctionnalités mal comprises ou peu appréciées des répondants	231
Figure 37 : Ce qui manque aux répondants pour chercher dans Gallica.....	235
Figure 38 : Classement par ordre de préférence/ priorité des propositions d'amélioration de la recherche dans Gallica.....	236
Figure 39 : Nombre de fois où chaque proposition s'est retrouvée en tête du classement (1 ^{ère} , 2 ^e ou 3 ^e place)	237
Figure 40 : Champs libre pour que les collègues puissent s'exprimer.....	241



Figure 1 : Organigramme de la BnF (mis à jour le 18 juillet 2024)

Transcription de l'atelier métier 2 : 16 avril 2024

Partie 1 : Les témoignages

Fanny : [...] Vous venez de départements assez différents. L'idée c'est que chacun puisse exposer ses usages qui vont être très différents. Bien sûr, il y a parmi vous des personnes qui font de la recherche uniquement iconographique, de la recherche sur les manuscrits, uniquement sur... Voilà l'idée, c'est que vous puissiez parler de vos recherches qu'elles soient spécifiques ou moins spécifiques. Des difficultés que vous rencontrez, des succès et aussi peut être et de quels sont les utilisateurs cibles que vous rencontrez, c'est-à-dire certains membres du service public ? Vous avez forcément des retours. On en parlait tout à l'heure beaucoup, quel est le public, de la même manière que quand vous faites des formations, comment elle est présentée, il y a un certain public et une certaine attente. C'est aussi ces usages là qu'on a besoin de comprendre. Alors bien sûr, il y a des choses que vous allez dire, on est déjà au courant, c'est pas grave, dites-les, il faut les dire. Comme le disait Eleonore, il y a vraiment, il n'y a pas, il n'y a pas de mauvaise réponse, pas de mauvaise expérience, retour, etc. Dans un premier temps, voilà comment on voudrait faire les choses. On vous propose de faire un retour d'expérience de Gallica à partir de vos recherches. Donc on exclut la partie billet de blog, des sélections. Mais après, si vous utilisez par exemple les pages sélection pour arriver sur un résultat de recherche, c'est aussi intéressant parce que peut être que vous avez passé déjà le moteur de recherche. Ce que je veux dire, c'est que c'est vraiment la recherche, le résultat dans Gallica, la présentation, la visualisation des collections de cette manière-là. Donc, on vous propose de nous parler de nouveaux usages, des difficultés que vous rencontrez, des succès que vous rencontrez, de ce que vous comprenez, de ce que vous ne comprenez pas, des stratégies que vous avez mises en place pour arriver à des résultats qui vous semblent cohérent. On va noter tout ça et Eléonore va noter ce que vous dites, les mots-clés, etc. Et puis après, l'idée, c'est qu'on essaye dans un second tour de table de catégoriser vos retours entre est ce que c'est spécifique est-ce que c'est général ? Est-ce que c'est positif ou négatif ? L'idée étant de placer ça sur... de faire un peu de matrice et de placer sur une carte et de se rendre compte de quelles sont les priorités demain ? Vers où on peut aller et donc ce sera le deuxième point. Le troisième point, je vous propose peut-être par groupe de trois pour discuter un peu plus librement, de penser suivant vos usages trois à quatre améliorations et évolutions qui seraient vraiment fondamentales pour vos recherches dans le moteur de recherche. Est-ce que ça vous semble clair ? Vous avez des questions ? Donc sachant que je voulais vous dire que vous venez de département différent, il

y aura un second atelier avec des départements différents, on aura sûrement les mêmes retours, mais voilà, lâchez-vous !

Arnaud : Juste pour compléter, on essaiera de faire une typologie, d'essayer de rassembler par exemple ce qui concerne l'interface, en ce qui concerne... Donc on regardera pour arriver à catégoriser selon une typologie qu'on va travailler ensemble justement. Et peut-être aussi toutes les stratégies de contournement que vous avez. Ça aussi, ça nous intéresse. Quels sont les chemins empruntés qui ne sont pas sans doute le chemin du parcours utilisateur tel qu'il a pu être pensé dans les interfaces ?

Fanny : Est-ce qu'une personne qui veut commencer ? Qui veut témoigner en premier ? Elle peut même choisir sa couleur. [Rires]

C. A. : Ça ne me dérange pas de commencer. Alors moi je me sers de Gallica pour plusieurs choses différentes. L'idée c'est que moi je valorise les collections de la BnF selon des thématiques, des actualités, des nouvelles mises en ligne, etc. Grosso modo, j'ai deux gros blocs de recherche. Un bloc : je recherche une image pour illustrer ce que je suis en train de raconter. Et un autre bloc : je cherche un document textuel. Alors je le dis comme ça pardon hein mais ça regroupe manuscrits, livres.

Fanny : Des documents écrits.

C. A. : Oui, documents écrits. Pour arriver à trouver l'image que je veux. J'ai différentes façons de le faire. Soit on va partir du principe que je ne connais pas cette image, parce que quand je la connais j'ai d'autres stratégies... Soit je me dis bon, je vais dans Gallica, je cherche une image qui illustre le printemps, je tape "printemps" et soit je vais chercher dans les sélections qui existent, soit je vais trier les résultats image dans les résultats de recherche...

Fanny : En utilisant une recherche simple.

C. A. : Oui, parce que la recherche avancée, ça ne marche pas très bien. En tout cas, moi je n'y arrive pas donc voilà. Autre chose, c'est que vraiment je n'arrive pas à trouver ce qui me plaît, donc je change de termes dans tous les sens. Quand je fais "printemps", je fais des "fleurs", je fais "amour" et un tas de trucs. Oui pardon hein.

Fanny : "Bourgeons"

C. A. : Oui voilà, "Chanson". Mais honnêtement, ça arrive assez souvent que je passe par Google Images. Je fais le "printemps BnF" ou "printemps Gallica". [Rire] Et là, souvent, il me

remonte d'emblée en fait un résultat de recherche visuel où là, moi mon gros problème, c'est de passer en revue toutes les pages. Vous voyez ce que je veux dire ?

Fanny : Oui !

C. A. : Alors qu'en terme d'expérience utilisateur, j'ai toutes les images d'un coup sur Google Images. Et souvent je ne vais pas jusqu'à la deuxième page de résultats. Voilà, ça c'était pour les images.

Pour le texte, comment vous dire que ? [Rire] C'est très différent. Donc, on va partir dans un cas de figure qui est un peu particulier, en fait moi souvent je vais rechercher une page en particulier, voire une citation ou une phrase en particulier dans le truc...

Fanny : Que tu connais à l'avance ?

C. A. : Ouais ! Et alors ça... comment vous dire que... [rire] j'ai souvent envie de mourir [rires] avant de commencer. [Rire] Mais en fait, j'ai là aussi différentes techniques. La première technique, c'est que je sais où ça se situe, à peu près dans quel chapitre, à quel endroit. Donc là, je cherche la table des matières dans le document pour remonter sur la page, etc. Après je me sers aussi, pareil, de la recherche Google en tapant "*Gallica plus la citation*" et souvent ça remonte avec l'OCR ou le document en .txt là ce que je veux. Ou alors ben je mets le document .txt et je fais un contrôle Find dedans.

Arnaud : Tu utilises les guillemets ?

C. A. : Non non. [Rire] Pardon hein mais...

Fanny : C'est bien, il n'y a pas de mauvais heu...

C. A. : Je ... Mais honnêtement je n'ai aucun doute sur le fait qu'il y a de meilleures techniques que ça. Mais je brasse un tel nombre de documents par jour que c'est, il faut que ça aille vite quoi.

Arnaud : Donc tu affiches le texte brut et tu recherches dedans ?

C. A. : Ouais et ça, ça marche super bien.

Fanny : Et est-ce que tu sais si dans ton service, ou département il y a d'autres stratégies ? Est-ce que vous avez déjà parlé de ça ? Est-ce que vous avez déjà parlé de ça ?

C. A. : Alors oui, si sur les images on s'en est pas mal parlé, sur les textes, c'est assez différent parce qu'on a des méthodes de pensée très différentes toutes les trois, notamment parce qu'on

n'a pas les mêmes formations et on n'a pas les mêmes origines culturelles. Donc [rire] on a des cerveaux pas câblés pareil. C'est vachement intéressant, mais en tout cas, sur les images, je sais que j'ai une collègue qui passe quasi systématiquement par la banque d'images plutôt mais en back office hein ! Pas en banque d'image... [rire] Pardon, moi je dis la vérité. Parce qu'en fait, quand dans la banque d'images, on passe par Iris et on fait une recherche soit par nom du document, soit même par terme, ça remonte super bien [rire] ce qui n'est pas toujours le cas en font, c'est assez magique. Et je sais qu'il y en a une des deux qui ne se sert pas de la recherche Google Image, comme moi. Je n'ai toujours pas compris pourquoi, parce que j'ai dit mais t'as vu, ça va vachement plus vite.

Fanny : Ok.

C. A. : Et je sais qu'après, sur les images, on se sert aussi beaucoup des sélections, c'est bête mais nous, ça nous fait gagner un temps infini. Et on a aussi... j'ai une collègue aussi qui passe pas mal par les essentiels parce qu'elle trouve des documents en rapport avec ce qu'elle est en train de raconter. Et après, elle recherche par cote dans Gallica. Ouais, c'est chelou hein ?

Fanny : [soupir] Chacun et chacune nous avons notre...

Arnaud : [pas compris]

C. A. : Ben ouais, mais en fait je ne sais pas pourquoi ?

Fanny : Pas forcément toujours, mais là.

C. A. : Voilà, donc c'est ça.

Fanny : Et tu parlais tout à l'heure de l'interface du coup ? Est-ce que tu utilises la vue mosaïque par exemple pour les images ?

C. A. : Oui, tu veux dire quand on est dans un document et on voit tout ?

Fanny : Alors, dans le document ou dans la liste de résultat ?

C. A. : Ah non, je n'ai jamais fait ça.

Fanny : En haut à gauche, la liste des résultats soit en liste soit en mosaïque par exemple pour les images, tu auras des résultats comme dans Google Images.

C. A. : Ok, je ne le savais pas.

Fanny : Non c'était pour savoir si tu l'utilisais, si tu penses que...

C. A. : Non, non, ça peut être pas mal

Fanny : Et ce n'est pas forcément visible, ok. Et tu parlais tout à l'heure du fait que t'allais rarement en page 2 de Google. Du coup, j'imagine que ça veut dire que tu vas en page 2 de Gallica ?

C. A. : Ouais. Ah ouais, vachement ouais.

Fanny : Est-ce que tu penses que c'est parce que tu ne trouves pas dès le début ou est-ce qu'il n'y a pas assez de documents qui sont présentés sur la même page ?

C. A. : Les deux et en fait. Il faudrait que je le refasse pour voir. Mais je pense que -souvent moi je cherche des images qui claquent un peu- voilà un truc pour les réseaux sociaux soyons honnêtes, et c'est difficile à évaluer sur Gallica au premier coup d'œil, alors que sur Google image...

Fanny : L'image est plus prenante.

C. A. : Je ne sais pas. Il faudrait que je le refasse pour vous qualifier pourquoi ? C'est plus simple pour moi sur Google image mais.

Arnaud : Et les résultats, combien de pages tu feuilletes dans Gallica ? Peut-être que tu affiches ?

C. A. : Ça dépend.

Arnaud : Est-ce que tu feuilletes 50 résultats et tu as 15 résultats par page ? Tu modifies ces paramètres ?

C. A. : Parfois ouais, quand je ne trouve pas rapidement ouais. Mais, ça dépend. Ça dépend franchement. Quand j'ai vraiment un truc précis en tête, c'est de donner forme à ma pensée. Ça dépend.

Fanny : Est ce qu'il y a des personnes pour qui ça fait écho à leurs recherches, à leurs expériences sur Gallica ?

XX : C'est vrai que parfois pour trouver la bonne image pour illustrer un document, on est obligés de feuilleter le document pour trouver, ça ne sera pas forcément l'image qui parlera le plus, qui sera en première page sur Gallica.

A. T. : Peut être juste sur la question de la mosaïque, pour donner un autre point de vue. Dans Mandragore, on n'a pas réussi à le développer. Ça a été très demandé par l'étude de public. Le

public nous a dit : “mais pourquoi vous n’avez plus de mosaïque ? C’est tellement plus simple !”

XX : La mosaïque c’est ?

A. T. : L’affichage des résultats sans avoir le texte, juste l’image. Sachant que dans Gallica, si c’est des livres on à la première page n'en a pas forcément les plus jolies...

Fanny : On n’a pas choisi.

A. T. : Oui voilà.

Fanny : Est-ce qu’il y a une autre personne qui vous livre son témoignage d’usage ?

A. B. : Moi, je veux bien prendre la suite. Oui, moi c’est pour... On a des bibliothèques numériques et on fait des liens, on envoie directement vers les documents qui sont sur Gallica. Donc par exemple, si je veux intégrer un article dans mon site, il y a des références à des documents, je vais aller chercher un document dans Gallica pour pouvoir mettre le lien dans mon article. Encore hier par exemple, le truc le plus simple que je fais, je ne vais pas aller très loin, je prends le titre et je vais coller dans le moteur de recherche Gallica. La plupart du temps, ça va marcher tout de suite. Je vais avoir le document dans les trois premiers résultats. Parfois ça ne fonctionne pas. Donc j’ajoute les guillemets pour avoir un juste le début du document et le trouver. Mais parfois ça ne marche pas et si j’ai le temps, je vais passer par le catalogue général. Je vais passer par des mots clés un peu plus précis d’auteurs ou d’autres, plutôt auteurs pour essayer de trouver. Donc ça, ça va être la première, le premier type de recherche.

Fanny : En passant par la recherche avancée du coup ? Avec le titre ?

A. B. : En recherche avancée, oui, ça m’arrive et après -le plus souvent c’est vrai- je vais aller voir dans le catalogue si la BnF possède vraiment ce document parce que pour être sûre de le trouver après. Après, j’ai aussi besoin d’illustrations. Donc là c’est pareil, je vais chercher par mots clés dans le moteur de recherche. Et pour le coup, moi je vais passer du temps à feuilleter les documents pour essayer de trouver la meilleure image. Je me note quand même l’astuce Google Image, je ne devrais peut-être pas ? [Rire] Là où je me suis retrouvée bloquée parfois, les documents en arabe, en langue arabe et la recherche, ça ne marche pas. On m’a aussi fait remarquer que la troncature ne marchait pas sur le moteur de recherche Gallica. J’ai testé ce matin, c’est toujours le cas a priori. Voilà. Après moi je travaille sur plusieurs langues donc ça peut être intéressant de voir avec l’ancien anglais avec “thy” au lieu de “the”, ce genre de choses si elles sont prises en compte ou non. Pour l’anglais en général je n’ai pas trop de problème,

mais dans les autres langues oui. Et comme on a pas mal de collections en langues étrangères sur Gallica, c'est important d'y penser. Oui. Et après nous on va aussi beaucoup utiliser mais là c'est plus dans la recherche, dans la recherche avancée par document, par partenaire en fait pour aller voir le document de notre partenaire, soit moissonné, soit intégré dans Gallica.

Arnaud : Qu'est-ce qui fonctionne bien et qu'est-ce qui fonctionne mal par rapport aux différentes situations de recherche que tu as ? Quelles difficultés tu peux avoir ?

A. B. : La principale difficulté que j'ai pu rencontrer, c'était par rapport à la langue arabe qui n'était pas du tout chercheable et on a dû faire appel à quelqu'un pour retrouver autrement les documents. Après dans la facilité, moi généralement je n'ai pas trop de problème à trouver un document quand on cherche le début du titre ou avec les guillemets pour préciser ma recherche, ça fonctionne assez bien.

Fanny : Parce que tu sais ce que tu recherches.

A. B. : Oui, quand je sais ce que je recherche, après quand je ne sais pas vraiment ce que je veux chercher, du coup, je vais y aller par propositions, regarder entre les dates d'édition de telle date à telle date, dans telle langue ou tel pays. Et dans ce cas, comme je n'ai pas d'idée précise de ce que je recherche, je suis toujours contente de ce que j'ai trouvé en fait.

Fanny : Ok.

Arnaud : Et au niveau de ton département ? Tu as des retours ?

A. B. : Euh non. Alors mes collègues m'avaient fait des retours, mais ils ne m'ont pas renvoyé, donc je peux faire un mail après avec leurs retours. Mais on va aussi chercher des beaux documents pour illustrer, par exemple Marion quand elle fait de la com.

Fanny : Est ce qu'il y a une autre personne qui veut parler ?

O. F. : Moi, par principe de proximité ? [Rires] J'ai des expériences assez diverses, mais disons que mon expérience principale, c'est avec le lecteur que je dois guider, à qui je dois présenter Gallica. Et donc il y a des situations extrêmement diverses. Une première catégorisation très rapide avant l'étape suivante. Il y a ceux qui sont complètement perdus. Je pense que dans Gallica, je pense que s'ils sont complètement perdus, c'est parce qu'il y a un défaut de -disons pour aller très vite-... de culture numérique. C'est-à-dire qu'ils ont du mal à se figurer le rapport entre ce soit la recherche ou la recherche avancée, a fortiori entre ces formulaires et ce qu'ils sont susceptibles d'attendre. Donc il y a cette difficulté là que je vois régulièrement dans les

ateliers les plus généraux. Ensuite, quand je les ai dans le public universitaire, la recherche est tout à fait différente. Ils sont souvent déjà un peu familiers de Gallica, mais leur demande c'est "comment faire une recherche systématique sur un sujet ?" qui est en général le sujet de leur mémoire de master ou de leur thèse. Et là, la demande est plutôt une demande de méthode. Parce que l'inquiétude, c'est "est-ce que je ne vais pas passer à côté de quelque chose ? J'arrive, je trouve plein de choses, mais comment être sûr que ma recherche systématique ?" Et j'ajoute pour une troisième expérience, c'est mon expérience personnelle, parce que je fais de temps en temps des billets de blog. Et là, mon problème, c'est plutôt "comment arriver à faire rapidement les billets ?" Parce que si on fait vraiment une enquête, il y a deux façons de faire un billet en soit, on a un sujet qu'on a prédéfini, puis on essaye de l'illustrer avec ce qu'on va trouver dans Gallica, c'est le cas le plus facile. Soit on fait une hypothèse de sujet et on commence à faire une enquête, on se rend compte que non, on ne trouve pas de ressources intéressantes, etc. Et donc ça, ça demande un travail très long d'interrogation pour voir si le sujet dont on a fait l'hypothèse a des retours intéressants ou pas dans Gallica, voilà. Et je complète par ça, avant de laisser la parole aux autres. Il y a évidemment deux types de recherche. Par exemple, chez les généalogistes dont je parlais tout à l'heure, ils font la recherche du type enquête policière parce qu'ils ont le nom d'un grand père, d'un arrière-grand-père. Et souvent ils sont à la recherche de l'apparition dans un journal de presse locale. Parce que je ne sais pas s'il a gagné un concours dans une fête locale ou ce genre de chose. Donc là ils cherchent des informations qui sont rares. Et puis l'autre c'est évidemment des recherches sur un... Ça rejoint ce que je disais avant. Des recherches sur un sujet et sur un thème "comment on fait pour ne pas être noyé dans les réponses ?". Il y a ceux qui sont en enquête policière parce qu'ils trouvent très peu et recherchent un peu une aiguille dans une botte de foin. Et puis il y a ceux, au contraire, qui sont noyés par les réponses et submergés : ça les angoisse, ça les inquiète, ça les décourage. Voilà pour un premier panorama.

Fanny : Ok, donc c'est plutôt... Cette masse de résultats est plutôt négative sur son site.

O. F. : Oui dans ce sens-là c'est "comment on fait pour réduire ?" Mais si on réduit le résultat pour que ça soit traitable, je laisse de côté tout ce qui est possible de faire du côté du Datalab. Comment réduire de manière qui ne soit pas arbitraire, c'est à dire que conserver ? Si vous avez 10 000 résultats, vous n'allez pas lire les 10 000 choses. Mais comment on fait pour réduire donc que ce soit pertinent ? En rajoutant un critère, etc. Et ça c'est une question très courante chez les doctorants, ceux qui essaient de se constituer comme une espèce de petit corpus provisoire indépendamment encore du Datalab car c'est encore une autre perspective.

Arnaud : Et du coup, quelles stratégies de contournement de recherche ?

O. F. : Alors je pense qu'il y a plusieurs aspects directement concernant Gallica. Moi j'utilise énormément la recherche par proximité et dans les critères, outre les critères classiques, titre, etc. J'attire leur attention sur le critère table des matières qui est très peu utilisé à tort parce que pour les livres -autant pour les périodiques, ça n'a pas l'air d'être très pertinent- autant pour les livres, c'est quand même assez efficace. Et après il y a un travail de réflexion qui est plutôt dans le dialogue entre le sujet et la réflexion sur le sujet et Gallica qui est la réflexion sur les concepts qu'on utilise. Il faut faire comprendre aux étudiants, aux chercheurs, etc qu'ils ne vont pas s'en tenir nécessairement strictement à ces termes. Mais c'est l'occasion, le dialogue, l'espèce de dialogue qui s'instaure à Gallica, c'est l'occasion pour réfléchir à ces concepts, pour les modifier, pour les faire varier. Donc ça, c'est les principales stratégies.

Arnaud : Et tu passes par le catalogue général ?

O. F. : Euh. Des fois oui, des fois mais pas systématiquement, mais sinon on passe à côté de beaucoup de choses. Par contre, je ne passe pas par Google, donc là ça m'ouvre des perspectives. [Rires] Je n'y ai pas pensé même si ça m'arrive de faire des articles...

C. A. : Désolée ! [Rires]

O. F. : Ah non non pas du tout, c'est pour ça, ça m'a ouvert des perspectives, ce n'était pas ironique parce que je fais des recherches d'images. La dominante, c'est quand même la recherche texte, mais en l'occurrence, on a peu d'iconographie et je sais que ce n'est pas facile parce que quand je fais un billet de blog, il faut que je cherche des images et je passe un temps fou, je trouve pour retrouver mes images. Donc je pense que je n'ai pas nécessairement la bonne stratégie. Donc voilà, je suis preneur de la méthode de ma voisine.

Matthieu : Si vous voulez, il n'y a aucun problème car on passe énormément de temps à donner à manger à Google pour qu'on soit référencés donc que les gens passent par Google pour aller sur Gallica ce n'est pas un problème. [Rires]

Fanny : Tant qu'on arrive sur Gallica. Est-ce qu'il y a une autre personne qui veut prendre la suite ?

XX : On peut faire dans l'ordre hein [rire]

V. P. : Moi, je passe beaucoup de temps à chercher des choses que je n'ai pas identifiées. Enfin, je fais des recherches assez vastes. Principalement de l'image, mais quand même beaucoup de

livres aussi puisque je vais chercher des illustrations dans des ouvrages scientifiques ou toute sorte d'ouvrages. Ce sont des recherches thématiques. Je prends l'exemple, par exemple, de ce qu'on fait avec la revue Relief. On fait de la papeterie en lien avec la revue. On fait trois gammes par an de papeterie avec eux, avec des thématiques très précises. Mais je travaille sur les deltas, les rivières, les sources par exemple. Donc je vais me retrouver à faire une recherche vraiment très large et je vais essayer quand même d'aller chercher aussi des choses qui ne sont pas numérisées. Donc aussi pour ça, je vais beaucoup travailler dans le catalogue. Et si je vais dans le catalogue, j'avoue aussi que c'est parce que la recherche thématique est plus facile. J'utilise les sujets. J'ai des rebonds sur les sujets, j'utilise des racines de cotes aussi par rapport à certains départements. Quand je sais qu'une racine de cote correspond à un sujet que je cherche. Par exemple, pour les estampes dans le guide Bouchot, je me repère comme ça. C'est vrai que sur Gallica, toutes ces fonctionnalités-là de recherche thématique sont plus compliquées... moins évidentes. Et il manque aussi pour moi la recherche par localisation puisque si je veux chercher des beaux livres 20^e avec des illustrations dedans sur des sujets précis, je vais chercher les livres de la réserve par exemple, ou le livre rare, et ça je ne peux pas l'affiner facilement sur Gallica ou quand je le fais sur Gallica, peut être que je ne fais pas bien, mais je n'ai pas trouvé le champ localisation vraiment. Je pense qu'on a souhaité gommer un peu la logique de département peut-être ? Je ne sais pas, je me trompe peut-être, mais donc je vais donc dans la recherche avancée et je vais aller taper entre guillemets par exemple "réserve" et "livres rares" pour voir les nouveautés, pour voir... Mais dans le catalogue, j'ai la fonctionnalité, de même que dans le catalogue si je passe par la recherche par type de documents, avec l'iconographie notamment, je vais avoir des effets de ... D'indexation qui est plus proche d'une indexation de livres en banque d'images plus proche de l'indexation de... (Je connais ça par cœur, mais j'ai tout oublié) du Thésaurus Garnier ! Qui est implémenté dans le catalogue général, alors que dans Gallica, j'ai une recherche Dewey, enfin d'indexation de bibliothèque. Donc le catalogue m'offre des fonctionnalités de recherche que je n'ai pas forcément dans Gallica. Et en même temps, c'est beaucoup plus amusant de chercher dans Gallica si on a du temps. Gallica est un outil formidable que j'adore. Donc j'y passe quand même énormément de temps. Je manque de temps parce que je suis toute seule sur mon activité et que je fais autre chose. Donc c'est vrai que je suis à l'affût de stratégies pour aller plus vite, mais en même temps, l'outil Gallica est super convivial et je n'ai même pas le temps de m'approprier pour l'instant tout ce que vous avez développé pour la mémorisation des recherches Gallica... (le nom m'échappe) mais Gallica...

Fanny : Rama ?

V. P. : Pardon ?

Arnaud : Gallicarama ?

V. P. : Gallicarama ! Carrément. Je ne m'en sers pas.

Arnaud : Ça vient d'arriver, c'est une nouvelle fonctionnalité

V. P. : Mais même s'il y avait avant ! Enfin, je ne connais pas toutes les fonctionnalités de panier... J'utilise très peu ça.

Arnaud : L'espace personnel.

V. P. : Voilà. Il faudrait que je le fasse parce que je gagnerais beaucoup de temps. Je pense que.

Fanny : Et est-ce que... Donc ok pour la recherche, mais en termes d'interface, pour visualiser un document par exemple une image, tu restes sur le catalogue général ou quand tu vas trouver le document qui te semble intéressant tu bascules sur Gallica ?

V. P. : Je bascule forcément sur Gallica, en fait je passe d'abord et par le catalogue et par la banque d'images et par différents outils Mandragore ou BAM. Pour avoir une idée un peu globale de ce que je vais pouvoir trouver. Et après, si je reste sur ce qui est numérisé, je vais essayer de travailler majoritairement dans Gallica. Autre souci que j'ai dans Gallica, c'est que je ne veux pas partir sur des pistes de documents que je ne nous ne détenons pas puisque je vais vendre la marque BnF et des images BnF. Et donc là, parfois, je suis en difficulté parce que même en choisissant Gallica, il y a encore des choses qui se sont greffées et qui ne sont pas des documents BnF. Et après, j'apprécie quand même d'avoir aussi des possibilités de recherche sur plein d'autres fonds, parce que ça me donne aussi des tas d'opportunités de ce que je vais retrouver dans le catalogue, des choses qui ne sont pas numérisées et qui proviennent d'une autre bibliothèque. En fait, il faut passer du temps pour que ce soit le plus... Le plus... que la récolte soit la plus ... la plus vaste et la plus pertinente. Mais tous les outils se complètent, simplement il me manque quelques clés de recherche, en recherche avancée dans Gallica qui me permettraient d'utiliser moins le catalogue afin d'avoir moins besoin du catalogue. En gros, c'est ça. Je ne sais pas si c'est utile ?

Fanny [acquiesce]

R. T. : Alors moi sur Gallica, je cherche un peu toutes les typologies de documents manuscrits, presse, c'est assez divers. J'utilise beaucoup le formulaire de recherche avancée puisqu'en fait, pour retrouver mes petits jeux, j'utilise beaucoup de champs de recherche différents. Alors avec

cette petite particularité que souvent moi je sais ce que je recherche et je connais déjà les cotes etc. Donc les ensembles de côtes. Donc je peux filtrer comme ça. J'utilise aussi beaucoup la fonction de date de mise en ligne parce que comme je fais des recherches sur des ensembles numérisés sur le très long cours, ça me permet de vraiment cibler les ensembles les plus récents. Par exemple. C'est ce que je recherche sur l'interface de Gallica, mais moi je couple beaucoup les facettes en fait. Moi pour retrouver vraiment ce que je cherche. J'ai quand même souvent dans mes recherches beaucoup de bruit. Alors là, qui est dû, je pense à la fois à la présence de nos documents sur les deux catalogues, ce qui fait qu'on peut avoir des documents qui, quand on fait une recherche de termes qui sont le catalogue général qui permet vraiment de circonscrire les recherches et qui une fois sur BAM, font que la recherche est trois fois plus importante ou au contraire, on sait qu'il y a plus et en fait on trouve moins. Donc souvent, j'ai la nécessité de combiner ces facettes et aussi la fonctionnalité de recherche d'un terme dans les résultats de recherche. C'est quelque chose que j'utilise beaucoup parce que ça me permettait de cibler les photos que sur Molière par exemple, alors que je vais avoir du bruit parmi d'autres.

Fanny : Est-ce que le bruit est toujours négatif ou pas ? Ou est-ce que parfois il peut être positif ?

R. T. : Non, parfois il peut être positif. Mais c'est vrai qu'il y a certains documents qui se retrouvent systématiquement dans mes requêtes et donc vraiment pour moi...

Fanny : Et qui sont systématiquement hors sujet ?

R. T. : Oui. Ça arrive parce que y'en a qu'on connaît un peu tous et qui reviennent tout le temps sur certains termes. Donc c'est vrai que cette recherche par terme dans les résultats de recherche me permet vraiment de mieux cibler et que je l'utilise beaucoup.

Arnaud : Sur le bruit. Est-ce que le bruit c'est parce qu'il y a trop de résultats ou il y a une incohérence dans l'ordonnement des résultats ?

R. T. : C'est plus une incohérence. Parce que en fait, en cumulant les champs de recherche, on arrive vraiment à cibler quand même, dans le formulaire de recherche avancée. On peut vraiment faire des requêtes très très précises, ce qui est très utile pour les sélections notamment. Mais, il y a certains termes qui génèrent automatiquement du bruit et pour lequel il faut vraiment ruser. Pour les résultats dans la requête. Donc voilà le formulaire de recherche avancée, pour moi il est très utile, mais c'est vrai que je pense qu'il devrait... C'est possible qu'on l'utilise dans ses pleines potentialités que si on connaît très bien ce qu'on cherche déjà en fait. Est-ce que

peut-être que, peut-être que quand on sait un peu moins, ça génère trop de résultats. Mais c'est vrai que quand on sait bien ce qu'on cherche, moi je trouve que c'est très très utile. Pour la localisation, bah je l'utilise en utilisant notamment "source", etc. Alors là... je pense qui n'est pas forcément dû qu'à Gallica même. Mais peut être aussi à la manière dont Gallica moissonne les données dans les catalogues. C'est qu'on va avoir des documents qui vont être regroupés sous des typologies documentaires qui ne sont en fait pas celles auxquelles on penserait quand on chercherait les documents. Je pense par exemple aux costumes. Qu'on ne va pas pouvoir retrouver en fait dans les facettes, facilement, ou même à des recueils de coupures de presse, etc. Qui vont être considérés comme des manuscrits ou même des photographies qui vont être considérées comme des manuscrits. Donc, je pense que le lecteur, s'il ne sait pas que sur cette collection-là, il y a cette particularité-là. Voilà. Mais bon, ça c'est, je pense, pas propre à Gallica même c'est la façon dont Gallica va moissonner dans les catalogues qui se pose particulièrement quand on a des typologies documentaires qui sont sur les deux catalogues en même temps.

Arnaud : Sur ce point, il y a des spécificités propres aux collections des arts du spectacle qui.. Ce qui pose des difficultés pour l'identifier comme tel fond ? Quelles sont les spécificités propres à tes collections R. T. ?

R. T. : Je pense qu'il y a la manière dont on a pu cataloguer en fait, à la base des ensembles iconographiques, des photographies par exemple, ou les ou les costumes qui sont pas... pour lesquels on a un peu changé aussi au cours du temps. Il faudrait faire, je pense, un travail d'harmonisation. Et puis il y a la fonctionnalité par spectacle aussi qui fait qu'on fait on a des notices de spectacle, du coup, on fait partie du spectacle à laquelle on rattache toutes les typologies documentaires qu'on connaît et qui des fois, sur Gallica, on va avoir plusieurs niveaux au moment de la numérisation, par exemple pour des affiches autour d'un spectacle, d'un festival, on va avoir qu'un résultat et on va se dire où sont toutes les affiches, on va voir qu'il faut vraiment cliquer dessus, choisir l'édition, etc. Et ça, ce n'est pas je pense aussi très très clair, même pour les collègues en fait qui des fois tapent directement le spectacle et trouvent qu'un recueil, mais le spectacle n'apparaît pas au niveau racine, donc se disent "ah c'est pas la bonne réponse, c'est pas le bon résultat". Et en fait non, c'est parce qu'il fallait cliquer, chercher l'édition, etc. Donc il y a aussi, je pense, ce niveau de détail, qui parfois n'est pas très bien rendu par Gallica.

Fanny : Et est-ce que tu fais du service public ?

R. T. : Oui.

Fanny : Pour les lecteurs qui viennent consulter les arts du spectacle, est-ce qu'il y a vraiment une particularité ? À quoi sont-ils... sont-ils confrontés en fait ? Qu'est-ce que tu vois comme retours ? S'il y en a.

R. T. : Mais alors non, en réalité, pas tant que ça. Parce qu'en général, nous on a beaucoup de documents qui sont sur Gallica Intramuros quand même, il faut le rappeler. La majorité en fait. Donc souvent, les lecteurs apprennent que les documents sont sur intramuros parce qu'on leur a refusé la communication de l'original et on leur dit de venir en salle de lecture et de nous demander. Voilà. Donc en fait, les lecteurs font rarement finalement leurs recherches sur Gallica, en tout cas en salle directement. C'est souvent suggéré par la personne qui est en banque de salle et qui va les aider à chercher. Après, quand ils savent faire une fois, quand on leur a montré une fois en général, ils arrivent à retrouver le document.

XX : Je voudrais juste appuyer le fait sur cette recherche. En effet, que de manière parfois quasi systématique, on va avoir le même document qui va réapparaître aux premiers résultats de recherche et après avoir ceux qui sont vraiment pertinents qu'on cherche.

Fanny : Oui. Est-ce qu'il y a une personne qui veut prendre la suite ?

A. V. : Alors moi, je vais utiliser Gallica dans deux situations : pour moi même quand j'ai des demandes, par exemple des réseaux sociaux, de documents pour un événement, une exposition à venir et ensuite pour présenter les collections aux chercheurs et aux donateurs. Donc dans le premier cas, quand c'est moi qui recherche, je vais passer par le catalogue général. Il faut dire aussi que nous à SVM, par rapport aux autres départements, on n'a vraiment pas du tout les mêmes volumétries qui sont sur Gallica pour des raisons de droits et spécialement pour moi pour la vidéo. Donc en fait, c'est beaucoup plus simple pour moi de passer par le catalogue général et de chercher "images animées" et de faire un clic localisation Gallica pour voir si j'ai des choses à mettre en avant ou si on fait des documents sur place que de faire la recherche directement dans Gallica. Donc ça c'est avec ma stratégie à moi. En revanche, quand je vais présenter aux étudiants, quand on a des présentations de collections, il va y avoir deux choses. Je vais quand même utiliser le moteur de recherche simple en utilisant le petit filtre vidéo à côté du moteur de recherche. Et ça, c'est vraiment très important pour moi, les types de présentation, je le fais parce que je vais le faire face souvent à des étudiants en cinéma et un des enjeux, c'est quand on a des licences ou des masters qui sont en train de chercher des sujets de mémoire, d'identifier des corpus sur lesquels ils vont travailler. En fait, nous en vidéo, c'est un peu plus de 2000 documents sur Gallica par rapport aux millions d'autres documents. Et donc

je vais leur dire bah “vous mettez le petit filtre vidéo et vous tapez votre... soit, votre sujet, etc. Et vous voyez tout de suite qu'on a en matière d'images animées.” Donc c'est une approche par typologie de documents qui est très importante. Après, je vais passer par les corpus, évidemment, sur les pages sélection, on a des corpus par cinéastes qui ont fait des dons, et donc ça leur permet, eux, d'identifier des ensembles. Parce que notamment, si je fais une recherche par cinéaste, même en mettant le filtre vidéo dans le moteur de recherche, une des difficultés, c'est qu'on va avoir des films qui vont être mélangés avec des conférences de la BnF auxquelles la personne a pu participer. Donc on a parfois un ratio qui n'est vraiment pas très bon entre un ou deux films sur ou sur un sujet, et puis plein de conférences qui vont être pas du tout le type de document qu'on cherchait.

Fanny : Parce que la personne est invitée c'est ça ?

A. V. : Ouais.

Fanny : Et donc du coup ça la recherche avancée permettrait de mettre en hauteur le cinéaste...

A. V. : Ouais mais moi je ne l'utilise pas la recherche avancée, je passe par le catalogue général dans ce cas-là.

Arnaud : Le fait justement de présenter des conférences qui sont associées aux films, c'est c'est plutôt du bruit, enfin c'est plutôt négatif ou au contraire, ça montre finalement qu'il y a une diversité de types de documents ?

A. V. : À ces moments-là c'est plutôt du bruit en fait parce que ce n'est pas le même type d'images animées, ce n'est pas le même type de document. Donc s'ils veulent se documenter sur un sujet de manière générale et qu'il y a des textes impR. T.és, il y a des images, il y a des conférences, très bien. Mais là, quand je suis face à des étudiants en étude cinématographique, ce n'est pas une conférence à laquelle la personne participe qui est intéressante. Évidemment, ça va dépendre, mais là, dans les cas dont je parle, c'est plutôt un problème. Et après, pour les donateurs, bah c'est extrêmement valorisant d'avoir une entrée sélection avec leur nom. Et derrière il y a une requête qui leur permet d'afficher directement l'ensemble des films qu'ils nous ont... pour lesquels il y a une cession de droits. Et je sais qu'ils partagent beaucoup des liens de requête auprès de leurs pairs, de leur réseau ne vont pas, eux, faire une recherche sur un article que j'ai donné, mais ils vont avoir tout de suite cette entrée qui est toute prête. Donc ça ça leur plaît bien. Comme R. T., je mets beaucoup en avant l'entrée presse de cinéma dans les formations parce que la presse, c'est vraiment une des difficultés récurrentes des usagers, y

compris en salle. Et souvent c'est très clair pour eux d'avoir ce corpus près du cinéma, avec en plus une couverture. Ils sont très contents après de pouvoir chercher dedans. Et moi à titre personnel, mais c'est plus tellement une question de recherche. Quand j'ai besoin de visionner des films, je vais systématiquement les télécharger à partir de l'archive numérique plutôt que sur Gallica, parce que ce que ça rame. Euh voilà. Mais c'est vrai que nous à SVM, c'est vraiment cet enjeu de volumétrie aussi qui nous force à avoir vraiment une approche par type de document pour mettre en avant nos collections.

Fanny : Est-ce que quelqu'un veut prendre la suite ?

A. M. : On a un côté ressources documentaires sur la DPV, éducation artistique et culturelle. Nous, pour nos métiers, médiateurs culturels, chargés de projets et même encadrement, nous allons utiliser Gallica plutôt par entrée thématique. C'est-à-dire qu'on va illustrer des supports Powerpoint à présenter aux élèves et aux profs. Ça peut être aussi une illustration dans le cadre de journées en région pour présenter par exemple des ressources. Donc là, je sais que, en termes d'encadrement, il va y avoir des entrées et des recherches par aire géographique sur Gallica. On va travailler aussi par typologie de documents, types de documents. Type de document, aire géographique et on va travailler pour tout ce qui est nourrir les ateliers pour les élèves. On va travailler sur l'index, l'index Gallica avec l'abécédaire qui démarre à A, on va, on va s'amuser à regarder. Est ce qu'on peut trouver des choses et des du corpus qui sont créés ? Donc on va passer par ça. Après, dans le cadre de des ateliers. Moi, en tant que médiatrice chargée de projet, je vais montrer des documents qui sont liés à des ateliers thématiques. Donc ça peut être des manuscrits, je vais entrer, par exemple, dans le cœur du document. Je vais leur montrer par exemple la grammaire de Champollion si c'est un atelier hiéroglyphes. Et puis on va parcourir. Je vais vraiment, en plus de l'atelier, sortir de mon Powerpoint et passer sur Gallica en complément du Powerpoint pour vraiment montrer comment est faite la page. Et là, je vais leur montrer le document en format mosaïque. Je vais utiliser la loupe parce que par exemple, dans le corps difforme de Montchenu qui est un manuscrit en forme de cœur, ce qui est très intéressant pour les élèves qui vont découvrir la forme des livres, c'est de rentrer dans le cœur du ... sans faire de jeu de mots... dans le cœur de l'ouvrage. Et là, on voit des erreurs, presque des moines copistes. On va voir des erreurs de réglure, on va rentrer dans une, dans un tampon, dans une estampille, dans la couleur de l'enluminure. Donc là, vraiment, on rentre. Donc la recherche a pu être parce qu'on connaît le document ou sinon ça peut être en général, je ne vais pas m'improviser une recherche en direct parce qu'effectivement il faut paramétrer le temps. Donc je vais faire une recherche en amont. En général, quand on fait du live, en fait devant le

public, on prépare vraiment structurellement notre recherche, donc on va chercher en amont. Et je vais utiliser donc la recherche simple ou la recherche avancée, et après une rubrique qui est très intéressante pour nous en termes de véracité de l'info, c'est la rubrique "En savoir plus" sur le bandeau de gauche. Parce que là, nous on fait un écho immédiat avec les fausses nouvelles sur internet, c'est-à-dire qu'on va valoriser le document et je vais montrer cette rubrique "En savoir plus" et notamment la cote. Et cette notice numérique, en fait, elle est très importante pour montrer le titre du document, l'auteur et la source. Et en fait pour le coup effectivement un quand les élèves plus que les étudiants vont faire des recherches sur Internet. Ils peuvent très bien ne pas sourcer les sources. Ils peuvent très bien aller prendre une information sans la vérifier. Donc nous on a un travail d'illustration, mais on a un travail de montrer la véracité et la fiabilité de la cotation. Et c'est pour ça qu'on parle de bibliothèque de référence. Ça c'est hyper important nous vis à vis des élèves et des profs.

Fanny : Et justement, cette partie "En savoir plus." Sur comment c'est vécu ?

A. M. : C'est une découverte !

Fanny : Parce qu'en fait il y a beaucoup d'informations très riches.

A. M. : Bien sûr, elle est très très dense cette notice. Donc moi je montre la rubrique "En savoir plus" sur le côté. Ça c'est ce que je montre où ça se trouve et après je ne vais pas tout, je ne vais pas rentrer dans développer la notice qui est beaucoup trop longue. Je montre le titre, l'auteur, un auteur, pas forcément les autres. Et puis je bascule directement sur la source parce que sinon c'est trop dense. Mais en tout cas, c'est quelque chose qui est important pour eux parce qu'ils ne savent pas forcément la date du document. Et pour des profs et des élèves, c'est hyper important.

Arnaud : Est-ce que tu fais des rebonds sur les documents associés dans cette colonne justement, où se trouve la rubrique "En savoir plus" ?

A. M. : Je ne vais pas faire de rebond immédiatement lors d'un atelier parce que sinon, bah on va partir sur d'autres sujets. Et puis moi j'ai mon le timing. Mais effectivement les rebonds sont importants en termes d'arborescence pour aller sur d'autres choses.

Arnaud : Tu l'utilises quand tu fais tes recherches ?

A. M. : Ah pour moi oui.

Arnaud : C'est une fonctionnalité qui te plaît ?

A. M. : Ah tout à fait, tout à fait, tout à fait. Mais vraiment, la loupe est importante. Après l'enregistrement, l'image aussi parce que l'image doit être téléchargée pour être réutilisée ailleurs. Et là on insiste beaucoup sur le rappel des sources, etc. Ne pas délier une source d'image dont on travaille beaucoup sur la presse et les fausses nouvelles, ça c'est quelque chose qui est important, même sur un document historique ou classique. Voilà. Donc recherche thématique, aire géographique, la recherche par type de documents, et puis, et puis de l'information et de la formation pour les profs et les élèves. Donc là je vais envoyer des liens, des liens par thématique. Mais moi l'idée c'est d'y coupler des tutos parce que ce n'est pas le tout de leur faire des recherches, il faut qu'après ils sont en autonomie pour trouver. Donc c'est là que je vais mettre des choses par lien sur le blog et puis sur les tutos que je trouve dans les tutos "tutoriels Gallica". Je fais des recherches, je leur envoie le numéro un, numéro deux, numéro trois. Parce que ça me semble simple pour eux, mais en fait ils connaissent Gallica de nom, ils ne sont pas sûrs que tout est libre de droit. Donc c'est là que je leur parle de Gallica Intramuros, mais ils ne sont pas sûrs que tout est libre de droit. Et moi je les invite quand je fais ma présentation ou même aux élèves. Je vous invite à faire vos travaux, vos exposés. Venez sur Gallica, tout est libre de droit. Vous avez cette rubrique.

Réactions diverses dans la salle : Non !

A. M. : Non mais tout est libre de droits à part le partenariat marque-blanc et tout ce qui vient des bibliothèques qui ne sont pas bibliothèques...

Réactions diverses : Non ! Même ! Il y a une cession de droits, ce n'est pas la même chose.

A. M. : Oui c'est ça.

Fanny : C'est un vrai problème qu'il faut éclaircir.

A. M. : Non effectivement quand je parle de ça, c'est par rapport à l'enseignant qui va l'utiliser dans le cas d'une ressource pédagogique en classe.

Fanny et autres : Il peut l'utiliser.

A. M. : Pardon, mais merci de faire la différence.

XX : Faut pas dire ça aux lecteurs.

A. M. : Voilà. Non mais là on est chez les petits donc y'a pas de problème. Mais si ça parle, je parle de ça à l'enseignant pour qu'elle, elle le réutilise dans le cadre d'un usage non-commercial, en classe, avec eux bien sûr. Et après on va parler aussi lors des visites de sites aux personnes

qui viennent visiter la bibliothèque et leur présenter Gallica aussi dans le cadre de la visite historique. Quand on est face à la maquette parce qu'on parle de volumétrie documentaire et du coup la requête de Gallica en plus du catalogue général de recherche.

Arnaud : Est-ce qu'il y a des difficultés particulières ? Quand on est face à un public comme ça qui ne connaît pas Gallica. Quelle est sa réaction ?

A. M. : C'est souvent des découvertes importantes sur les personnes qui viennent visiter la BnF qui sont des tous public. Ils n'arrivent pas à croire que depuis chez eux, à Clermont-Ferrand, où... ils puissent avoir accès à des ressources si précieuses. Du coup, on insiste là-dessus. Même pour le champ social, pour le champ social des personnes qui sont pR. T.o arrivantes ou qui sont en apprentissage du français. Elles peuvent rentrer par l'iconographie même de documents qui vont être verbalisés ou dont les notices sont simplement rédigées de manière à pouvoir s'approprier des notices simples. C'est des choses dont on parle. Quoi d'autre en termes de... J'ai d'autres témoignages de personnes et comme c'est assez dense, je vous les enverrai par mail plus tard.

Fanny : Ok, on essaie de finir.

Arnaud : En même temps c'est important de vous écouter. [Rires légers]

E. M. : Oui donc moi c'est plutôt des recherches assez classiques. Soit. On cherche des textes, soit j'ai une thématique très générale. J'ai besoin de me faire des revues de presse assez générales autour d'un événement ou d'une personnalité. Il y a la fonctionnalité du rapport de recherche que j'utilise pas mal pour la mise en contexte. Évidemment, ce qui manque dans la presse pour l'instant, c'est la segmentation par article. Pour faciliter cette revue de presse, c'est encore un niveau de complexité [rire]. Voilà. Soit un cas de figure différent où je suis à la recherche d'une information très précise. Alors je sais précisément ce que je recherche, soit parce que j'ai le titre du numéro. Je sais que c'est un quotidien du 18 janvier 1839 et lorsque je vais chercher vraiment dans le numéro et je feuillette et je retrouve assez facilement les informations mais ça nécessite d'avoir en amont des recherches, ce n'est pas Gallica qui va fournir la solution, je sais d'abord ce que je vais trouver. Là où Gallica m'est utile pour rechercher vraiment des informations précises c'est avec la recherche par proximité du terme. C'est sans doute la fonctionnalité que j'utilise le plus. Parfois, je tâtonne un peu sur le nombre de mots, l'espacement entre deux termes et il n'y a pas de règle absolue. Parfois, c'est une lettre de dix mots. Résultat, parfois il faut mettre 30... Donc je serai assez preneuse de stratégies pour un peu réduire ce tâtonnement à chaque fois aussi. Le filtrage par recherche, par typologie documentaire m'aide parfois, au

contraire il ne m'aide pas et il faut que je décoche Presse et revues et ça peut mieux faire remonter les résultats et aboutir à ce que je cherche. Ce que je trouve un peu difficile avec la recherche avancée, c'est l'absence de règle générale et c'est vraiment tâtonnant qu'on s'y retrouve.

Fanny : C'est-à-dire que quand Presse et revues est cochée, les résultats sont moins pertinents ?

E. M. : Ça arrive, de faire remonter les résultats moins pertinents que quand je décoche. En général, je commence toujours par chercher par typologie documentaire et après je relance la recherche en ayant décoché ça. Parfois, je peux trouver des résultats en plus. Ce qui me serait très utile dans cette fonctionnalité de recherche par proximité, ce serait de pouvoir associer plus de mots. C'est-à-dire par exemple je recherche les robes de Paul Poiret, mais précisément de Satin. J'aimerais bien pouvoir associer Paul Poiret, robe et satin et pouvoir comme ça faire plusieurs recherches par proximité combinées.

Arnaud : Effectivement, tu peux ajouter une recherche par proximité, mais ça ne fonctionne pas bien.

E. M. : Oui, c'est ça. Si cette fonctionnalité-là pouvait être affinée, ce serait...

Arnaud : Tu préfères avoir trois termes ?

E. M. : C'est ça.

Fanny à O. F. : C'est ce que vous disiez, vous utilisez beaucoup la recherche par proximité ?

O. F. : C'est ça.

Fanny : Est-ce que c'est quelque chose qui vous a manqué ?

O. F. : Euh oui, il y a une hésitation. Par exemple, est ce que -que ce soit à gauche ou à droite du nombre de mots- Est-ce qu'il est possible d'utiliser une expression parce que deux mots ne suffisent pas ? Donc apparemment ça marche, mais il y a un doute. Il y a quelquefois ça ne marche pas, ça ne marche pas. Enfin, il y a une incertitude sur ce point-là, parce qu'il ne suffit pas souvent de réunir deux mots. On a besoin de quelque chose d'un peu plus complexe. Alors soit, comme vous dites, on aimerait avoir une recherche par proximité, mais un peu complexe ça devrait associer trois mots, trois termes. Mais il y a aussi associer un mot, par exemple un nom d'auteur et une expression. Par exemple, je ne sais pas par exemple qu'il y a un étudiant "Rousseau et Amour de soi", des expressions assez caractéristiques et des fois et de bons retours. Des fois c'est un peu incertain.

Fanny : Donc à la fois l'expression et à la fois la possibilité de chercher des expressions et d'ajouter un troisième terme.

E. M. : En revanche, ça va très bien fonctionner. Par exemple, sur des personnalités du monde du théâtre, je cherche précisément un rôle qu'elle a interprété, je cherche le rôle et la comédienne et ça fonctionne très bien dans ce type de cas de figure. Le filtre par date également, je l'utilise beaucoup. Je parlais aussi du rapport de recherche que j'utilise notamment, alors ça, c'est plus une stratégie de contournement parce que parfois j'ai besoin d'avoir -si mettons que je fais des recherches sur une personnalité, mais qui a aussi été rédacteur ou rédactrice dans la presse. J'aimerais trouver les articles non pas qui parlent d'elle mais qu'elle a rédigés. La distinction sujet ou auteur ne fonctionne pas avec les filtres de la recherche avancée. En revanche, on peut trouver des stratégies de contournement. Typiquement, je pense à quelqu'un qui s'appelle Marie Marvin, je cherche les articles qu'elle a été amenée à écrire en tant que reporter de guerre et je trouve. Il y a aussi beaucoup et la plupart des articles qui parlent d'elle comme sujet. Donc dans ce cas-là, je mets Marvin, par mademoiselle Marvin parce que j'essaie plusieurs combinaisons et ça va me faire ... En regardant ensuite les résultats dans le rapport de recherche, de pouvoir voir si son nom est en signature ou en tête d'article. Et ça me permet de cibler assez vite un article qui parle d'elle mais qui reste très spécifique.

Fanny : Ça c'est très spécifique à la presse. Est-ce que la recherche par article ça fonctionne ? [Rires] Pour la coup la recherche par article ...

Matthieu : XX C'est compliqué.

E. M. : On est sur un niveau de granularité assez fin.

Fanny : Oui mais c'est intéressant.

E. M. : Je pense que c'est un cas d'usage assez fréquent que je cherche à faire. Voilà ensuite bon la recherche d'images, en particulier dans la presse, est très complexe. Là je n'ai pas encore développé de stratégie de contournement. C'est souvent un peu, par hasard, ce n'est souvent pas un besoin, sauf dans le cas d'illustration. Et c'est vrai que pour le filtre images dans la recherche avancée ne marche pas bien. En général, mais je vais essayer de croiser un peu les facettes comme j'ai un peu de chance. Ce n'est pas du tout une science exacte.

Matthieu : C'est le rapport de recherche qui marche le mieux ?

E. M. : Oui. Voilà, c'est principalement mes usages. Ensuite. Quand je suis, quand je fais du service public et que j'ai besoin d'aider les lecteurs qui en général viennent avec en tête un

journal précis qu'ils recherchent, je fais une recherche dans le catalogue général. Et puis en formation, je vais assez facilement montrer les sélections (la presse de spectacle, la presse de cinéma, ...). Je trouve que c'est vraiment un mode d'entrée qui est le plus simple pour ceux qui découvrent.

Fanny : Ok, est-ce que D. B. tu veux ... ?

D. B. : Oui, donc moi maintenant j'ai un usage, je pense beaucoup moins quotidien et assidu de Gallica que d'autres ici parce que les collections dont je m'occupe ne sont pas dedans, donc je l'utilise principalement en service public. Alors ça rejoint beaucoup des choses qu'on a déjà été dites faire des recherches très pointues sur les soulèvements indigènes entre telle et telle période. Et dans ce cas-là, la recherche par proximité aide énormément. Je passe beaucoup par les sélections de presse aussi et beaucoup quand on demande des titres de presse par le catalogue général parce que j'ai besoin de voir vite l'état des collections et je trouve que je les vois moins dans Gallica. Une des difficultés et limites quand les gens veulent une constitution de corpus systématique, c'est que je ne sais pas qu'est-ce qui est indexé en plein texte ou pas. Mais bon, ça je ne sais pas s'il y a vraiment de solution à part connaître la politique de numérisation de la BnF. On ne sait jamais si les résultats c'est... à quel niveau de granularité on est, si c'est égal et je ne sais pas toujours répondre à ce type de questions. Ensuite, au niveau du service -alors je ne sais pas si Gallica et Gallica intramuros c'est le même moteur de recherche ou c'est deux... ?

Fanny [acquiesce]

D. B. : Oui, c'est le même moteur de recherche, donc je peux faire aussi des réflexions sur un Gallica Intramuros. Ah oui, juste parce que ça n'a pas été dit. Je trouve l'auto-complétion et le ... parfois c'est un peu long, ça rame, mais je trouve ça très utile. Et aussi, je ne sais pas, les recommandations ou les suggestions qui sont parfois dans une espèce d'encart quand on fait une recherche, je trouve ça assez utile aussi et souvent c'est plus pertinent que les trois premiers résultats de recherche. Oui et au niveau de Gallica Intramuros. Donc nous, on a une problématique un peu spécifique dans le service, c'est que donc j'ai un collègue qui s'occupe de tout ce qui est dépôt légal des livres numériques et que pour l'instant, pour tout ce qui est donc ... (amené à prendre beaucoup d'ampleur) et pour l'instant en fait on a pas du tout les fonctionnalités bibliothèque numérique, on ne peut pas faire une recherche plein texte ni sur l'ensemble des collections, ni dans un livre. Et en fait on est plutôt sur de la page par page ou

au chapitre quand vous faites du contrôle F, il y a en fait... on est, on est, on est bloqués par... on a, les livres y sont pas, on a pas de texte ou on indexe pas vraiment le contenu en fait.

Matthieu : On indexe mais XX

D. B. : Oui oui, tout comme. Oui, on l'indexe, d'accord, parce que je ne savais même pas quand il m'a montré ce qu'il faisait dans les Epub, on en fait on un...

Matthieu [interagit]

D. B. : Donc en fait on ne s'en sert pas vraiment.

Matthieu : XX

D. B. : Voilà surligné. Donc en fait on est dans des fonctionnalités zéro de technique de consultation d'Epub. Voilà. Du coup, j'ai passé un peu de temps avec lui hier pour regarder. Donc je pense que beaucoup de choses sont liées à l'outil d'accès qui est *virtual browser* et où il y a peut être il y a très peu de fonctionnalités. Alors lui, quand il cherche, il y a cette question de : "comment je récupère, enfin comment je consulte l'intégralité des fonds, etc." Donc il passe par le petit "nés numériques". Euh voilà qui est un peu contradictoire avec le reste des champs, parce qu'il y a Epub, nés numériques. Enfin voilà, les catégories se recoupent partiellement dans ce qu'on va choisir. Il aimerait... et il aimerait pouvoir avoir C. A. ment un PDF et Epub seulement en nés numériques. Voilà. Voilà. Donc il y a cette grosse question de voilà, on va proposer un catalogue très riche d'ouvrages et en fait quelles fonctions de recherche on propose dessus et quelles fonctions de filtres ? Alors qu'il y a déjà énormément de choses dans Gallica de différentes périodes. Et en plus, est ce que voilà, du né numérique, ça va parler aux lecteurs ? Entre le numérisé et le né numérique. Donc ça c'est les préoccupations un peu...

Arnaud : Né numériques ou c'est que c'est un format ?

D. B. : Il est né numérique au sens où il n'est pas issu d'une numérisation qu'on va recevoir par dépôt légal numérique, c'est des documents qui sont... qu'on nous fournit sous forme numérique originellement. Mais effectivement, comme il y a des Epub dérivés de numérisation, c'est très compliqué. Donc voilà. Donc une question qui va encore ajouter, je pense au bruit et à la complexité de ce qu'on va trouver dans Gallica. Euh. Est-ce que Oui mais.

Fanny [lance un regard à A. T.]

A. T. : Oui, donc je termine, je me retrouve quand même dans un certain nombre de choses qui ont déjà été dites au cours de la matinée. [Rires] Alors, sachant que je suis plutôt sur des

manuscripts, il y a un certain nombre de spécificités, notamment entre ces questions d'OCR, de recherche à proximité, et ça ne se pose pas, mais c'est pas forcément évident pour le lecteur qui a l'habitude de... pour certains qui ont l'habitude de faire une recherche à proximité, ils comprennent pas que tu ne peux pas chercher à l'intérieur du manuscrit. En salle de lecture, en salle de lecture ou via Sindbad. Donc les gens qui viennent me voir sont des gens qui n'arrivent pas à faire leur recherche ou alors qui n'ont pas tout trouvé. Il y a un certain nombre de gens qui cherchent un document précis, donc ils connaissent la cote, mais ils vont taper dans la recherche d'un texte en fait mais ils ne vont pas trouver. Donc ça, c'est quelque chose qu'on voit effectivement beaucoup. Les gens n'ont pas l'habitude, n'ont pas le réflexe de passer par la recherche avancée avant qu'on leur dise, il faut leur dire deux, trois fois avant qu'ils le fassent régulièrement. Quand ils font des recherches et recherches par auteur, c'est un peu compliqué aussi parfois à cause de l'état du catalogue, mais aussi parce que, au moins dans les manuscrits ou dans les auteurs, et on ne les a pas. La recherche par date pose problème quand on est avant l'an 1000 et même un certain nombre de choses avant l'an mil, voire avant l'an 0. C'est compliqué aussi. Et si parfois cette question la cote et il y a aussi la question de trouver le département, alors ce n'est pas tous les lecteurs, mais effectivement, il y a un certain nombre de lecteurs qui cherchent des manuscrits de la BnF, voire du département des manuscrits de la BnF et on peut le trouver dans le champ source. Mais il faut savoir, il faut connaître le fonctionnement et ils ont tendance à cocher les manuscrits et à ne pas comprendre pourquoi ils retrouvent en premier ou pour la moitié des manuscrits partenaires qui ne sont pas des manuscrits de la BnF. Ce qui fait qu'en fait, dans trois cas sur quatre, je conseille de passer par le catalogue plutôt par BnF, par les archives et manuscrits, mais aussi par le catalogue général quand c'est sur des documents imprimés. Et pour le reste via le moteur de recherche Gallica je prépare des recherches, notamment pour des réponses Sindbad pour dire "voilà dans le cas qui vous intéresse. Vous pouvez... (comme une sélection) vous pouvez faire la recherche suivante et normalement elle devrait s'enrichir des nouvelles numérisations". Mais c'est une équation que j'ai construite en connaissant l'état des métadonnées et une chose qui est assez difficile je pense à faire quand on n'a pas idée de ce qu'il y a dans le catalogue, ce qui du catalogue va dans Gallica et commencer à dresser un état des lieux. Voilà pour les pour les recherches d'images, j'ai tendance à renvoyer vers Mandragore. C'est un peu prêcher pour ma paroisse aussi [rire], mais c'est vrai que pour l'instant, même en sachant qu'il y a dans Gallica des documents qui ne sont pas dans Mandragore, c'est vraiment très compliqué à trouver. A moins encore une fois de savoir exactement ce qu'on cherche et de passer par exemple par un catalogue papier. Et ensuite de poursuivre la recherche ?

C. A. : Moi aussi, quand je recherche un manuscrit en particulier. Je crois qu'il y a une expo dans le manuscrit. Je passe beaucoup plus souvent par Archives et manuscrits que par Gallica. Quand je sais duquel il s'agit. Et je crois que c'est parce que dans Archives et manuscrits t'as le résultat de recherche, tu vois la qualité de la numérisation tout de suite.

A. T. : Oui, alors ça, c'est aussi quelque chose que les lecteurs demandent.

C. A. : Ouais, et pour moi c'est super important aussi parce que je sais si je vais pouvoir les utiliser ou pas pour les réseaux sociaux, tu sais quand t'as les vieilles numérisations ?

A. T. : Oui. Pour certains usages, elles sont très utiles. Mais sinon, c'est aussi quelque chose qui gêne les lecteurs, notamment quand ils font une recherche par cote, ils arrivent à trouver, à trouver la cote. Il y a parfois plusieurs numérisations parce qu'il y a à la fois le microfilm, l'original, ...

C. A. : Exactement !

A. T. : Parfois plusieurs microfilms parce que c'est un trop gros manuscrit, alors là il fait deux bobines et parfois il va y avoir plusieurs originaux parce qu'on a la restauration, on l'a restauré. On se dit que quand même, maintenant qu'il est restauré ... et c'est vrai que c'est une demande d'arriver à pouvoir séparer les numérisations en noir et blanc des numérisations en couleur. Alors c'est pareil, on peut le faire mais quand on ne connaît pas... Dans le catalogue, il y a écrit numériser à partir du document original, non ? Si on me le demande à moi, je peux le faire, mais je ne pense pas que ça soit... (à moins de bien connaître le catalogue) que ça soit évident de...

Fanny : Et du coup C. A. tu dis que tu passes par Archives et Manuscrits ? Mais tu ne passes pas forcément par le catalogue général pour faire des recherches ?

C. A. : Bah, quand tu cherches un manuscrit, comment est-ce que je pourrai dire ça ? Un manuscrit... Du coup tu sais lequel tu es en train de chercher. Je trouve que ça va beaucoup plus vite pour voir si le manuscrit que par le catalogue général, je ne saurai pas dire pourquoi parce que souvent j'ai la cote en plus.

A. T. : Après archives et manuscrits c'est fait pour chercher des archives et manuscrits.

C. A. : Ouais, bah ouais.

A. T. : Ça correspond mieux.

C. A. : Ouais.

Fanny : Ok !

D. B. : Juste une petite remarque d'ordre général. Moi parfois j'hésite à utiliser Gallica à cause des temps de réponse, mais je sais qu'en ce moment le réseau est très lent. Mais c'est vrai qu'en ce moment quand je suis en salle et que ça, ça peut jouer dans l'équation, quand il faut faire une recherche...

O. F. : A midi surtout

D. B. : Même le soir, parfois.

Matthieu : XX [rires]

D. B. : Et je me souviens si le service reproduction avait fait de la pub avant Noël pour commander des illustrations, je me suis dit que c'est une super idée. Puis j'avais lancé une recherche et là ça mettait quinze secondes et j'avais abandonné [rire].

Fanny : Ok, bah merci à toutes et à tous. [Regard vers Arnaud]

Arnaud : Oui, il faut qu'on ajuste un petit peu [le déroulement de l'atelier].

D. B. : J'ai l'impression que c'est le réseau, c'est le réseau. Mais c'est vrai que parfois, c'est vraiment complètement incroyable de mettre 10 secondes à afficher une page mais c'était moins long avant non ?

XX : C'est par période.

Passage à l'étape suivante (les améliorations à apporter à Gallica).

Partie 3 : Les améliorations

Fanny : Il y a certaines de vos remarques qui sont très spécifiques. Et le public va avoir d'autres besoins. C'est pour ça qu'on a envie de savoir pour vous, quelles sont les évolutions qui pourraient être intéressantes, pour votre usage, ce que vous avez constaté des usages en tant que professionnels au sein de la BnF ? [...] Est-ce que A. T. et D. B., vous voulez commencer ?

D. B. : Oui, elles ne sont pas dans l'ordre, on ne les a...

Fanny : Ce n'est pas grave, allez-y.

A. T. : Oui il nous semble qu'il y a quelque chose à revoir sur la typologie des documents, à revoir ou à expliciter.

D. B. : A désambiguïser peut être parfois, parce que c'est vrai que pour image, on a l'impression qu'on cherche que dans les images. En fait, tout ce qui est, tout ce qui est catégories, qui se recourent ou... ??

A. T. Dans le même ordre d'idée, dans un formulaire où en facette, ça serait bien d'avoir des choses sur la qualité de la numérisation, que ce soit une numérisation noir et blanc, en couleur ou nativement numérique. Déjà entre ces trois-là, ça c'est des résultats de recherche très différents et qu'on va vouloir trier.

Fanny : Ok.

Arnaud : Et peut-être la qualité de l'OCR ?

A. T. : Et peut-être aussi mode texte ou pas mode texte ?

D. B. : Et sinon après on s'était dit que des fonctionnalités de recherche images spécifiques. Mais ça je crois que c'est c'est en projet. Mais mieux distinguer la recherche textuelle et avoir des fonctionnalités spécifiques pour la recherche d'images. Et éventuellement des suggestions de mots-clés proches en fait. En se fondant sur des répertoires d'autorités qu'on a déjà ou je ne sais pas mais ça peut être une aide à la reformulation.

Fanny : Ok.

D. B. : Et améliorer la recherche par proximité enfin, l'approfondir et permettre à plus de... des "ou" enfin de permettre de lancer sur plusieurs... plus de termes et aussi en incluant des ou et des et.

Fanny : Ok.

A. T. : Après, il y a toujours un point sur la pertinence de l'ordre des résultats mais bon après...

D. B. : Oui on s'est dit...

A. T. : Bah on n'a pas non plus de solution toute prête ...

Fanny : Mais donc il y a quand même un sujet sur la pertinence et l'ordre de présentation.

D. B. et A. T. [approuvent]

Arnaud : En termes d'affichage aussi ?

D. B. : Ouais. En fait je ne sais pas s'il y a possibilité de surpondérer certains titres star par exemple. Parce que je sais que si on passe beaucoup par les... les pages de sélection sur la presse etc, c'est qu'en fait ça ressort mal en premier quand on fait une recherche. Et donc je ne sais pas si ça existe dans certains moteurs de recherche d'avoir... qui apparaissent en encart ou quand... ça le fait un peu de temps en temps.

Fanny : Bah Arnaud vous expliquera tout à l'heure comment fonctionne le ... Enfin rapidement ce qu'il en a compris. L'ordre de présentation et oui peut-être que c'est une des choses qu'il faudrait améliorer ? En tout cas questionner. Ok, est-ce qu'on passe à R. T., V. P., A. B. ?

V. P. : Alors nous on avait mis les critères de pertinence, peut-être aussi en troisième position, mais c'est vrai que pour les résultats... pour une meilleure lisibilité des résultats, on s'était dit que ce serait bien de jouer là-dessus. Et ensuite, donc les deux premiers points, c'était peut-être (une des pistes d'amélioration possible) ce serait de rajouter des fonctionnalités de recherche dans Gallica qui soient peut-être plus en phase avec celles des catalogues. Notamment en rajoutant des critères de recherche, on a parlé par exemple de... une personnalité comme auteur, comme donateur... etc. Qui permettrait de moissonner plus d'informations, des catalogues et de ne plus passer par un catalogue général pour s'assurer que on a bien tel document, ce qui permettrait de réduire le bruit aussi. Et en dernier point, peut-être renommer certains critères de recherche pour les rendre plus accessibles parce que peut-être un peu jargonant pour le grand public. Voilà les idées (sources et localisation). Source et cote du document ça ne parle peut-être pas à tout le monde ?

Fanny : En vrai sachez que le grand public la recherche avancée ça ne lui parle pas du tout, elle est trop compliquée [rires].

Arnaud : C'est 3% des usages.

V. P. : Combien tu dis ?

Arnaud : 3% des usages.

V. P. : Ah oui d'accord ! [Rires]

[Chacun donne son commentaire].

Fanny : L'utilisation personnelle, voilà ce que c'est que du document, à part peut-être pas tout le monde. En bref, sachez que le grand public, la recherche.

V. P. : Mais pour cette raison-là, parfois je les perds mes partenaires. Ils se lancent dans Gallica, s'ils ne font que de la recherche simple, au bout de peu de temps ils sont ravis parce qu'effectivement on rebondit sur plein de choses, etc. Mais ils se rendent compte qu'ils ne pourront pas s'en sortir tout seul en recherche simple.

Arnaud : Juste le groupe 2, l'organisation et la priorisation c'était ça hein ?

A. B. : Non la 2 était en 1. On recherche juste pour le groupe deux. Le nommage était en 2e.

V. P. : A. T. a parlé de la couleur. C'est une de mes grandes attentes. Si un jour on pouvait dire trier.

Fanny : Ouiii... On vous en parle à la fin [rires]. C'est aussi une de mes attentes. On vous en parle à la fin, c'est une évolution qui est en cours.

V. P. : Ça c'est super.

Fanny : Dans les projets de recherche pas l'image mais c'est un autre mode de recherche et du coup dans des corpus au début restreints puis c'est des choses qui évolueront avec le temps.

V. P. : Oui parce qu'il y a des tas d'astuces hein, ça ne marche pas tout le temps et après on va dépouiller des tas de livres.

Fanny : Je fais un encart à la fin pour vous parler de Gallica Images et vous dire où on en est et ce à quoi ça pourrait ressembler. Est-ce que vous voulez prendre la suite ?

C. A. : On a plein de trucs nous. On n'a pas priorisé.

Fanny : C'est pas grave.

C. A. : Donc alors sur la question des filtres, on a deux morceaux : un morceau, l'aspect des filtres, on les trouve peu lisibles, peu visibles, trop petits. Et après on a un aspect contenu des filtres sur la priorité des filtres proposés. Remonter le filtre par recherche exacte et voilà, revoir la priorité des filtres qui sont proposés. (ça c'est fait). Ensuite. On a une partie interface où quand on arrive sur Gallica, on trouve que le moteur de recherche est noyé dans les autres valorisations. Voilà, vraiment le mettre en avant.

Fanny : Alors ? J'en parlerai, j'en parlerai à la fin. Du coup, un des grands projets. Il y a plusieurs projets en cours du côté de Gallica, il y a le projet sur le moteur de recherche effectivement, un projet de refonte de la page d'accueil : technique, graphique et ergonomique et le moteur de recherche, dans la première maquette, on ne peut pas le rater, il est très présent. Donc ça, c'est

ça, c'est en cours. Il y a, il y a aussi, on en a parlé, je le dis, il y a aussi une refonte qui est en cours et une réflexion sur la refonte enfin une nouvelle arborescence pour les pages sélection. Aujourd'hui les pages sections, elles sont très départementalisées et il y a des choses qui ne se recoupent pas forcément (par exemple la presse). Je prends un exemple qui est peut-être totalement faux, mais la presse de théâtre est peut être uniquement dans la presse et pouvait être aussi dans les pages du théâtre. On est en train de réfléchir à comment mettre ça à plat pour justement permettre des entrées beaucoup plus thématiques, pas sûrement uniquement thématiques dans les pages sélections et donc dans un résultat. Et puis supprimer certains étages aussi. Parce que, si vous avez déjà utilisé les pages sélections quelques fois, on peut parfois cliquer très longtemps avant d'arriver à un résultat. Bon après on ne pourra pas supprimer tous les niveaux non plus. Et donc j'y viens au dernier sujet (comme ça je clos) un des projets des trois années à venir pour Gallica, c'est la recherche par l'image, donc un autre mode de recherche là où aujourd'hui on cherche dans le texte et le plein texte. Là on pourra rechercher par l'image. Alors bien sûr, ça ne va concerner que certains corpus au début...

V. P. : Ce n'est pas GallicaPix ?

Fanny : Ça s'appuie sur la même chose, l'expérimentation c'était GallicaPix que vous allez voir, qui est toujours en ligne sur cinq corpus, qui étaient : Vogue, la presse de la Première guerre mondiale, le miroir des sports et les papiers peints. On pouvait rechercher par mot clé à l'intérieur, vous cherchez dans Vogue par exemple les chapeaux et vous avez les affichages de toutes les images qui contiennent un chapeau. Parce que le logiciel a compris que c'était donc là Gallica Images ça s'appuiera sur l'intelligence artificielle pour reconnaître des formes. Du coup, il y aura des corpus... des premiers corpus sur lesquels on va travailler en ce moment, les collections spécialisées en premier pour entraîner l'IA, pour qu'après elle puisse passer, par exemple dans la presse ou dans d'autres formes, des livres pour reconnaître les images. Et donc là, oui, l'attente de la couleur, tu te doutes V. P. que nous l'avons déjà formulée, J. V. et moi, c'est une priorité de pouvoir chercher du bleu dans tout Gallica. Ça va être très compliqué et nécessiter beaucoup de développement.

V. P. : Ah moi, je ne pensais pas forcément par couleur, mais (même si c'est super), mais au moins... d'avoir la certitude d'ouvrir à une reproduction d'un ouvrage en couleur.

Fanny : Ok, donc ça c'est d'autres choses.

[Chacun y va de son commentaire et de sa question]

Arnaud : Il y a deux choses en fait : la reconnaissance de formes (au lieu de rentrer par mot-clé dans l'image, on va donner une image et on retrouve par la similarité de formes) et l'identification des images dans les monographies et dans les presses. Donc là il faut segmenter les documents pour trouver où sont les images et ensuite, pouvoir les interroger par ce qu'on identifie dans la légende (comme dans GallicaPix) soit par similarité de contenus. Il y a le développement de ces deux pistes là en parallèle avec dans un premier temps l'entraînement du modèle de reconnaissance de similarités de formes par similarités d'images.

Fanny : des estampes...

V. P. : des cartes aussi ? Des plans ?

Fanny : Je pense qu'à moyen terme oui mais au début ça va être compliqué. Et puis reconnaître une forme, reconnaître une licorne au Moyen Âge et dans la presse du XXe c'est pas deux images différentes donc il faudra vraiment entraîner sur des corpus bien précis, des formes bien précises. C'est les 4 chantiers en cours du côté de Gallica avec ma nouvelle page d'accueil, normalement, fin d'année, vous ne pouvez pas la louper.

C. A. : Après on a mis, mettre d'abord l'affichage mosaïque sur l'interface, changer l'affichage des résultats...

A. V. : En fait on voudrait une fonctionnalité plus visible mais ça ne conviendra peut-être pas à tout le monde.

C. A. : Pour la recherche avancée, on avait mis à hiérarchiser, comme le catalogue général parce que... les gens remplissent tout comme un formulaire, comme s'il fallait tout remplir.

Fanny : Oui donc quelque chose qui pourrait s'ouvrir avec des rubriques ?

A. V. : A déplier, replier.

C. A. : Et dans la question de l'affichage de résultats, il y avait aussi le résultat par pertinence, mais on ne sait pas ce que c'est par pertinence ?

Arnaud : Oui c'est un critère

O. F. : pas transparent... Et je défie ceux qui ne connaissent pas de nous éclairer [rires], de comprendre ce que c'est [rires].

Arnaud : La pertinence c'est un calcul. Après, ça dépend de quel point on se place...

O. F. : Oui mais il faut que l'objectivité et soient perceptibles, disons consciemment ou subjectivement, à la réflexion.

Fanny : On écoute les dernières, enfin si vous avez fini ?

C. A. : Ouais

A. M. : Bon moi c'est un métier à part.

Fanny : Oui oui.

E. M. : Bon, il y a des choses qui ont déjà été dites : affiner les possibilités de recherche par proximité, on a dit aussi l'amélioration des fonctionnalités de téléchargement, alors ça, c'est plutôt un retour que j'ai eu par des utilisateurs et des chercheurs qui travaillent sur la presse. La possibilité a minima de pouvoir voir tous les numéros disponibles dans le journal et pas de devoir passer par le calendrier. Le mois qui ralentit beaucoup et ce serait la version plus à minima et idéalement pouvoir télécharger d'un coup tous les numéros.

Fanny : Ça je pense que, il faut avoir un bon réseau. [Rires]

E. M. : Très concrètement, je me base sur l'exemple d'une amie qui avait eu besoin de faire ça pour une recherche et son copain qui était informaticien lui avait un petit algorithme.

Arnaud : Oui, ce serait possible avec les API.

Fanny : Oui.

E. M. : Mais voilà, elle s'en est beaucoup servie. Je pense que si elle devait donner une nouvelle fonctionnalité, ce serait celle-là. Voilà... Et puis globalement, oui, un peu expliciter via des tutoriels ou des capsules vidéo comme c'est déjà fait trouvables sur YouTube...

V. P. : Chaque nouvelle fonctionnalité trouvable sur Gallica...

E. M. : Ouais

V. P. : De manière à ce que pour chaque outil, pour chaque nouvelle chose on ait quelque chose qui permettrait de naviguer dans l'outil, dans la banque, la bibliothèque numérique. Un tuto dédié à chaque nouvelle fonctionnalité...

E. M. : Et puis expliciter des stratégies de recherche un peu expertes. Peut-être avoir des tutoriels gradués par niveau de familiarité numérique, du plus simple au plus complexe et ne pas réserver ça à une expertise de bibliothécaire...

A. M. : Bah oui puisque dans les retours d'expérience, on voit bien que les retours et les mots que vous utilisez sont des éléments métier très propres à la bibliothèque, à la recherche bibliographique, sur des métiers de la médiation culturelle et de la pédagogie. On n'est pas du tout à ce niveau-là. En fait, on est en amont, on est vraiment dans une promotion et une information de l'outil. Et quand je parle de formation pour le premier degré, ce n'est pas une formation au niveau bibliothécaire, c'est une information de ce que c'est l'outil pour les professeurs, on n'est pas du tout dans une analyse fine de résultats comme vous le développez. On est bien en amont, on est déjà dans une explication pour les profs de ce que c'est Gallica en fait. On ne va pas plus loin. Donc on est... ça je pense que nous... Donc c'est pour ça que moi j'avais noté donc améliorer l'interface, peut être créer un petit onglet sur “Gallica enseignant”, Gallica... ou dedans, on pourrait mettre les tutos premier niveau de présentation de l'outil et puis toutes les petites formations qui arriveraient derrière pour savoir...

XX : Il y en a déjà ?

A. M. : les nouvelles fonctionnalités, parce que eux, ils sont complètement perdus. Ils connaissent de loin Gallica, mais ils ne savent pas... ils ne savent pas vraiment comment naviguer dedans, même des enseignants, ils ne savent pas. Et moi je pensais même créer peut-être... (Tout à l'heure, vous parliez de d'arborescence avec Gallica), avec Gallica Image ça va être super pour nos métiers, parce qu'il y a vraiment cette arborescence par corpus et par mot qui va aller chercher plein de choses. Moi j'imaginai peut être aussi un corpus par niveau scolaire, alors peut être que ça va se recouper avec Gallica images, mais je sais que nous, par rapport aux B.O, au Bulletin officiel et aux programmes du primaire, du secondaire et plus, on va avoir des grandes thématiques de recherche, parce que vous savez, il y a “se représenter le monde”, il y a aussi l'EMC (éducation civique), il va y avoir l'EMI (éducation aux médias)...ça, ça peut être aussi des corpus qui pourraient rentrer dans quelques...

Fanny : C'est de l'éditorialisation ...

A. M. : Oui, de l'éditorialisation de corpus/ programmes scolaires, comme ça, ce serait super parce que nous, en fait, là, on fait quelque chose de très et de très pertinent qui va s'ancrer. On va lier les grandes thématiques par niveau scolaire au corpus image que vous allez faire et après ça va être tout bénéf pour les profs.

C. A. : Peut-être que ce qui est intéressant là, tu sais que y a des vellétés dans vos départements sur certains sujets très précis, il n'y a jamais un ensemble... commun [rire].

Fanny : Transverse ?

C. A. : [rire] Je n'ai pas osé.

Discussion de certains sur les essentiels.

V. P. : Il n'y a pas forcément le lien vers Gallica dans les essentiels.

Fanny : Oui, parce que les essentiels, en tout cas pour l'instant, c'est vrai qu'ils sont beaucoup sur l'actualité BnF parce qu'ils les construisent comme ça, peut-être que ça arrivera ? Un des buts de remettre à plat arborescence des pages sélection c'est justement d'amener à des thématiques "transverses". Le sport par exemple, il y a des images de sportifs aux estampes, dans la presse, il y a du sport au LLA, etc, etc. On va y arriver.

A. M. : Toutes les disciplines scolaires, en fait, ça se calque très bien sur Gallica. Et en fait, il y a un classement évident entre les disciplines transverses de la BnF -tout type de supports, toutes aires géographiques, toutes thématiques- en lien tout de suite avec les disciplines pédagogiques qu'on tous.

E. M. : Je pense que c'est important de disséminer l'information hors de Gallica. Je sais que moi, à titre personnel, je passe beaucoup par Youtube et assez spontanément, je pense que j'irais taper "faire une recherche particulière dans Gallica" sur YouTube pour avoir cette information là sans rentrer dans Gallica.

D. B. : Je suis désolée, je m'aperçois que je n'ai pas redis ici, mais peut-être on peut remplacer par une de nos remarques le fait que ce serait bien que les DL numériques soient chercheables en plein texte, peut-être à la place de la recherche d'images vu que c'est déjà en cours pour la recherche d'images spécifiques. C'était permettre la recherche plein texte dans tout ce qui va arriver par le DL numérique, des livres, des éditeurs, le dépôt...

Arnaud : C'est ce qui va arriver par flux ?

D. B. : Oui, ça va être massif. Et en fait, pour l'instant, il n'y a pas de fonctionnalité de recherche, en fait c'est comme si on avait juste une liseuse avec un catalogue, on va avoir les livres [rire] C'est une liseuse sur ordinateur. Mais oui ! [Rire]

O. F. : Arnaud, tu peux rappeler le... tu l'avais dit une fois mais j'ai un doute, le pourcentage des textes sur Gallica qui sont numérisés en plein texte ?

Arnaud : Oh, c'est de l'ordre de 80 %, de 75 %, de 80 %.

O. F. : Ah ça c'est beaucoup !

Arnaud : En fait, ça dépend comment on calcule. C'est-à-dire comme il y a des livres qui sont moissonnés, ces livres qui sont moissonnés ne sont pas accessibles en mode texte. Donc on va les exclure des résultats. Si tu cherches dans le moteur de recherche uniquement, les livres en plein texte, il va te sortir un taux très faible. Mais évidemment, il y a tous les partenaires dont les livres sont moissonnés et qui par défaut ne sont pas accessibles en mode texte. Donc, pour ce qui est intégré dans Gallica, les livres intégrés dans Gallica, c'est à peu près 75 %...

O. F. : C'est les trois quarts

Arnaud : Oui, les trois quarts.

O. F. : J'aurais cru que c'était un chiffre beaucoup plus bas.

Arnaud : Ça c'est pour les livres et pour la presse, et plus encore, c'est 80 %...

O. F. : Et il est prévu de faire en sorte que la totalité un jour soit ...

Arnaud : Oui, un des objectifs du travail c'est effectivement l'océrisation rétrospective, donc ça, c'est une collègue du département de la conservation du service num qui envisage dans les cinq ans à ce que l'ensemble de la collection d'imprimés et monographies soit accessible en mode texte. C'est l'objectif, mais après, il y a des difficultés en qualité d'OCR. Il y a des choses qui ont été numérisées, qui ne sont pas dans Gallica parce que l'OCR comporte trop d'erreurs. Donc on ne peut pas tout verser dans Gallica, même si ça a été océrisé à cause de la qualité des papiers et la qualité de la numérisation, est-ce que c'est d'après microfilms ? Il y a ça aussi.

A. T. : Oui, on a déjà une idée originale. Ce n'est pas la peine de passer du temps et de l'argent sur microfilm.

Arnaud : Mais c'est des chaînes en fait, ce n'est pas des choix au titre.

D. B. : En fait, ils disaient que ça prenait plus de temps de trier entre ce qui est déjà numérisé...
[Rire] On va avoir des doublons beaucoup...

A. V. : Il faudrait aussi ajouter l'océrisation des génériques de films.

Arnaud [répète]

A. V. : Un exemple de recherche : des étudiants ou des chercheurs qui travailleraient sur les femmes dans des postes techniques au cinéma. Cette info ne figurera sur la notice bibliographique évidemment. Par contre, si le générique était océrisé, ça serait génial.

Arnaud : Ça en fait, il faudrait nous le faire remonter.

R. T. : On a le même cas pour les programmes de spectacles en fait qui sont numérisés en mode image.

Arnaud : A l'occasion de cette consultation-là, faites-nous remonter les besoins même si effectivement, ça ne nous concerne pas nous directement. Même si on ne peut pas vous apporter une réponse mais on est quand même dans une démarche de collecter les besoins donc on en tiendra compte autant que possible et en tout cas on aura conscience de ces besoins là et de ces difficultés-là surtout par spécificité de collection car on les connaît mal et on n'en connaît pas les usages.

Fanny : Merci pour cette première partie.

[Présentation du schéma par Arnaud]

[...] Sauf que on ne peut pas combiner ces expansions. En fait, elles s'additionnent alors qu'elles pourraient s'inclure de manière à désambiguïser. Donc ça, c'est une idée, une piste sur laquelle on réfléchit.

D. B. : Mais, quand tu dis ça, ça te retourne... Du coup, tu fais ta requête et ça te retourne quelque chose d'enrichit les résultats d'une requête enrichie et entendue. Ou alors ça pourrait, tu pourrais choisir "étendre à" et puis on cocherait parce que c'est très différent en fait, s'il y en a un, on renvoie beaucoup de bruit par rapport à ce qu'on a demandé et l'autre on suggère des moyens d'élargir la recherche a posteriori et c'est l'usager qui a la main dessus en fait.

Arnaud : Alors, effectivement, ça c'est une très bonne remarque, une très bonne piste. C'est à dire que. Pour le moment, les deux fonctionnent en même temps et en parallèle, et il y a un algorithme qui va calculer la pertinence, c'est à dire qui va prendre les dix meilleurs résultats selon son calcul de probabilités à la fois dans l'expansion par lemme et dans l'expansion orthographique. Et du coup, les résultats... ça va cumuler les résultats. L'objectif, enfin, c'est une recherche élargie par défaut, avec toutes tous les résultats possibles. Et en fait, comme le dit le DSI, la beauté de l'algorithme, c'est que à partir de cet ensemble de résultats, il va y avoir un calcul de pertinence et un ranking qui va permettre d'ordonner les résultats. Et donc effectivement, on pourrait être dans une recherche ciblée en fait par défaut et proposer à

l'utilisateur une expansion par lemme ou par synonyme ou par orthographique, à la demande. Peut-être que c'est une des solutions justement pour améliorer l'interprétation des résultats. En tout cas, le problème c'est qu'effectivement ces expansions elles s'additionnent et ensuite il y a un algorithme qui va venir -Troisième point algorithme de ranking, c'est un de classement des résultats calculés en fait une pertinence- en fonction de 1 : l'auteur, 2: le titre et 3 : le sujet et en fonction de ces catégories-là, le nombre d'occurrences des termes. Le terme que vous avez choisi, c'est d'abord dans le titre qu'il va s'afficher ci-dessus et ensuite en tant qu'auteur, et ensuite dans le texte en fonction de l'occurrence des termes. Et donc il y a un calcul qui va permettre de pondérer et de remonter en fait les résultats en fonction des paramètres qu'on peut ajuster. Donc on peut choisir de valoriser les résultats partenaires ou les résultats à la BnF. Et donc on a ce qu'on appelle des boosts, des manières de à partir de ce calcul, de pousser en fait certains résultats. Mais pour être tout à fait honnête avec vous, en fait, c'est une interface très complexe de paramétrages et de réglages, que ça a été fait et maintenu sur une dizaine d'années avec des collègues qui se sont succédés et c'est vrai qu'il y a des réglages et des paramétrages qui n'ont pas forcément été documentés et dont aujourd'hui, et le travail avec vous, c'est d'identifier effectivement aussi des cas d'usage et donc des exemples très précis pour voir, côté système, comment est-ce qu'on peut améliorer justement la pertinence en s'appuyant sur des cas d'usage et sur des exemples précis pour nous-mêmes d'essayer d'ajuster ces réglages. Et donc on a une stratégie à court terme qui est, à partir de ce matériau qu'on construit, de pouvoir voir comment on peut améliorer les choses. Et puis, à plus long terme, comment s'inscrire dans une perspective d'évolution, d'augmentation avec des nouvelles fonctionnalités, voir le changement de solution ? A. B. ?

A. B. : Juste pour rebondir. Euh, je vois bien que ce n'est pas un truc qu'on utilise beaucoup et c'est quand on va chercher quelque chose de précis mais est-ce que c'est un choix de ne pas avoir mis à troncature ou... ?

Arnaud : Oui, j'allais y venir parce qu'effectivement, vous ne l'avez pas proposé dans les améliorations.

A. B. : Si justement, moi j'en ai parlé.

Arnaud : Oh pardon A. B.

A. B. : Oui justement, j'ai dit que ça ne marchait pas.

Arnaud : Oui, on s'attendait à ce que ça revienne très souvent parce que c'est une attente. Alors le problème de la troncature, c'est le temps de calcul en fait, c'est énorme. On l'a tellement demandé que là aussi, on va voir avec nos collègues du DSI. La stratégie pour nous, c'est s'appuyer sur les collègues et sur les usagers pour prioriser des améliorations et essayer de voir quel coût technique ça a pour pouvoir les mettre en œuvre. Et donc c'est la première phase, un peu de notre travail avec vous. Donc on a bien, on a bien en tête donc, pour les collègues qui font de la recherche bibliographique dans Gallica, la question -surtout pour la création de corpus pour le Datalab c'est nécessaire- d'intégrer cette fonctionnalité de la troncature.

V. P. : Surtout pour le permettre pour le sans cote. Parce que parfois, à une erreur près, le libellé de la cote, ben on galère et c'est fichu. Alors que si on rentre une racine de cote, finalement, on va retrouver en plus toute une série de documents qui potentiellement va être....

Fanny : C'est vrai que juste l'expansion orthographique par exemple, vous n'en avez pas du tout parlé et on a fait le même atelier avec notre service un peu élargi et l'expansion orthographique c'est ce qui est revenu le plus souvent. Donc je pense aussi qu'on a des usages différents malgré le fait qu'on est... voilà on est 2000 à la BnF et donc voilà, avec ces ateliers, on rend compte de la complexité du moteur de recherche aussi des résultats différents. C'est vrai que l'expansion orthographique est ressortie comme étant un vrai problème parce que quand on cherche une typographie, une typologie et qu'on a typographie.

D. B. : Oui, mais inversement, comme on est habitués à ça, j'imagine qu'on ne voit pas les erreurs d'OCR. Du coup, il y a beaucoup de documents qu'on voit et qu'on trouve parce que les erreurs d'OCR sont corrigées...

Fanny : Normalement, il n'y en a plus beaucoup

D. B. : Il n'y en a plus beaucoup, parce qu'ils ont été corrigés ?

Fanny : Oui oui, le pourcentage d'OCR qui est bien meilleur que à l'époque où on avait besoin des extensions orthographiques pour justement corriger l'OCR...

D. B. : Mais c'est rétrospectif, du coup les documents ont été numérisés il y a longtemps ?

Fanny : Oui !

D. B. : Ah d'accord, ok. Ah ok.

Fanny : Et aussi avec 1 million de documents...

D. B. : D'accord.

Fanny : et aujourd'hui on en a dix fois plus.

D. B. : Oui, donc effectivement ça vaudrait le coup...

Fanny : Justement, ce que le DSI trouve "beau" [rires] c'est une difficulté pour nous mais bon.

D. B. : Et c'est vrai qu'une des questions, c'est vraiment aussi la pondération entre les mots de la notice et puis le plein texte. Et de faire cette distinction parce que finalement, parfois, on voudrait juste faire un bouton et dire je veux juste dans la notice en fait, via le formulaire de recherche avancée, tous les mots de la notice. Mais une fois qu'on a des résultats, peut-être pouvoir affiner et dire juste ce qui est dans la notice puisque c'est vrai que la pondération 1 de titre, auteur, c'est normal, sinon on ne trouverait rien je pense. C'est des paramètres qu'on ne peut pas bouger c'est dans tous les moteurs de recherche. Mais après entre le sujet, entre ce qui est dans une notice et ce qui est dans du plein texte, c'est quand même des modes de recherche très différents. Je ne sais pas comment...

Fanny : Alors oui c'est différent mais à la fois, moi je prends la casquette grand public, le grand public il cherche par sujet, il ne cherche pas par type de document, donc il se retrouve avec des documents justement...

D. B. : Pour moi il y a quand même deux catégories. Parfois ils savent qu'ils cherchent un document et veulent le trouver et parfois on fait une recherche thématique. Fin ce n'est quand même pas la même...

V. P. : Même pour thématique, il faut finalement taper même dans le champ titre. Moi je fais les deux à la fois dans le sujet et dans le titre parce que sinon...

D. B. : Bah moi je mets tous les mots de la notice dans ces cas-là. Parce que oui, effectivement, parfois il y a une forme.

Fanny : L'exemple du printemps de C. A. tout à l'heure, c'est vrai que le printemps en titre de document, il y en aura beaucoup moins que en sujet. Bon après ça questionne aussi la manière dont sont... l'état des notices en fait, est-ce que les sujets sont bien indexés. Par exemple la statue de la liberté, mais en fait, on ne trouvait pas puisque le champ sujet était en anglaise et moi j'avais écrit liberté et pas liberty. Et ça ne marchait pas, il y a plein de choses qui font que...

D. B. : Bah oui. Non mais oui, mais je suis d'accord comme d'accès principal mais peut être comme option de filtre, ça peut être très efficace effectivement d'enlever parce que ça permet d'enlever du bruit.

A. T. : Mais il y a une facette qui dit "trouver dans" qui permet de savoir si l'expression est trouvée dans la côte ou dans... justement parce que la plupart des gens font une recherche plein texte et... en tout cas ça, c'est peut-être quelque chose qui...

Fanny : En tout cas à vous entendre il y a quand même une amélioration des facettes à avoir.

V. P. : Ah oui. Des facettes et même des fonctionnalités de recherche.

Fanny : Oui, des fonctionnalités de recherche, c'est vrai que les facettes, elles vont répondre au besoin des personnes qui vont chercher par recherche simple, par recherche avancée.

V. P. : Oui.

R. T. : Mais si les facettes fonctionnaient dans leur plein potentiel, on n'aurait pas besoin de passer par la recherche avancée.

D. B. : Alors oui, c'est ça l'autocomplétions et tout ça, c'est une sorte de recherche avancée dans la recherche simple, sans passer par auteur/ titre. En fait, c'est une façon de remettre les facettes aussi. C'est une façon de remettre de la recherche avancée dans la recherche simple.

Fanny : Mais on le voit dans les usages numériques, on est quand même très habitués aux facettes.

V. P. : Et ça, c'est dans les logiciels de banque d'images. On a beaucoup vu à un moment donné que on simplifiait les champs de recherche et on misait beaucoup sur les facettes. Moi, j'étais absolument contre ça quand j'étais sur le chantier de la banque d'images, parce que oui, c'est bien, mais en même temps affiner en amont et ça permet quand même d'être précis. Et je pense que là, ça bouge de ce côté-là. Au niveau des agences.

Fanny : Je pense qu'il y a des usages aussi différents.

Arnaud : Peut-être que je termine le schéma et après on discute. [...] Des questions, des remarques ?

[Chacun commente]

D. B. : Tu as dû le dire Arnaud mais c'est développé entièrement en interne ? Ou c'est à partir d'une ressource externe ?

Arnaud : Alors non, c'est une brique logicielle, c'est une prestation. [...]

Fanny : Et un des objectifs aussi de recueillir les besoins, c'est de savoir est-ce qu'est facilement modifiable via du paramétrage. C'est-à-dire est-ce qu'on peut satisfaire ce besoin parce qu'on a les paramètres qui permettent de. Ou est ce qu'il y a vraiment à se comprendre quelque chose de technique soit Le moteur de recherche ne peut pas répondre à ça. L'extension orthographique, ça a été un choix a été fait. On peut revenir en arrière...

A. B. : Et dans ce cas ce serait carrément un changement de solution ?

Fanny : De moteur de recherche ?

A. B. : Oui.

Fanny : C'est possible.

Arnaud : Déjà, pour être honnête. C'est vrai que ça fait 10 ans, que c'est une solution qui est maintenue en interne mais qui est plus leader sur le marché aujourd'hui, on est aussi dans un contexte de renouvellement des modalités d'accès à l'information avec l'IA. Il n'y a pas d'IA dans notre moteur de recherche. On considère qu'on a une mine de données avec nos catalogues et nos autorités qui est peu exploitée, mal exploitée. On voudrait faire de la recherche par le concept, de la recherche par sujet. Il y a déjà des extensions qu'on souhaiterait et qui ne sont pas proposées par les autres de recherche et donc on sait qu'il y a des solutions sur le marché qui existent. La logique aussi, c'est s'inscrire dans un plus long terme et effectivement commencer à réfléchir à ce qu'on veut et comment est-ce qu'on peut faire évoluer les fonctionnalités et quelles sont les solutions qui permettraient peut-être Solr hein, je regarde D. B. parce que c'est sans doute la solution qui est utilisée, en tout cas de manière expérimentale, autour des archives du web. Il y a d'autres types de moteurs. On a pour objectif d'avoir aussi un état de l'art, un état des lieux de ce qui existe pour pouvoir partager avec vous. De toute façon ou non, dans un processus un peu... On reviendra vous voir, vous présentera les résultats des enquêtes d'usage, mais aussi de l'audit interne pour...

Fanny : Peut-être un peu finir sur ça, sur dans quoi s'inscrit toute cette réflexion. On n'a pas détaillé encore.

A. V. : Juste pour info, solaire c'est aussi un moteur de recherche qui est utilisé sur le portail de diffusion des documents audiovisuels en Haut-de-Jardins. [...]

A. T. : Mandragore aussi c'est Solr. Et BAM aussi.

Fanny : On va dire que, en fait, changer de moteur, de logiciel, ce serait du long terme. Donc mettre en place avant, est ce qu'il n'y a pas ce moyen terme sur le paramétrage de quelques petits points ? Prévoir une facette pour un usage que tout le monde a mais qu'on n'a pas... Moi je le vois comme ça comme moyen terme.

Arnaud : Il y a peut-être des... comment dire ? Des paradigmes à changer. On est dans une phase de recherche élargie qu'on va ensuite affiner, filtrer. Est-ce que finalement on ne va pas changer ? On commence avec peu pour ensuite pouvoir ouvrir, ouvrir, ouvrir ? Mais la recherche avancée... Dans un audit sur d'autres plateformes on nous a montré, d'ailleurs c'est le travail qu'avait fait Eléonore, de la "recherche progressive" pour déployer progressivement.

Fanny : C'est un peu ce que vous aviez dit non ? La recherche avancée qui peut se déplier.

V. P. : C'est un peu comme un catalogue.

Arnaud : Au niveau de l'interface de Gallica. On ne touche pas au moteur. Si on modifie les logiques d'affichage, on ne touche pas au moteur donc il y a des choses sur lesquelles on a la main.

Fanny : Plus directement.

O. F. : Voilà un point qui me revient seulement maintenant à l'esprit, c'est cette question de mise en relation, de mise en réseau des documents. Dans les résultats, il y a sur la gauche, dans la colonne de gauche, des documents qui sont affichés en disant "Document..." alors je ne me souviens plus de la formule exacte. Alors ça, ça ne fonctionne pas du tout. C'est très mauvais comme résultat. Est-ce que ce ne serait pas mieux de l'enlever ? [Rires].

Fanny : Je sais, parfois il y a des choses...

O. F. : Voilà.

Fanny : Alors en fait, ça fonctionne... [rires] Il y a certains champs de la notice et notamment description, sujet et auteur et s'il y a d'autres documents, ça les propose à ce moment-là. Alors pour auteur ça peut être pertinent. Pour sujet parfois ce n'est pas tout à fait pertinent. Et alors pour la description c'est pas du tout, mais alors pas du tout pertinent. C'est souvent lié aux collections. Il y a un vrai travail sur ce point On voit et on se rend compte qu'on est habitué aussi, je pense, par nos usages numériques autres que les moteurs de recherches et tout ça à utiliser les facettes. C'est quand même... Vous allez sur n'importe quel site de e-commerce, avez des recettes. [Rires] Et on a aussi des renvois à d'autres produits et on est habitué à ça.

Mais là, pour Gallica, ça fonctionne assez mal en l'état. Alors pour l'auteur, je pense que ça va encore. Le sujet ça dépend, si c'est un sujet très précis, on va avoir...

D. B. : Alors on pourrait avoir... Enfin je ne sais pas si ça marchera, mais j'imagine pour certains fonds spécifiques, une sorte de... des étagères virtuelles avec feuilletage par cote, c'est-à-dire faire défiler par proximité dans le fond les documents.

Fanny : Je pense que les rebonds en fait, les renvois vers des documents, l'expansion de la recherche après avoir vu un document.

A. V. : Même par auteur, je crois que ce n'est pas super. Parce que sur un corpus d'une vingtaine de documents, j'étais en train de les passer un par un, il me proposait toujours les mêmes deux, trois du même auteur.

Fanny : Oui, voilà alors que ça aurait été beaucoup plus varié.

A. V. : Mais c'était comme si... on avait l'impression de rester bloqué sur les mêmes,

Fanny : Ouais. Ça j'ai... ça aussi, je ne sais pas exactement comment ça fonctionne.

Arnaud : En fait, là aussi c'est de nouvelles requêtes. Un des travaux qu'on va faire aussi avec le DSI c'est que sur chaque page il y a une multitude de requêtes qui sont lancés et des précisément les rebonds ce sont des requêtes, donc ça affiche le document qui est dans un des champs.

Fanny : Il y a auteur, sujet et description. J'ai identifié ça.

V. P. : Est-ce qu'il a un travail ? Est-ce que c'est par exemple le mapping entre le catalogue du catalogue et ce Gallica va être remis à plat ? Au niveau du moissonnage des données ?

Arnaud : Alors, on a en tête l'idée d'une convergence de fonctionnalités. Donc c'est vrai qu'on en a discuté avec D. B. pour l'interface d'accès des archives du web. C'est vrai que sur le catalogue général et sur BAM. La difficulté, c'est qu'on n'est pas exactement dans la même logique. Le moteur de recherche de Gallica il combine à la fois la recherche simple par mot clé et une recherche de type plutôt recherche sur le web, même si ce n'est pas le même fonctionnement que Google, mais c'est quand même plutôt ce moteur de recherche là et une recherche par catalogue ou par bases de données, c'est à dire de chaînes de caractères dans un champ. Donc ce n'est pas exactement le même fonctionnement et on essaie de combiner les deux, en fait de manière transparente pour les utilisateurs, mais en fait c'est deux logiques différentes. Donc on remettre le catalogue général, donc Gallica, en fait c'est...

V. P. : Moi ce n'est pas ça que je voulais dire, c'était surtout faire en sorte que le nom et le mapping, il y a peut-être des champs dans le catalogue qui ne sont pas mappés avec Gallica et qui permettrait des recherches. Je n'ai pas d'exemple hyper précis là.

Arnaud : Je comprends.

A. T. : Ceci dit on a le problème inverse avec BAM parce que si on mappe tout, on va se retrouver avec beaucoup d'informations et donc beaucoup de bruit.

V. P. : Pas tout, mais il y a peut-être quand même pour la recherche icono des petites choses à améliorer au niveau du mapping.

Arnaud : C'est ça. C'est en fait ce qu'il se passe c'est que les champs du catalogue général de BAM sont compressés, combinés dans un format de quinze champs qui est le Dublincore. Donc on a dans le champ description du Dublincore tous les éléments des métadonnées du catalogue général. Et ça, effectivement, c'est un format, interopérable... On n'a pas la main. Est ce qu'on va changer de format ? On ne sait pas. Là, on est...

V. P. : C'est un travail qu'on a fait quand même pas mal pour la banque d'images de mapper comme ça. Et moi je crois vraiment qu'il y a des pertes d'informations entre le catalogue et le... notamment pour l'icono.

Fanny : Mais peut-être que ça vaudrait le coup que tu nous donnes des exemples.

V. P. : Oui, il y a des choses très spécifiques. Je pense que si on arrive à les faire se combiner un peu, ça peut améliorer la recherche par sujet ou par thème. Même par rapport à certains fonds...

Fanny : Mais en tout cas, le lien avec le catalogue général est beaucoup ressorti dans le premier atelier qu'on a fait avec le service. Moi, j'avoue que je ne passe jamais par le catalogue général mais jamais donc pour moi c'était un non-sujet que la recherche commence par le catalogue général pour finir dans Gallica. Mais en fait, on a eu beaucoup de retours. Je pense qu'il y a quand même une efficacité que l'on retrouve dans le catalogue général et qu'on ne retrouve pas sur Gallica et il va falloir comprendre le pourquoi. Qu'est-ce qui fait qu'on a ...

V. P. : Il y a une perte d'information !

Fanny : Et c'est pour ça que les exemples précis, en fait ça peut nous aider (quels champs sont intéressants...).

V. P. : Il y a même des choses à faire à partir d'une racine de cote pour transcrire ... une racine de cote peut correspondre à un thème. Et ça, on gagne un temps quand même considérable.

Fanny : Je vous propose qu'on finisse rapidement pour vous dire où on va avec ce travail-là. Déjà ce printemps, on va faire plusieurs ateliers, vous avez participé au premier, il y en aura un deuxième avec d'autres départements. [...]

Fanny : Est-ce que vous avez des remarques ?

[Question de V. P. sur les formations futures pour la recherche dans Gallica]

[Question d'E. M. sur le calendrier du projet]

Transcription de l'atelier métier 3 : 30 avril 2024

Temps 1 : Témoignages

Fanny : Ce que l'on attend dans cette prise de parole de chacun c'est à la fois vos cas d'usage personnel, de votre entourage, les besoins dans vos recherches. Qu'est ce qui fonctionne bien ? Qu'est-ce qui fonctionne moins bien ? Si vous avez des exemples précis de recherches récentes, et puis aussi ce que vous avez mis en place pour pallier à certaines défaillances que vous constatez. Donc ça peut être des stratégies, ça peut être passer par un autre site, ça peut être chercher dans un corpus que vous connaissez bien, ça peut être ça, ça peut passer par la médiation de Gallica, etc.

Arnaud : Précisez aussi les spécificités de vos collections. On avait dans l'atelier précédent des collègues, voilà des arts du spectacles, de SVM (son vidéo multimédia) qui nous signalaient pour leur collection les spécificités d'accès.

Fanny : Sachant qu'on sait que vous êtes sûrement un très bon connaisseur et très bonnes connaissances de Gallica et que vous allez avoir un usage et un usage déjà un peu expert par rapport au grand public. Donc il y a des choses qui seront à mon avis de la dentelle dans la recherche par rapport à ceux qui l'utilisent pour la 2e ou 3e fois. Je vous laisse la parole ? Donc on se dit qu'on a une bonne heure, 1 h 15. Est ce qu'il y a une personne qui veut commencer ? Alors, vas-y, tu étais le premier.

J. O. : Je m'appelle J. O.. Département des monnaies, médailles et antiques, je m'occupe des monnaies grecques. Et entre autres choses aussi, des questions de numérisation et de médiation numérique. Et alors, c'est très bienveillant, mais c'est assez sévère.

Fanny : On s'en doute.

J. O. : Alors nous, on a plusieurs types de collections dans Gallica, évidemment des imprimés de manuscrits et des archives, mais surtout des monnaies et des médailles qui sont depuis maintenant une grosse dizaine d'années on en a un quart de million. C'est destiné à augmenter ou à doubler. Et le moteur de recherche pas du tout adapté à ce type d'objets. Il y a eu beaucoup de progrès, notamment avoir deux images endroit/envers et ça, c'est déjà assez pratique pour visualiser rapidement les résultats des requêtes. Mais comme la recherche est avant tout iconographique et qu'on n'a pas ces informations dans Gallica, la recherche est très peu, très peu efficace. Est-ce que tu disais que sur les deux facettes, il y a des gens qui sont très content de trouver plein de choses et puis finalement qui ont des difficultés à trouver ce qu'ils

cherchent : je pense que c'est exactement ça le problème qui se pose de manière générale, quand je peux me laisser surprendre, je vais voir Gallica parce qu'on va trouver plein de trucs très différents avec une recherche simple et souvent c'est pas toujours intéressant. Par contre, quand on sait ce qu'on cherche, là c'est plus compliqué. Moins d'avoir la cote ou l'Ark en réalité. Je vois bien avec -des fois avec mes collègues, mais plus encore avec les chercheurs qui fréquentent les départements- un des principaux problèmes, c'est de leur expliquer qu'il ne faut pas passer par Gallica. Il faut passer par le catalogue général où sont les notices avec les datas et ensuite elles arriveront dans Gallica et là, effectivement, ça marche beaucoup mieux. Et ça, c'est un message qui est très difficile à faire passer parce que la vitrine des catalogues de bibliothèque, ça reste Gallica. Ce qui est assez étonnant. Et pour terminer, on a une stratégie de contournement assez puissante qu'est le web sémantique de plus en plus et chez nous et dans d'autres institutions à l'internationales se créent des portails, une typologie bibliographique peu importe qui en fait, agrège les collections d'un certain nombre d'institutions dont la nôtre, et de plus en plus. (Je pense que c'est plutôt une bonne chose, de plus). En fait, c'est par ce biais là que les gens arrivent sur l'outil, c'est qu'ils n'utilisent plus l'outil Gallica, donc en tant que tel. Voilà.

Arnaud : Quels sont les portails, en l'occurrence dont vous parlez ?

J. O. : Il y en a une foultitude par exemple on peut regarder OCRE (O.C.R.E) heu c'est Online Coinage of the Romane Empire. Donc c'est un portail qui est développé à l'American Numismatic Society à New York, qui décrit tous les monnayages des empereurs depuis Auguste jusqu'à la fin. Il y a une soixantaine d'institutions partenaires qui ont leurs connexions là-dedans. Et donc là, c'est un outil qui est conçu pour cette recherche spécifiquement, qui est multilingue et qui donne accès à (je crois que c'est) 180 000 monnaies à travers le monde. Et donc là on trouve nos collections.

Fanny : Et donc quand tu le disais tout à l'heure, pour vous c'est principalement de l'iconographie. Ça veut dire aussi que les pièces n'étant pas... Est-ce que tu es amené à faire des recherches sur le sujet de la monnaie en particulier la personne représentée. Ça peut être assez facile, mais c'est d'autres petits pas, d'autres détails, etc. Ça peut arriver. C'est pas dans la notice ?

J. O. : Oui, c'est ça, on aura évidemment les autorités, mais les autorités, quand on a affaire à un objet produit en série comme la monnaie, c'est compliqué si vous demandez les monnaies d'Hadrien, vous allez en avoir tout de suite 3 000 à 4 000.

Fanny : Oui...

J. O. : Pour Gallica, ça va se limiter ensuite à la dénomination au métal ce n'est pas assez fin. Tandis que dans le catalogue général où l'iconographie normalement est indiquée. Et donc effectivement, pour savoir si vous avez Roma ou un morceau de légende éventuellement mais voilà comme ces données ne sont pas dans Gallica...

Arnaud : Par ailleurs, vous n'êtes pas amenés à faire de la recherche en plein texte sur les monographies, les visuels ?

J. O. : Si et effectivement, ça, on le fait. Par exemple, on a un problème de numérisation depuis plusieurs années sur le catalogue de vente. Et là, l'outil Gallica est extrêmement efficace. Parce que oui, quand on fait de la recherche plein texte par exemple, ça marche extrêmement bien. Mais il y a aussi la petite complexité, comme vous disiez, nous on est là et on sait comment ça marche mais pour expliquer à un lecteur que si vous trouvez un catalogue de vente, c'est plus pratique de passer par le catalogue général en rentrant la date du catalogue parce que dans Gallica ça ne marchera pas et que pour faire la recherche plein texte sur des morceaux de ce texte-là, il ne faut pas passer par le catalogue parce que là ça ne marchera pas. Et ça, ce n'est pas évident à comprendre. Cette articulation, elle pose beaucoup de questions, même à des chercheurs assez familiers avec des outils numériques.

Fanny : Merci beaucoup, est-ce qu'il y a une autre personne qui veut prendre la suite ?

L. L. B. : Moi, je veux bien continuer un peu dans le même esprit d'ailleurs. Je suis L. L. B., chef du service de la conservation au département des Estampes. Alors je précise que je ne suis pas un expert de Gallica, et que je prends la suite, j'assume l'intérim. Sur les questions de service public. Suite au départ de L. B., qui a beaucoup travaillé sur Gallica. Voilà donc pour la situation au département des estampes. Je ne sais pas quelles sont les pratiques des collègues. Je pense en effet que la recherche passe plutôt par le catalogue dans lequel on a une connaissance ou une bonne connaissance des données. C'est par là, je pense, qu'on retrouve les numérisations. Les sélections qui ont été élaborées pour l'exploration de Gallica. Celles-là, je pense qu'elles sont quand même relativement bien pratiquées et signalées aux lecteurs. Les images, puisque c'est plutôt l'image qu'on s'occupe, elles sont brassées et exportées : pour les images libres de droits vers la banque d'images du département de l'image et des prestations numériques. Et je pense que les recherches iconographiques se font beaucoup sur cette plateforme qui combine les images issues de Gallica et les images produites par le DIP stockées donc dans leurs silos. Donc Gallica se trouve au milieu d'une chaîne, comme ça, à la fois d'informations et de transferts de

fichiers. Donc les recherches sont multiples. Pour l'usage des ressources qui sont dans Gallica, on essaie aussi de simplifier l'accès aux utilisateurs, aux chercheurs, en proposant les accès à un certain nombre de documents dans les pages du département. Là, on va proposer les outils principaux pour la recherche d'image, à savoir le guide du lecteur du département qui est dans Gallica, interrogeable en plein texte. On va aussi signaler tous nos inventaires de collections qui eux aussi sont interrogeables en plein texte. Une chose qu'on n'a peut-être pas encore signalée ? Je ne crois pas, mais en tout cas, on pratique, très souvent, c'est le comment dire ? Ce sont... les pardons... des registres de création des cotes qui, pour nous, servent à retracer, on va dire l'histoire d'un document. Donc tout ça est numérisé et pour nous, c'est une ... Voilà, c'est un outil absolument essentiel. Les registres sont conservés à la réserve des estampes, donc avec un accès extrêmement limité. Et puis finalement, chacun peut accéder de son bureau à toutes ces archives. Qu'est-ce qu'on peut dire encore ? Pour la recherche d'images, ça ne fait pas à mon avis directement dans Gallica, mais plutôt de façon indirecte. Sauf dans les sélections. Et pour les recherches plein texte, donc je vous ai parlé de l'exploration des imprimés spécifiques au département. A titre personnel, j'utilise aussi beaucoup les recherches textes dans la presse par exemple, avec quelques difficultés toujours, mais aussi avec de bonnes surprises, particulièrement avec l'affichage des listes de résultats que j'ai découvertes il n'y a pas très longtemps et qui, je trouve, sont très très utiles, pratiques donc il y a tout un travail de visualisation des résultats et qui est à mon avis vraiment bénéfique.

Fanny : Ok.

Arnaud : Vous pensez au rapport de recherche ?

L. L. B. : Oui, absolument. Et aux possibilités d'exportation qui sont vraiment très bien.

Fanny : Tu as des questions ?

Arnaud : Je ne sais pas si c'est le bon moment ? Mais est ce qu'il y a des attentes ou des besoins spécifiques propres aux collections d'estampes ? On vous parlera après un peu plus de la fouille d'images un projet qu'on a en cours d'instruction. C'est aussi l'occasion de faire le tour des besoins des départements.

L. L. B. : En fait, je vais parler à titre personnel. J'ai tellement peu l'habitude de faire mes recherches sur Gallica que j'aurais du mal à vous répondre.

Arnaud : Plutôt une convergence de fonctionnalités entre le catalogue général et Gallica ?

L. L. B. : Voilà. Je pense faire un atelier spécifique sur la recherche d'images.

Fanny : Et est-ce que sur la banque d'images (on a déjà eu cette... j'ai déjà entendu que la banque d'images pour chercher des images, ça pouvait être plus intéressant). Alors est-ce que c'est sur le mode de recherche à proprement parler, c'est-à-dire les mots qu'on va taper la pertinence des résultats, etc. Ou est-ce que c'est plutôt sur l'interface ou la présentation du document, le rebond entre les documents.

L. L. B. : L'interface de la banque d'images et je dirais très simple en même, on peut combiner des critères, mais on arrive assez vite au résultat. Peut-être aussi parce qu'il y a moins de résultats.

Fanny : Il y a moins de documents.

J. O. : Peut-être aussi parce que dans la banque d'images, ce sont des images, alors que dans Gallica, ce sont des livres. Donc on va sur la banque d'images pour chercher des images.

Fanny : Mais on a aussi eu la remarque que l'interface était beaucoup plus agréable pour chercher. C'est aussi pour ça que j'interroge, c'est effectivement, et Arnaud nous en parlera tout à l'heure, il y a aussi les paramètres de la recherche et puis aussi la manière de proposer ses résultats, les rebonds entre les documents, etc. Et c'est un peu deux axes différents d'amélioration.

L. L. B. : Mais ça vaudrait la peine effectivement de se pencher sur la recherche d'images spécifiquement dans Gallica.

Fanny : Oui oui, ça on a prévu. Est-ce qu'il y a une personne qui veut parler ?

M. D. : Alors M. D., département droit éco politique. Je suis coordinatrice de la numérisation et de la valorisation. Du coup, j'utilise beaucoup Gallica pour des recherches de documents à numériser, par exemple faire voir ce qui est numérisé, pas numérisé, etc. Et aussi pour la valorisation des recherches sur justement trouver des illustrations, trouver des articles de presse un peu tous azimuts. Du coup, je voudrais différencier ce qui relève du modèle de données du moteur de recherche lui-même, parce que dans ce qui a été dit, il y a pas mal de choses qui relèvent finalement des données et pas du moteur. Donc effectivement, c'est des données pauvres en Dublincore appauvri par rapport à la puissance du catalogue, pour les recherches, évidemment, ça n'a rien à voir et c'est très dommage. Mais voilà, la recherche en plein texte bah tout n'est pas océrisé donc pareil c'est très puissant, mais il n'y a pas tout. Du coup, je vais faire quelques remarques moi plus sur l'interface de recherche. Au niveau de la recherche par expression exacte. Euh voilà, si je cherche quelque chose. Alors en recherche simple, voilà...

Mais même en mettant titre, auteur etc, en mettant les guillemets, ça ne marche pas très bien. Donc il faut vraiment cliquer “expression exacte” en bas et là ça marche mieux.

Fanny : On vous en reparlera.

M. D. : Voilà, il y a aussi ce filtre qui est intéressant de trier les résultats par titre ou numéro de fascicule pour la presse. Mais du coup, ces deux possibilités, elles sont toutes en bas et elles sont un peu cachées et ce serait peut-être bien de les remonter vers les termes de recherche en fait. La partie sous... Alors la recherche par proximité : Super ça ! J'adore. La partie qui est juste en dessous je crois, qui est : recherche par site ou bibliothèque. Alors là c'est juste incompréhensible. Je n'ai pas encore compris la différence.

Fanny : Alors juste pour dire la différence [rires] les bibliothèques c'est les partenaires qui ont versé leurs documents dans Gallica, c'est intégré dans Gallica et les sites c'est les sites de consultation donc moissonnés. C'est pas clair. Parce que la Ville de Paris est intégrée et moissonnée.

M. D. : Ah, je ne savais pas ça.

Fanny : Mais je suis la première à dire que ce n'est pas clair. Mais effectivement, mon idée, c'est que si on garde ces deux encarts là il faut qu'ils soient plus clairs dans leur intitulé et est-ce que ça a un intérêt d'avoir deux encarts différents ? Je ne sais pas ...

M. D. : Mais effectivement. Et s'il pouvait y avoir une recherche par autocomplétion dedans, ce serait super parce qu'il faut avoir vraiment le nom de l'établissement exact. Et la liste étant très très très longue avec tellement de variantes. Et j'aimerais bien qu'il y ait des filtres a posteriori de la recherche qui permettent de faire des exclusions aussi ça ce serait pas mal. Dans la to do heu dans la liste au Père Noël. Euh. Je sais plus qui a parlé d'export ? Moi je n'arrive pas faire d'export de Gallica, je ne peux pas faire d'export. Justement, ça m'intéresserait beaucoup de pouvoir exporter. Alors qu'est-ce que je voulais faire ? Oui, des documents qui sont bloqués dans Gallica intramuros (heu, peut-être que c'est dans Gallica Intramuros qu'on ne peut faire d'export ? C'est bloqué ?)

Arnaud : Dans Gallica intramuros, ça n'a pas été déployé en tout cas pas en version en Rez-de-jardin. Alors on peut y accéder ponctuellement, je pourrai vous accompagner parce qu'on peut y accéder sur la plateforme de développement et on ne l'a pas mis en service pour des raisons de droits parce que ça fait des exports de vignettes et pour le moment on n'a pas validé avec le service juridique le fait...

M. D. : Mais juste l'export des données tabulées, il n'y a pas non plus ?

Arnaud : C'est possible mais pas en interne, en passant en interne ça n'a pas été déployé sur vos interfaces et vos accès.

M. D. : D'accord et du coup c'est parce que moi, mon objectif c'est de libérer des documents. Du coup, je voudrais bien pouvoir faire un export depuis Gallica Intramuros directement. Là je suis obligée de passer par le catalogue général, donc j'ai trouvé la parade. Mais bon, ça me complique un peu la vie.

Arnaud : Je pourrai vous montrer comment on fait. En passant par la PNG [?].

M. D. : D'accord, mais je ne suis pas sûre d'y avoir accès.

Arnaud : Je pense que c'est possible, on pourra regarder ensemble, je viendrai vous voir sur ce point en particulier.

M. D. : Ok, ça me ferait une étape en moins dans ma requête. Alors bon après voilà, on retombe un peu sur une histoire de données plutôt. C'est pour pouvoir trouver des illustrations dans du texte parce que ce n'est pas indexé en tant qu'illustration. La recherche OCR quand on est dans un document et qu'on veut chercher des termes dans le document. Alors déjà, le truc tout bête, c'est qu'on est obligé de cliquer dans la petite fenêtre, le curseur n'est pas positionné par défaut dedans. C'est, c'est bête, mais c'est très pénible.

Arnaud : Je n'ai pas compris.

M. D. : Bah, on clique sur la loupe, la petite vignette s'affiche et le curseur n'est pas automatiquement positionné dans la zone de recherche. C'est bête mais chiant [rire] et ce n'est pas... on ne peut pas faire de recherche par expression exacte. Donc parfois c'est moins performant et puis c'est tout petit sur la gauche enfin c'est... Quand je suis en télétravail, c'est juste impossible en fait de naviguer dans ce petit carré latéral. Et quand on cherche vraiment une expression exacte, il faut repasser par la recherche avancée avec la référence exacte du document et relancer la recherche par expression exacte dans le document. [Soupire] Euh. Et je passe sur les temps de réponse. [Laisse un blanc] Et j'aimerais savoir comment fonctionnent les rebonds qui sont proposés dans la partie gauche justement, si ça marche par des sélections ou si ça marche par un rebond sur les mots clés, l'indexation ?

Fanny : Alors il y a les auteurs, les sujets et l'encart qui s'appelle description. Et ça, c'est fait de manière...

M. D. : Et comment se fait la sélection ? Parce que des fois il en présente que deux ou trois ?

Fanny : Voilà, des fois, ce n'est pas très logique. Et des fois effectivement il va présenter, il va faire un rebond vers la description alors qu'il aurait pu faire un rebond vers l'auteur. Ça c'est un peu obscur de mon côté, je pense aussi de ton côté [en regardant Arnaud] je ne sais pas, c'est pas trop à notre niveau, je pense que c'est. Bon en tout cas moi j'ai juste posé la question de la même manière que toi et je ne sais pas trop si on a la main sur pousser les auteurs en particulier ou autre. Donc ça s'applique en fait, ça fait partie aussi des réflexions pour savoir est-ce que ça rentre dans le moteur de recherche ou pas ?

Arnaud : Ouais, moi je n'ai pas de précision, ça fait partie des choses qu'on va demander au DSI pour mieux comprendre comment ça fonctionne.

Fanny : Le but de tout ça, c'est aussi qu'on arrive à avoir une différence entre ce qui est du moteur de recherche et ce qui est du paramétrage moteur de recherche. Et ça, Arnaud vous l'expliquera bien sur comment il fonctionne aujourd'hui sur le moteur de recherche et la manière dont il fonctionne et comment on le paramètre.

Arnaud : Et puis aussi la manière dont ça s'affiche.

Fanny : Et la manière dont ça s'affiche.

Arnaud : C'est des calculs de pertinence, donc sur ces paramétrages de pertinence, il va choisir les trois plus pertinents qu'il va afficher. Mais après, quels sont les critères du calcul de pertinence ? C'est là où ça devient... C'est compliqué à comprendre.

Fanny : En tout cas les rebonds sont sur les auteurs, les sujets et ce champs description qui moi je trouve, est celui qui est le moins pertinent quand il est affiché. On ne sait jamais ce qu'on va avoir.

M. D. : Bah il est alimenté par des tas de champs divers et variés.

Fanny : Ok.

M. D. : Voilà, c'est tout.

Fanny : Merci. Et Arnaud, est-ce que tu as des questions ?

Arnaud : Non, c'était très riche, mais sur certains cas d'usage très précis, peut-être qu'on pourra regarder ensemble, pour bien comprendre où sont les ... J'allais dire les blocages, en tout cas les difficultés pour faire un retour. Parce qu'un de nos enjeux aussi, c'est de s'appuyer sur

l'expertise des des collègues pour ensuite avoir des cas d'usage et expliquer, et voir avec le DSI pour voir comment régler certains paramétrages pour améliorer le résultat. Donc ces cas d'usage, ça nous intéresse, parce que ça nous permet d'avoir des exemples concrets d'évolutions fonctionnelles.

M. D. : Ok !

Fanny : Et tu as parlé beaucoup de la recherche avancée. Est-ce que tu utilises la recherche simple ou ... ?

M. D. : Oui quand j'ai des recherches basiques ou comme tu dis exploratoires.

Fanny : Il y a une personne qui veut prendre la suite ?

X. L. : Je veux bien continuer, X. L., je suis chef de produit data.bnf au département des métadonnées. Alors je suis très content d'avoir entendu parler de web sémantique parce que c'est un peu, c'est un peu mon quotidien. Je vais parler en tant que fonction, mais aussi beaucoup pour des cas d'usage que j'ai pu identifier en tant que personne faisant du service public. Et puis ce que j'ai sondé un peu, alors pas tout le département des métadonnées, mais mes collègues les plus proches, qui sont les membres de l'équipe de correction du catalogue général et qui ont des attentes très spécifiques. Alors les usages dont je veux parler sont des usages effectivement assez spécifiques de recherche d'information plutôt que de recherche de constitution de corpus par exemple. Moi quand j'ai besoin de faire un corpus pour dire bah tiens, ça serait intéressant qu'on essaie de proposer des visualisations de données à partir des collections de la BnF, je vais passer par data.bnf. C'est mon réflexe initial. Quand j'ai l'occasion d'échanger avec des chercheurs qui ont ce genre de besoin aussi, je vais leur proposer de passer par data.bnf par soit par l'interface, soit par des requêtes SPARQL qui permettent d'aller très loin dans les données. Mais voilà, moi ce dont je peux parler, le type d'usage dont je peux parler, c'est vraiment des recherches précises d'information. Par exemple, en service public, y'a pas très longtemps, une personne m'a demandé de trouver le numéro précis du périodique *La Fronde*, fondée par la Marguerite Durand dans lequel il était fait mention d'une personne qui était récipiendaire d'un prix. Et elle savait où. Elle savait que c'était la fronde. Elle savait le nom de la personne, mais elle savait plus trop l'année où à un ou deux ans après. Donc en fait, là, ça a été très utile pour moi de pouvoir faire une recherche plein texte. Donc j'ai dû commencer par rechercher dans tous les numéros. Je ne sais plus si j'ai cherché d'abord dans tous les numéros ou dans une ou deux années, puisqu'on peut chercher les numéros listés dans les listes de résultats. Et on a pu assez rapidement finalement identifier le numéro et le passage où il était fait mention de ce prix.

En service public pour la recherche plein texte sur les périodiques, en particulier c'est un outil très intéressant. Mais encore une fois, c'est une recherche très ciblée. On va chercher une granularité d'information qui est dans le texte et qui est dans un document précis. Côté collègues de l'équipe de correction du catalogue, évidemment leur usage premier c'est de passer par le catalogue et d'accéder au document de Gallica par le catalogue. Mais il y a un cas d'usage auquel je ne pensais pas du tout, par exemple, une collègue qui utilise la recherche par proximité pour trouver les dates biographiques de personnes en recherchant prénom, nom, nécrologie, décès, mort, etc. Même genre de recherche avec la recherche par proximité pour désambiguïser quand on a deux homonymes et qu'on veut les distinguer prénom, nom et noms de métiers de l'un, nom de métier de l'autre. Et on fait des sondages comme ça, ce qui permet d'identifier des personnes précises. Autre cas d'usage qui me vient là plutôt d'une, parce que comme beaucoup de collègues, j'ai été enseignant avant. Si, si, je sais que la BnF propose 10 millions de documents. Je peux m'attendre à ce qu'elle propose aussi des grands textes littéraires en Epub. Donc je vais vouloir trouver des Epub. Et donc par exemple, je vais vouloir trouver plus précisément les lettres portugaises en Epub. J'essaie de chercher les lettres portugaises dans la recherche simple, je ne trouve pas tout de suite. Alors les lettres portugaises ça ne me donne pas grand-chose. J'ai trop de bruit. Si je cherche "lettres portugaises Guilleragues", j'ai encore un peu trop de bruit je crois et si je cherche Guilleragues j'en ai deux. Comment est-ce que je fais la différence ? Je clique sur le premier et j'ai une liste d'occurrences, de documents numérisés qui sont effectivement des lettres portugaises. Ok, si je remonte sur l'autre, je relance ma recherche, je prends la deuxième suggestion. J'ai aussi une autre liste, donc peut-être pas complètement différente d'ailleurs des documents numérisés qui me proposent des lettres portugaises de Guilleragues. Mais comment est-ce que je m'y retrouve dans tout ça ? C'est un peu un peu compliqué. Alors si je cherche un Epub, je vais avoir tendance à aller vers la recherche avancée. Il y a des fonctionnalités si je me souviens bien, pour pour filtrer, mais il faut, il faut avoir identifié précisément l'ouvrage que l'on veut filtrer ensuite par la recherche avancée. Et c'est compliqué de savoir rapidement, sauf à passer par les sélections 150 Epub identifiés par l'Education nationale. Je veux savoir est ce qu'il y a un Epub ? Pour quel type de documents, sur quelle période ? On trouve facilement des Epub...

Arnaud : Dans la recherche avancée, on peut chercher par format en fait, c'est un peu enfoui mais on on peut filtrer par Epub.

X. L. : Oui mais il faut le savoir, il faut, il faut le savoir, il faut passer par là. Et ça marche bien quand on a déjà identifié assez précisément ce qu'on veut, avec ou sans Epub. Voilà, pour avoir

une vue sur le corpus d'Epub disponibles dans Gallica, je suppose que si on fait une recherche, ça doit sortir la liste mais que je n'ai pas testé.

Fanny : Et après oui effectivement, c'est dans la recherche avancée. Après, j'aurais pu vous le dire dès le début, c'est que là les dernières statistiques que l'on a, la recherche avancée, c'est 3 % de l'usage. [Rires collectifs] J'aurais peut-être dû le dire au début ? C'est pour ça que vous êtes experts même si vous n'êtes pas experts.

M. D. : En gros c'est nous quoi. [Rires]

Fanny : C'est 3% des usages et la recherche simple c'est environ 20%. Le reste c'est de la recherche qui passe directement par Google, on arrive sur le visualiseur directement.

M. D. : Ou le catalogue !

[Rires]

Fanny : Non c'est 3%.

[Rires]

M. D. : Je rêve ?!

Arnaud : Non, c'est équivalent à la recherche avancée.

Fanny : Et la recherche dans les emprises de la BnF est très très faible. C'est du 0,8 %, quelque chose comme ça. Donc nous on est sur une recherche extérieure avec, même étrangère parfois, et hors Ile-de-France avec des personnes qui en majorité arrivent sur Gallica via Google en arrivant sur un document ou alors en recherche simple, c'est 20/25 % et quand on est en recherche avancée c'est 3 %. Et donc la recherche par proximité, c'est bien que vous en parliez parce qu'on a eu les mêmes remarques lors du précédent atelier, beaucoup de personnes sont conquises par la recherche avancée alors que je pense que dans les usages c'est infime. Dans ces 3%, qui utilise la recherche par proximité ? Donc ça aussi c'est aussi cette voie-là d'expert, même si voilà, vous avez dit vous n'êtes pas experts, mais quand même en fait un peu.

X. L. : Et alors du coup justement dans la recherche simple, ce serait intéressant d'être un petit peu guidé peut-être d'avoir un élément d'explication, une aide contextuelle qui me dise bah voilà, c'est ça, qu'est ce qui va se passer ? Sans entrer dans les détails, bien évidemment.

J. C. Chaponneau : Moi, ça me rassure beaucoup ce que vous dites Fanny parce que j'avais l'impression d'être... En fait je suis J. C. Chaponneau, conservateur, chargé des manuscrits

médiévaux à la bibliothèque de l'Arsenal. Et puis là, j'avais l'impression d'être... Je me suis dit "Ah la la, moi, je ne fais pas du tout ces choses, ni recherche avancée, ni recherche par proximité." En vérité, on utilise beaucoup Gallica, mais on n'utilise jamais le moteur de recherche. En fait, c'est un peu le système. En tout cas, moi, d'après ce que je relève, de mes collègues à l'Arsenal, c'est vraiment la façon de l'utiliser. Dans mon cas en particulier même, je dirais que le moteur de recherche, c'est Google en vérité, l'utilisation principale de Gallica, c'est quand je cherche un manuscrit en particulier qu'il soit à l'arsenal ou dans d'autres départements. Ce qui est un peu troublant, c'est quand on tape la cote du manuscrit dans Gallica, c'est un peu compliqué pour finir par le retrouver et quand on le tape simplement dans Google, on met la cote plus "BnF" ou la cote plus "Arsenal" directement on tombe sur le manuscrit dans Gallica et ça c'est l'observation principale que je peux faire aujourd'hui. Je ne sais pas exactement ce qu'on pourrait y faire ? Mais c'est dans les usages qu'on a, que ce soit des manuscrits médiévaux ou des manuscrits plus tardifs (je parlerai des images après), mais c'est vraiment le besoin principal, ce serait de pouvoir accéder dans Gallica (qui est un peu utilisé comme une sorte de catalogue) directement par la cote à un manuscrit.

M. D. : De mémoire, ça marche bien avec les guillemets pour le coup.

J. C. : Eh ben moi pas toujours.

M. D. : En mettant "source : cote" en recherche avancée.

Fanny : On a eu cette remarque lors du précédent atelier. À la fois d'A. T. et à la fois des personnes qui font de la recherche iconographique qui savaient exactement en étant passées par le catalogue, que dans ce livre, il y aurait sûrement des choses intéressantes et effectivement, que la recherche par codes était moins facile. Mais du coup, je ne savais pas du tout que ça marchait sur Google.

J. C. : Je le fais quotidiennement.

Fanny : Ok c'est ...

M. D. : Moi j'utilise pas mal de sources côte avec le nom du département, ça ne marche en plus.

X. L. : Vous ne passez jamais par BAM ?

J. C. : Ah si si, bien sûr. Non, bien sûr, je passe par BAM, mais en fait, je pourrais quasiment m'en passer, on peut tout faire presque directement (pour les manuscrits). En fait, donc pour les usages, en fait, il y a trois types d'usages de Gallica : la première, c'est donc la recherche de

manuscripts pour des raisons, pour des questions de recherche iconographique, en vérité (Et donc là, ce qui est quand même... Comment dire ? Ça, c'est indépendant du moteur de recherche, en fait, c'est qu'on n'arrive pas à aller sur une image en particulier à partir du moteur de recherche, mais c'est normal à 100 %. Il faudrait qu'il y ait des liens entre Mandragore et Gallica mais ça c'est un autre, un autre sujet). Donc ça, c'est le premier, le premier point. L'autre type de recherche, en fait, qu'on fait, c'est (et c'est très embêtant cette question de la cote pour les manuscrits de type manuscrit Archives de la Bastille, sur les manuscrits qui ne comportent pas d'images) donc on cherche par exemple essentiellement les dossiers par exemple de prisonniers qui sont bien indexés dans le catalogue à partir du catalogue, on y arrive facilement avec la cote, c'est plus difficile quand on passe directement dans Gallica. Et là, tout ça, je ne sais pas en fait si... je ne sais pas mais par le moteur de recherche en fait on n'y arrive pas. Donc là, c'est là qu'il a l'intérêt de passer par BAM, c'est plus simple, pour les lecteurs, c'est plus compliqué. Et puis l'autre, le troisième cas d'utilisation, bah c'est la recherche plein texte avec les guillemets, c'est relativement efficace. Un exemple : oui, un exemple, là je n'ai pas tout compris, mais c'est une collègue qui m'a expliqué qu'elle avait été un service public. Elle était en service public avec une lectrice, une chercheuse qui cherchait donc dans les premiers Figaro, le premier Figaro, donc avant 1830, des caricatures. Le mot caricature et en fait elle cherchait une caricature en particulier. Mais la lectrice ne savait pas quelle caricature, elle avait le type de la caricature, mais elle n'avait pas le numéro du Figaro en question. Par Gallica. Elle a tapé "caricature" et puis elle a cherché dans l'ensemble des numéros. Elle y est arrivée. Mais ce qu'elle a, ce qu'elle a eu du mal à comprendre, c'est que quand on est dans les périodiques, quand on fait une recherche périodique, ensuite ils s'affichent par année, il y a une ligne, mais pas par numéro. C'est-à-dire qu'il y avait, je ne sais pas le 1^{er} janvier et le 1^{er} juin. Sauf que si on cliquait 1^{er} janvier, on avait l'ensemble des numéros des six premiers mois et si on cliquait 1^{er} juin, on avait l'ensemble des numéros des six derniers mois. Et du coup elle n'a pas compris. En fait, elle s'est dit "ils n'ont numérisé qu'un numéro au premier janvier et un numéro au premier juin". Et du coup, elle a laissé tomber. Mais après, elle a fini par comprendre mais la première impression...

X. L. : C'est là que la recherche dans tous les numéros, dans les numéros, dans la liste, c'est ce qu'il faut faire mais ça ne se voit pas très bien.

Fanny : Non, mais c'est vrai qu'au premier coup d'œil, on voit que... enfin pour un public... C'est pas évident. Peut-être qu'il faut le signaler sur le calendrier.

J. C. : Voilà.

Fanny : Mais oui, c'est effectivement un problème, merci J. C..

Arnaud : Juste, sur les archives de la Bastille, vous aimeriez avoir l'ensemble du fond ? Comment améliorer en fait l'accès ?

J. C. : L'ensemble du fond n'est pas numérisé. Il y a quelques types de dossiers qui sont numérisés...

Arnaud : Mais en tout cas une visualisation de ce qui est numérisé dans le fond ? C'est ça en fait le besoin ?

J. C. : Oui.

M. D. : Ah, ça je l'ai demandé. [Rires] Et j'en profite pour le recaser, j'ai demandé ça au comité OAI. Parce que les notices du catalogue général, quand un document appartient à un ensemble, il y a un rebond vers les notices du même ensemble. Et on a demandé pour BAM de pouvoir créer un rebond vers les documents numérisés dans un même instrument de recherche à partir d'un document donné et on ne l'a toujours pas. Parce que ça, ce serait vraiment extrêmement précieux.

Arnaud : De toute façon, on fait un recueil de besoins ? Mais je ne m'engage à rien, c'est vraiment...

Fanny : Ça dépasse nos, enfin ça dépasse mes compétences. Est-ce que J. C. ou X. L., vous avez d'autres choses à ajouter ?

X. L. : Oui, c'est ce qu'a dit tout à l'heure Arnaud sur le besoin de convergence de fonctionnalités entre les outils de recherche BnF. J'ai une remarque sur le fait qu'à l'heure actuelle, quand on veut filtrer par date de publication dans Gallica, ça s'appelle une date d'édition et c'est partiel alors que dans le catalogue qui s'appelle date de publication et c'est entre 1900 et 1903, enfin c'est par tranche d'années quoi. Ça perturbe les personnes qui jonglent entre les différents moteurs. Après c'est des choix de vocabulaire qu'on rencontre. On le voit aussi avec quand on décide qu'on va nommer telle facette ainsi, est ce qu'on prend en compte prioritairement le grand public ? Est ce qu'on prend en compte prioritairement les professionnels des bibliothèques ?

Fanny : Non, mais alors effectivement, il y a la question de l'appellation date d'édition, de publication etc. Mais il y a aussi que, je pense qu'on pourrait être plus précis, sur le XX^e siècle par exemple avoir, par exemple moi j'imagine très bien (et vous avez vu ça sur tous les sites de

e-commerce), mais une fourchette. Les baskets entre tel prix et tel prix, c'est assez facile d'avoir une fourchette de dates dans les facettes et du coup, de sélectionner les bornes chronologiques des documents qu'on veut afficher.

Arnaud : Ce qui est compliqué avec la date, c'est qu'il y a beaucoup de documents qui ne sont pas datés dans Gallica parce qu'on n'a pas la date de publication, ce sont des estimations. Il y a aussi qu'elles ne figurent pas dans la tranche chronologique.

Fanny : Et d'ailleurs, tu n'en as pas parlé...

J. O. : Les dates avant Jésus Christ. [Rires]

Fanny : Oui, on l'a en tête. Oui, oui, les dates avant Jésus Christ, on le sait. Il y a plein de documents qu'on ne peut pas rechercher.

J. O. : Oui, il faut se retourner le cerveau pour les retrouver.

Fanny : Oui oui, on a eu la remarque, on l'avait en tête. Effectivement, la date il y a des gros, enfin des petits paramétrages à faire. Est-ce qu'il y a une personne qui veut prendre la suite ?

S. L. Lévêque : S. L. Lévêque donc je suis chargée de numérisation au département littérature et arts. Et ça ne fait pas longtemps que je suis sur ce poste. Et je vais vous décevoir, mais je n'utilise jamais les moteurs de recherche Gallica, je passe uniquement par le catalogue général en fait parce que mon travail en fait, c'est plutôt de sélectionner une notice une charger... Et en fait moi je vérifie dans le catalogue général si c'est numérisé et après j'ai l'exemplaire numérisé. Donc en fait, moi je travaille uniquement comme ça. Les rares fois où je suis allée sur Gallica en fait. Alors la recherche avancée est une catastrophe en fait, je n'ai jamais trouvé ce que je cherchais et du coup c'est pareil en service public, je passe systématiquement par le catalogue général.

Arnaud : Qu'est-ce que vous cherchez en particulier ? Vous avez un exemple ?

S. L. : Euh non. Alors bon, en public c'est plutôt sous très souvent des périodiques, ça, ça va, c'est assez facile. Et après moi, dans mon travail de tous les jours, je charge des notices d'imprimés. Du coup, moi c'est plutôt de l'imprimé...

Fanny : Et donc dans le catalogue général, vous cherchez par quoi ?

S. L. : Alors par cote, beaucoup cote, auteur, titre, beaucoup par cote en fait. Moi il faut que je regarde si c'est numérisé ou pas, mais après je vais sur Gallica. Pour les périodiques, je regarde

les années et du coup en fait moi je n'utilise jamais le moteur de recherche en fait. C'est pour ça, quand on m'a demandé de participer à des ateliers, j'ai dit d'accord, mais je ne suis pas du tout une spécialiste. Alors dans le département on a déjà eu des formations, on m'a montré plein de trucs formidables pour trouver justement des fonds spécifiques qu'on a. Mais je veux dire un lecteur, il ne peut pas trouver, ça c'est sûr. Moi, si je ne le fais pas tous les jours, j'oublie et je sais plus non plus. Et voilà, enfin quand on me parle de recherche par proximité tout ça. Moi je n'en fais pas du tout en fait. Quand je vais sur Gallica directement, je fais de la recherche simple, c'est souvent des choses sur un auteur, un titre voilà. Et sur le catalogue général, je suis sûre de trouver ce que je veux et en dernier recours je vais sur Google. Voilà "ma recherche Gallica" et souvent on trouve ce qu'on veut en fait. Voilà, c'est un peu décevant. Je suis désolée.

Fanny : Non, non. C'est intéressant et vous êtes pas du tout la seule à utiliser le catalogue général et je pense que c'est un usage interne, tout a sa place en fait.

S. L. : Mais même du coup au lecteur, en général, lecteur en service public je... En fait, même moi je ne pense pas à aller sur Gallica, en fait, je vais directement sur le catalogue général.

Arnaud : Du coup je vous faites pas de recherche en texte ?

S. L. : Non non non non, je n'en ai pas l'utilité en fait, enfin dans mon travail de tous les jours.

Fanny : C'est vrai que c'est peut-être plus en service public ? Je ne sais pas...

S. L. : Oui.

M. D. : Après, ce qui n'est peut-être pas clair pour tout le monde, y compris nous-même professionnels. C'est la distinction entre ce qu'on peut trouver dans le catalogue général, et dans Gallica, en fait dans le catalogue général on ne trouve pas tout parce que dans Gallica c'est beaucoup plus large, il y a BAM, les partenaires moissonnés. Il y a ça et ça je ne sais pas qui a conscience de ça. Peut-être 3 % aussi des utilisateurs.

S. L. : En revanche, je peux utiliser BAM. Par contre quand je sais que ce sera archives.

J. O. : C'est vrai que c'est étonnant parce que finalement, le seul point d'entrée unifié de tous les catalogues... de pratiquement tous les catalogues (c'est Google) et c'est éventuellement Gallica pour ce qui est numérisé, alors même que la plupart d'entre nous on ne se rend pas compte du nombre de catalogues différents avec des points d'accès différents. Peut-être qu'il manque d'abord un point unique, parce que ça c'est un problème que nous-mêmes on expérimente. On a du mal à... [rires]

Fanny : Mais effectivement et c'est vrai que les collections partenaires ça représente aujourd'hui à peu près 20% des 10 millions. Et là, dans les derniers chiffres, les collections partenaires sont à peu près consultées à 15 %. Donc en fait, elles sont quand même assez consultées. Donc c'est aussi pour ça que c'est un usage externe qui prédomine et que les gens utilisent le moteur, effectivement pas dans les mêmes typologies de recherche que nous en interne, mais c'est aussi intéressant parce qu'on est tous utilisateurs et utilisatrices de Gallica. Il faut prendre en compte aussi cet usage-là, même si c'est très professionnel. Vous vouliez prendre la suite ?

L. A. : Alors oui, L. A., chargée de collections pour le XVI^e siècle à la réserve des Livres rares. Il m'arrive de répondre aux questions des lecteurs via Sindbad ou la messagerie électronique, et je travaille aussi à la coordination du prêt aux expositions. Et donc, ces deux points d'entrée, les questions des lecteurs et le prêt aux expositions amènent un certain nombre de cas d'usage de Gallica. Ce qui nous intéresse, nous, à la réserve en tant que fond patrimonial c'est d'utiliser Gallica, non pas pour la recherche plein texte, très peu pour la recherche bibliographique, mais bien pour la recherche d'exemplaires de substitution à des copies physiques. Ce qu'on remarque de façon quotidienne, c'est que les lecteurs n'ont absolument pas conscience de la différence entre exemplaires et éditions. Moi, je parle vraiment d'édition imprimée donc pas du côté des manuscrits, donc d'édition d'exemplaires répétés pour une même édition. Mais Gallica ne permet pas du tout au lecteur d'avoir cette lisibilité sur le fait que là, il est en train de consulter UN exemplaire de TEL fonds qui appartient à ... donc évidemment, si on clique sur description détaillée et qu'on va chercher dans les toutes petites lignes la cote du document, la source, etc, on sait d'où elle vient, mais le lecteur n'a absolument pas cette conscience-là. Et donc il arrive très souvent que lecteur ou emprunteur nous demandent un tel exemplaire parce que c'est celui qu'il a vu sur Gallica et qui n'est pas du tout conscient que d'autres exemplaires peuvent tout à fait se substituer, et même beaucoup mieux, à celui qui est présent sur Gallica pour sa propre recherche. Voilà. Et donc c'est presque un constat (je partage complètement ce que vous avez dit sur le fait que le lecteur n'a pas conscience de tout ce qu'il peut en voir, que le Gallica complète ce qui figure au catalogue général), mais c'est aussi le contraire. C'est-à-dire que Gallica ne donne qu'une vision très parcellaire de ce que nous conservons à la Bibliothèque nationale. Et donc les lecteurs arrivent en croyant avoir fait le tour (les lecteurs et les chercheurs novices et de plus en plus de jeunes étudiants) ...

XX : Je confirme

L. A. : ...arrivent en ayant, en croyant avoir fait le tour des exemplaires qui sont conservés à la BnF parce qu'ils ont fait une recherche, même peut être une recherche avancée sur Gallica. Et ça vient aussi du fait d'une, je pense d'un manque de lisibilité dans l'affichage, du bruit aussi qui résulte quand on fait une recherche simple du fait que les recherches plein textes, enfin les résultats d'une recherche plein texte sont mélangées avec les recherches de volumes, donc on a à la fois les volumes et les citations de ce volume dans tous les outils qui sont à disposition. Et ça, pareil, je trouve que c'est assez mal indiqué, mais c'est ce qui est, c'est ce qu'on trouve aussi maintenant dans les bibliothèques universitaires, avec des outils de découverte, etc. Ça donne un bruit très très fort qui est difficile à gérer pour un débutant. Et ça, gêne vraiment le lecteur. Ah oui, et cette chose aussi que vous avez -je crois- soulevé, le fait que quand on fait une recherche avancée (j'ai quand même pas mal travaillé la recherche dans Gallica parce qu'on prépare une rencontre). Et en fait, si on ne décoche pas (par défaut, ce qui est coché c'est "afficher dans tous les ..." enfin tous les volumes s'affichent sous le même intitulé avec la même vignette unique, alors que ce sont des exemplaires qui ne sont pas tous les mêmes). C'est... en fait j'ai mis du temps à trouver cette petite case qui permet d'afficher les volumes en liste. Et ça, c'est vraiment un problème. Parce que Gallica, par essence, c'est quelque chose qui montre les exemplaires et qui ne montre pas les éditions. Donc masquer tous les exemplaires d'une même édition, sous le même intitulé, voilà qui est absolument, extrêmement trompeur. Je trouve que mettre ça par défaut, c'est compliqué. Alors je comprends que ça puisse avoir un intérêt pour les périodiques, mais pour les autres, pour les éditions, pour nous, c'est catastrophique. Parce qu'en plus, les premières numérisations qui apparaissent, enfin celle qui est affichée en vignette c'est, j'ai l'impression, la plus ancienne numérisation qui est donc la numérisation à partir du microfilm et donc c'est catastrophique. Ce n'est pas celle sur laquelle vont cliquer les utilisateurs.

Arnaud : Par défaut. Les résultats sont regroupés au titre, donc non pas l'exemplaire mais le fascicule. Ce choix a été fait pour clarifier la liste de résultats pour les utilisateurs. Effectivement pour les périodiques, il peut y avoir énormément de fascicules différents affichés. Alors si vous souhaitez avoir la visibilité de l'ensemble, vous pouvez effectivement le décocher dans la recherche avancée et grouper/ dégroupé. Soit quand vous exportez les résultats dans le rapport de recherche, vous l'avez forcément en mode dégroupé et donc les résultats...

L. A. : Et par la recherche ? Si on fait par la recherche simple ? C'est automatiquement groupé ? [Arnaud acquiesce] Et ça je trouve que c'est vraiment un vrai problème quand on est dans Gallica, enfin quand les gens utilisent Gallica pour chercher des titres et des exemplaires, aussi

pour chercher des exemplaires qu'ils ont vu dans un article de recherche, etc. Ils tapent le titre, ils tapent l'auteur et ne tombent pas sur cet exemplaire-là, alors que parfois il arrive qu'ils soient numérisés, donc c'est un vrai problème. Nous, la conclusion, c'est que les lecteurs, on les incite toujours à passer par le catalogue général. On ne peut pas leur dire de faire une recherche par Gallica parce que ces paramètres-là, ces fonctionnalités, elles sont beaucoup trop loin dans le moteur de recherche pour que ce soit facile, alors que l'interrogation du catalogue est quand même relativement fiable par le système de filtre, de tri, ...

Fanny : Du coup, cette présentation regroupée, évidemment qu'on ne peut pas s'en détacher.

L. A. : Évidemment, mais il pourrait y avoir une distinction pour les exemplaires parce qu'on ne décrit pas les monographies comme on décrit les périodiques. Ça devrait pouvoir se... enfin je ne sais pas mais dans la liste au Père Noël, on devrait pouvoir le modifier.

Arnaud : La distinction que je fais mal dans votre exemple c'est entre les volumes et les exemplaires. C'est normalement regroupé au volume par exemple l'encyclopédie où il y a dix volumes mais une seule entrée et là, en fait... ?

L. A. : Non, en fait ma recherche c'était ce matin : Dante, Landilo, 1481. J'ai cherché notre ZYD17 qui est numérisé et le premier qui apparaît c'est une numérisation d'un microfilm d'un autre exemplaire mais pas celui-ci. Et donc du coup je ne comprenais pas pourquoi je n'y arrivais pas. J'ai essayé toutes les manières possibles de chercher et pour qu'il sorte jusqu'à ce que je tombe sur des petites cases en bas. C'est une pratique très très fraîche de la recherche avancée de Gallica.

M. D. : Et il apparaissait pas du tout sinon ?

L. A. : Et sinon il n'apparaît pas, il apparaît sous...

Fanny : Une fois qu'on a cliqué ?

L. A. : Oui, une fois qu'on a cliqué sur la première ligne comme la presse. C'est un vrai masque en fait. C'est impossible... enfin, pour un lecteur non averti, c'est impossible de deviner. Je veux dire, moi j'ai cliqué par hasard en disant ah en fait c'était là ?

Fanny : Ok.

J. O. : La remarque de L. A. disant que beaucoup de lecteurs qui consultant Gallica et considèrent que tout est là et qu'il est vraiment importante et ça vaudrait sans doute la peine de dire que c'est 10 millions, c'est sur je ne sais plus combien de centaines de millions de

documents à la BnF ? C'est une fraction des collections de la bibliothèque et sans ça, c'est pas du tout intégré normalement.

Fanny : Oui oui, normalement, dans la refonte qu'on est en train de penser, il y aura une page sur la collection de la BnF, la collection de Gallica par rapport à la collection de la BnF à visualiser en disant qu'est-ce que ça donne ? Après, c'est très complexe parce que on parle de la collection de la BnF en général ? Est-ce qu'on parle de la collection de la BnF libre de droits ? En même temps sur Gallica, tout n'est pas libre de droits. Donc on a vraiment... La comparaison est difficile à faire, il faut voir...

L. A. : Mais sans faire de comparaison, inviter les lecteurs à aller consulter les catalogues de la bibliothèque, enfin je ne sais pas, mais même intellectuellement, c'est important pour qu'on sache que tout ce qui est numérisé ne représente qu'une toute petite fraction du savoir qui est conservé à la bibliothèque. C'est vraiment, mais c'est fondamental et c'est une inquiétude vraiment sur la manière de faire, les recherches, etc de voir que les gens considèrent que ce qui est numérisé, c'est l'existant et que tout le reste, tout le reste, n'a pas d'existence.

J. O. : C'est peut-être provocateur mais je me demande s'il ne faudrait pas qu'il n'y ait pas de moteur de recherche dans Gallica. [Rires]

Fanny : Mais oui, je comprends.

L. L. B. : Et aussi, certaines ressources, comme on dit, ne sont pas librement accessibles. Je pense aux périodiques de Retronews dont les résultats s'affichent dans Gallica.

XX : Ça c'est un autre problème.

L. A. : C'est un problème de compréhension majeur de la part des lecteurs. Ils ne comprennent pas qu'ils aient pu être affichés un jour et qu'aujourd'hui ils ne le soient plus. Pour certains vieux lecteurs ...

Fanny : Ils le redeviendront.

L. A. : Ils le redeviendront mais dans les questions des lecteurs c'est "Je ne comprends pas pourquoi ?"

Fanny : Oui, surtout que des fois ça rejoint des collections partenaires que l'on a et qui sont aussi indiquées en étant sur Retronews. C'est ça dépasse notre service, c'est des politiques BnF-Filiales. Après ce qui est sûr, c'est que au bout de dix ans, ces périodiques retomberont dans

Gallica mais le fait qu'ils soient affichés dans les résultats et qu'on vous renvoie vers un autre site qui n'est pas accessible... C'est un problème.

L. L. B. : C'est une bonne publicité.

XX : Gratuite...

L. L. B. : C'est intéressant de savoir que c'est là, quelque part. On ne sait pas comment y arriver, mais on sait que ça existe. C'est bien déjà bien en théorie.

M. D. : C'est accessible, il faut juste créer un compte gratuit.

Fanny : Si on a le lien direct, oui. Après si effectivement on a perdu le lien direct, et qu'on veut faire de la recherche, ça c'est payant. Oui effectivement, enfin, c'est gratuit d'accéder à la ressource si on a le lien, après la recherche avancée est payante. Et par rapport au moteur de recherche de Retronews, on retrouve souvent la remarque de personnes qui ont des abonnements, qui y ont accès via leur université hein, il faut se dire que c'est aussi accessible via les universités. C'est que le moteur de recherche était plus performant. Oui, parce qu'il y a une typologie des documents : la presse. En fait, c'est toute la difficulté du moteur de recherche de Gallica. C'est la diversité documentaire.

M. D. : Oui, et ils font tout un travail sur les entités nommées, etc qu'on ne fait pas etc. Donc forcément, c'est plus performant puisqu'on parle de la presse. Est-ce qu'il y aurait un moyen (parce que justement, quand on affiche le calendrier, quand on est dans Gallica, on voit le calendrier des numéros qui sont libres de droits ou assimilés) et si on veut voir ce qu'il y a dans Gallica Intramuros, il faut changer l'URL pour aller voir le calendrier plus complet, etc. Il n'y avait pas moyen de mettre un petit bouton en voir plus dans Gallica Intramuros ? Qu'on sache qu'il y en a plus.

Fanny : Oui le lien, ça c'est quelque chose qu'on va travailler je pense après la refonte de la page d'accueil pour voir de quelle manière on fait ces liens entre Gallica et Gallica Intramuros aussi pour que Gallica Intramuros soit plus consultée, parce que ce n'est pas si consulté que ça en interne, enfin pas à notre niveau mais au niveau lecteur on va dire. Et je pense qu'il y a une incompréhension de ce qu'est Gallica Intramuros, ce n'est pas très très clair en fait. Donc oui, j'ai déjà pensé de la même manière que peut-être... déjà uniformiser peut-être les deux pages d'accueil, avoir un message qui dit clairement qu'on est dans Gallica Intramuros. Mais c'est très complexe parce qu'on ne peut pas dire que l'une est libre de droit, l'autre n'est pas libre de droits parce qu'on a des documents qui sont sous droit de mention sur Gallica.

M. D. : Et j'ai une question pour les gens qui font du service public, dans le portail documentaire, quand un lecteur ... c'est quoi qui s'ouvre, c'est Gallica ou Gallica Intramuros ?

X. L. : T'as tout, t'as toutes les entrées à plat.

XX : Il y a des briques.

M. D. : Oui mais justement, quand on demande Gallica, il y a les deux carrés où il y a un seul carré ?

XX : Et je pense qu'il y a le carré Intramuros. Parce que comme c'est uniquement en salle.

M. D. : Bah oui, du coup, ce serait plus logique qu'on arrive directement Intramuros puisqu'on est sur site et pas dans le Gallica, enfin voilà.

X. L. : Alors oui mais je crois qu'il y a des trucs qui sont Gallica et qui ne sont pas dans Intramuros.

Fanny : Non,

X. L. : Non, ce n'est pas possible, ok. Il n'y avait pas des ... ?

Fanny : Non, ou alors c'est une erreur. [Rires] Non, vraiment, il faut imaginer un selon la version enrichie de Gallica avec des documents sous droits qui ne sont pas accessibles dans Gallica parce que c'est sous droits. Mais normalement non, après on a du sous-droits sur Gallica aussi, c'est des conventions, c'est grâce à des conventions et des partenaires qui ont pu signer des conventions aussi avec des ayant-droit. Vous avez sûrement vu ces dernières semaines le musée national du sport, il y avait plein de dessins sur le sport et en grande partie c'était un dessinateur qui est décédé dans les années 1960 et qui est encore sous droits mais avec lequel ils ont une convention et ils sont donc sur Gallica.

Arnaud : Les dessins de Plantu, les fiches de lecture de Michel Foucault.

Fanny : Le théâtre, les manuscrits d'André Breton, ...

X. L. : Puisqu'on a parlé du portail documentaire. J'ai quand même le sentiment qu'on a une problématique d'ensemble pour toutes nos applications au vu de l'offre foisonnante qu'on propose à la BNF, qui est de se repérer, enfin de proposer un endroit où on explique aux lecteurs qu'on peut trouver telle chose là, telle chose là... C'est un vieux sujet. On y a repensé pour data récemment parce qu'on est en train de revoir la page d'accueil. On s'est dit ben voilà, il faut qu'on trouve un moyen de montrer clairement que dans data, on peut trouver ceci, cela et que

le reste se trouve à un autre endroit. Ce portail documentaire est une tentative pour essayer de mettre à plat l'ensemble de l'offre documentaire BnF, mais je ne suis pas sûr que ce soit suffisamment clair encore. Et est ce qu'il faut dupliquer ce genre d'initiative sur l'ensemble des applications ? Je ne sais pas, je n'ai pas de réponse tirée du chapeau, mais en tout cas, c'est quand même une problématique récurrente qu'on entend pratiquement toutes les interventions que de dire ben voilà, ici vous allez trouver ceci et pas cela, il faut aller là si vous cherchez cela.

Fanny : Oui, oui, il n'y a pas d'uniformisation. Je pense que le travail qui va avoir lieu sur la page d'accueil de Gallica, qui est un travail technique, ergonomique et graphique. Ça va être aussi un travail de communication dans le sens : ok, ce texte que l'on a aujourd'hui, qui sont dans le footer de Gallica, la collection de Gallica, c'est des textes qui ont 10 ans, qui n'ont jamais été revu parce qu'il faut qu'on explicite un peu plus franchement ce qu'est Gallica, ce qu'on y trouve, mais ce qu'il y a aussi dans la constellation de la BnF parce qu'il y a plein de choses. Et très souvent on a la question. J'ai eu la question assez récemment d'iconographes et vraiment foncièrement, c'est quoi la différence entre les collections de la Banque d'images et les collections Gallica ? Qu'est-ce que je vais trouver ? Qu'est-ce que je ne vais pas y trouver ? Et en fait, c'est des questions très spécifiques, mais c'est vrai que c'est écrit nulle part. En tout cas, je pense qu'il faudrait que ça soit porté par l'établissement sur les pages de la page de la BnF dans les pages d'informations. Mais nous, sur le terrain, il faut qu'on soit plus clair sur le site de Gallica.

M. D. : C'est typiquement, le genre d'information que personne ne va lire. [Rire] Je ne voudrais pas être pessimiste, il faut le dire ok, il faut que ce soit très visuel en fait, un genre de cartographie...

Fanny : Pour la collection, on va essayer de faire quelque chose de vraiment très visuel parce qu'il existe aujourd'hui des informations sur la collection avec Gallica en chiffres, ce sera différent après c'est beaucoup de blabla aussi.

L. L. B. : En fait, il faudrait un Gallica pour les nuls... [rires]

Arnaud : Justement, c'est ce qu'on voudrait éviter en fait. On veut garder un accès de Gallica pour tous, toutes et tous. Peut-être un mot sur la représentativité de la collection numérique, là, vous en avez parlé tout à l'heure et c'est vrai que pour les chercheurs aussi, c'est important d'avoir... Alors, c'est un sujet qu'on avait discuté avec X. L., parce que ce qu'on voudrait, c'est arriver à savoir, à avoir une projection de ce que numérisé part par rapport à ce qu'on a dans les catalogues, il y a de manière et... en fait c'est très divergent, différents en fonction des

collections des départements. Donc, par exemple, nous, on estime qu'il y a 10% des collections de la BnF qui sont dans Gallica, c'est les éléments qui sont donnés en formation. Et après ça a été très différent selon les départements, c'est seulement 3% ou 4% des titres de presse, mais c'est 80 % des manuscrits en tout cas, qui ont une très large ouverture du microfilm.

Fanny : Donc c'est 10 % du libre de droit ou 10% de l'ensemble des collections pas libres de droits ?

Arnaud : De l'ensemble de la collection. Sachant qu'il y a 10% de la collection dans Gallica Intramuros.

Fanny : et qui sont amenés à être versés demain sur Gallica un jour.

Arnaud : Donc effectivement on essaie de voir comment... ça fait partie de notre réflexion c'est donner à voir ce qui est numérisé parmi l'ensemble de l'offre en tout cas de ce qui est signalé dans les catalogues.

J. O. : C'est aussi beaucoup une question d'affichage, on devrait pouvoir expliquer aux gens ce qu'est-ce que c'est que Gallica, mais c'est aussi parce que c'est là le principal point d'entrée pour n'importe quel lecteur, et même toujours un peu loin quand on pense venir à la BnF, on va penser Gallica d'abord. Après, on a aussi beaucoup communiqué sur le fait que dans Gallica il y a tout, enfin de tout en tout cas un foisonnement documentaire, etc. Du coup, on ne sait pas très bien ce qu'il y a dedans et simplement sur la page d'accueil dire les grandes typologies de documents par exemple les monnaies [rires] pas hasard, enfin voilà il y a des monnaies, de la presse etc avec des gros boutons peut-être avec une jauge en disant (même si c'est vague), juste pour dire : il y a 10% de la presse sur un bouton, il y a la presse, cliquez là, vous avez la presse et c'est quelque chose, 10 % de la presse, il y a les modèles, puis la moitié, c'est que de l'affichage, personne n'ira forcément utiliser ça. Mais au moins on dit de manière claire ce qu'on trouve et ce n'est qu'une partie des collections. Et pour moi c'est essentiel parce que effectivement, la perception n'est pas du tout celle-là aujourd'hui.

L. L. B. : Il faudrait des histogrammes, des jauges...

J. O. : évolutifs !

Fanny : Est-ce que les personnes qui n'ont pas parlé veulent prendre la parole ? Il nous reste un quart d'heure.

T. L. Laurent : Je vais dire un petit mot qui va aller dans le sens de ce qu'on a beaucoup entendu. J'ai envoyé, c'est pour ça que j'ai pris mon ordi. Comme je ne vais pas m'attarder là-dessus, j'ai envoyé à Arnaud. Alors je vais me présenter, donc je m'appelle T. L. Laurent et je suis responsable de collections en sociologie au département PHS. Et donc voilà, j'ai récupéré un certain nombre de ... enfin, j'ai consulté mes collègues qui m'ont donné un nombre de cas d'usage assez précis, que du coup j'ai envoyé à Arnaud hier et je ne vais peut-être pas revenir dessus. Je pensais le faire mais on n'a pas beaucoup de temps et on pourra en reparler si vous avez besoin d'avoir des précisions, je trouve que c'est des cas assez intéressants. Mais la principale réponse que j'ai eu de la part de mes collègues et qui correspond (je dois dire) à l'essentiel de mon usage aussi, c'est que les gens passent par le catalogue général. C'est un grand classique parce que je pense que pour deux raisons : d'une part parce qu'il est très bien, très structuré, on sait exactement à qui on a affaire. Et puis aussi tout simplement pour des raisons d'habitudes. C'est-à-dire que voilà, le moteur de recherche de Gallica est mal connu et en fait, de mon point de vue. Mais en fait, je rebondis sur ce que vous avez dit tout à l'heure, à savoir il faudrait que les pages moteurs de recherche. Moi, mon idée c'est qu'en fait il faudrait qu'il y ait que le moteur de recherche avancée. Je me dis que, en fait, le moteur de recherche dans Gallica, il est intéressant à partir du moment où on commence à utiliser les fonctionnalités précises, ce qui ne veut pas dire qu'il n'est pas sans défaut même à ce niveau-là, mais au moins là, il... on trouve des choses intéressantes. Je pense en particulier que la recherche par proximité que j'ai découverte grâce à Arnaud, eh bien je crois que je ne connaissais pas ou je n'utilisais pas. Je trouve qu'effectivement c'est un usage intéressant pour le texte intégral et qui parfois permet de contourner des difficultés d'ailleurs signalées par certains de mes collègues sur certains usages ou sites combinés. Et quand on fait des recherches en combinant certains mots, je pense à un exemple d'un collègue qui qui cherche, qui faisait une recherche assez précise sur le vinaigre et qui cherchait "vinaigre" et "Maille" et il n'y arrivait pas parce qu'en fait y faisait une recherche en combinant plusieurs fois plusieurs critères. Et en fait, on tombait systématiquement sur "vinaigre", mais "maille" disparaissait. Et en fait, dans la recherche par proximité, j'ai fait le test et là par contre, ça fonctionne parfaitement. On a les deux et les deux éléments qui s'appliquent. Donc voilà, c'est des détails comme ça. Mais tout ça pour dire qu'effectivement j'avais le sentiment qu'en fait, ce qui serait vraiment intéressant, au-delà de toutes les demandes qui ont été faites, c'est en fait, c'est d'assumer le côté assez complexe de la recherche dans Gallica et donc de proposer d'emblée non pas une recherche simple parce qu'en fait elle, elle n'est pas satisfaisante, mais de proposer d'emblée une recherche avancée qui soit une recherche, donc d'une certaine façon experte, mais surtout une recherche guidée. J'en avais un petit peu

parlé à Arnaud parce que je sais qu'il y a des y a des comment dire ça ? Je pensais à quelque chose qui dans un univers mais complètement différent et qui n'a rien à voir. En plus. Je sais pas du tout comment ça a évolué, mais à une époque où je m'occupais de ressources électroniques, il y avait un quelque chose. Parce qu'on connaît bien au DEP qui est Factiva et effectivement il y a une recherche experte, je ne sais pas, et il y a une recherche guidée pour aider vraiment les lecteurs et pour les accompagner à tous les moments de la recherche. Je ne sais pas trop ce que ça pourrait donner en cas de Gallica, mais je me dis qu'il y aurait quelque chose à faire dans cette direction-là. Et moi, pour ma part, les usages et usages que j'ai de Gallica, c'est principalement dans deux temps, dans deux cadres de mon activité. C'est d'une part pour la constitution d'un corpus, donc je fais un corpus sur les origines de la sociologie. Et là effectivement, dans ce cas-là, je passais presque toujours par le catalogue parce que je cherchais des auteurs en particulier, et je voulais repérer ce qui était, euh, comment dire, déjà numérisé dans Gallica, ce qui ne l'était pas. Et dans ces cas-là, bien évidemment, c'est le catalogue qui m'aidait à le faire et sinon, j'ai un usage, je dirais presque ludique de Gallica, c'est qu'en fait, ici, une fois par an, je fais une intervention auprès d'enseignants pour leur présenter en fait les ressources de la BnF. Donc dans le cas d'une intervention, peu importe et. Et donc dans ces cas-là, je fais une présentation. En choisissant des exemples sur un thème particulier qui est défini à l'avance par le cadre dans lequel j'interviens. Et et en fait, là, dans ces cas-là, bon, bah je dirais que, on est en plein dans la sérendipité, je cherche vous raconter des choses comme ça, par proximité, trouver des images de tel ou telle, et là je bricole en quelque sorte et là ou Gallica me convient. Mais c'est vrai que je vois bien que la limite, la difficulté intervient dès que l'on cherche quelque chose de précis dans une collection ou chez un auteur. Mais ce qui revient beaucoup chez mes collègues, c'est effectivement le problème du bruit. En plus, je sais par Arnaud d'une certaine façon, d'où vient la difficulté. Les collègues s'attendent en fait à trouver un terme précis, par exemple et parfois il y a des termes proches. Je n'insiste pas là-dessus puisque vous...

Fanny : Oui, on va en reparler tout à l'heure. A. H. ?

A. H. Heuqueville : Oui, très rapidement sur la recherche exacte, quand on cherche un terme ce n'est pas forcément facile pour l'utilisateur de savoir que si on n'est pas dans la recherche exacte, il va avoir par exemple, pour mère d'un enfant, il y aura aussi mer des démons. Un autre assez amusant, c'est pour Citron si on fait dans toute la notice on va avoir Citroën. Donc je ne sais pas. Comment ça marche ? Il y a aussi sur l'assemblage de la troncature étoile dans Gallica parce qu'effectivement, souvent on jongle entre le catalogue général où on peut faire la

recherche par troncature et Gallica où on peut faire la recherche plein texte, mais il n'y a pas de possibilité de l'utiliser. Donc là, on en avait pour un projet au Datalab, on avait un projet sur la forêt, donc on ne peut pas faire "Forest" + troncature donc pour "Forest", ancien terme pour la forêt, forestier, forestière et toutes les déclinaisons dont foresterie et donc ce qui peut compliquer la recherche quand on va faire par exemple, on cherche des textes juridiques, donc on va faire "arrêté" ou "code" ou "loi" ou "décret" et forest toutes les déclinaisons. Donc ça ce n'est pas possible. Il faut passer effectivement par des voies contournées, soit limiter aux sources droits, économie, politiques, soit aux thèmes droits et société. Ça peut exclure des documents qui sont conservés au département PHS mais qui seraient juridiques ou on peut faire un export avec les rapports de recherche et du coup faire une recherche contient les termes qu'on recherche. Donc même dans le même style, donc avec forêt toujours, il y a parfois des termes significatifs comme "exploitation", "aménagement" ou "culture". Donc là on est obligés de décliner toutes les associations de termes type "sylviculture" ou "foresterie", ou "exploitation forestière" ou "exploitation en forêt" ou "exploitation de bois", etc etc. Mais on ne peut pas faire simplement une série de termes et "forêt". Et donc sinon, beaucoup de beaucoup d'espoirs et d'attentes concernant la recherche image puisque contrairement à d'autres collègues pour les estampes, de notre point de vue, on a l'impression que la recherche justement d'estampes, de photographies, d'affiches, de cartes et plans est beaucoup plus facile dans Gallica, par rapport à effectivement, nous on a des ouvrages avec de belles planches, etc. La presse illustrée, scientifique et médicale, mais ça c'est absolument impossible, très difficile à repérer, alors qu'effectivement on a... Donc là c'est un exemple simple, un lecteur en salle qui cherchait des planches anciennes représentant des agrumes. Donc si on fait une recherche "agrumes" en limitant les images dans Gallica, on ne va pas du tout trouver les bons ouvrages illustrés avec des planches anciennes. Donc voie de contournement mais qui pour un utilisateur qui ne connaît pas, bon ... Donc nous on a un parcours Gallica, la nature en image qui recense les plus beaux, les plus beaux ouvrages illustrés en botanique, en zoologie, en minéralogie, etc. Donc c'est une voie qui permet d'identifier des ouvrages illustrés, mais c'est pas du tout possible actuellement de le faire à partir de Gallica et de dire je cherche, je cherche des représentations de singes, une espèce de singe en particulier ou ce genre choses et voilà, c'est vrai que voilà des projets dont on a l'écho nous laissent espérer. Voilà, ça ne sera peut-être pas forcément tout de suite et pas non plus tout de suite et de façon exhaustive. Mais c'est vrai que voilà, on aimerait beaucoup qu'une recherche sur soit une recherche thématique large ou très précise puisse on puisse associer avec "cette page contient des images".

Fanny : Ça c'est, pour vous en dire deux mots, c'est un projet qui s'appelle Gallica images, ça ne sera pas mis en production avant 2026, ça demande du temps. Donc ce serait fait via de l'intelligence artificielle, on a un prestataire qui a été retenu. C'est fait. Ce qui sera le plus facile, c'est évidemment de commencer par les corpus qui sont déjà des typologies d'images. Parce que du coup ce sera beaucoup plus facile de reconnaître sur une photographie qu'on a un homme avec un parapluie par exemple, et d'entraîner la machine sur cette typologie. La machine n'aura pas besoin de comprendre que c'est déjà une image parce qu'on lui donne comme information. Et ensuite, oui, l'idée, c'est que ça puisse s'étendre à la fois aux imprimés, à la presse et potentiellement aux manuscrits par la suite. Mais tout ça, ça va demander du temps d'affinage. Parce que la représentation, je prends l'exemple de l'agrume d'un citron dans un manuscrit du Moyen-Âge, dans la presse en noir et blanc et dans un ouvrage illustré de botanique, ne peut pas être le même. Et puis ça pose des soucis, dans une page de presse, par exemple. On peut par exemple page, un article dans lequel il y a une photo, et sur cette photo, il peut y avoir un tableau, par exemple qui est représenté, quelle est l'image qu'on sélectionne ? Dans quoi on sélectionne en fait ? A quel degré d'image ? Mais ça, c'est un projet. Et je pense que les premiers corpus seront déjà des corpus iconographiques, des affiches, etc. Mais on l'a bien en tête et oui, ça aiderait tout le monde de faire une recherche par image.

M. D. : Et ça pourrait être couplé avec les projets sur les modèles de langues qui peuvent reconnaître justement identifier dans le texte qu'il y a autre chose que du texte ?

Arnaud : Ça, ça fait partie des évolutions du moteur de recherche, on pourra en dire un mot tout à l'heure, on est dans un processus avec deux grandes étapes : une première étape sur qu'est-ce qu'on peut améliorer dans le moteur de recherche actuel ? Et puis après on se pose la question de l'évolution, voire du remplacement du moteur de recherche à d'autres outils. Et on a la première phase, c'est qu'on a demandé à un cabinet d'expertise et de conseil un l'état de l'art sur les moteurs de recherche aujourd'hui, et justement les nouvelles fonctionnalités qui sont permises par les grands modèles de langue. Mais c'est ce qu'on utilise en partie pour la fouille d'images. Il y a deux étapes aussi. Ce que disait Fanny, c'est une première étape de reconnaissance de formes. Quelles sont les formes à l'intérieur de l'image et comment, par similarité de forme, peut produire des résultats équivalents ? comme vous le trouvez dans Google Images. Et donc ça, c'est un entraînement d'image. Puis après, la deuxième étape, c'est d'identifier l'image à l'intérieur du document ou dans la presse, dans les manuscrits, dans les monographies, ... Et ça, c'est... On devrait quand même avoir les premiers prototypes à tester cette année sur la reconnaissance de forme. Et cette reconnaissance de forme étant couplée avec

l'indexation pour pouvoir mieux cibler et notamment avec les légendes, les images qui seraient dans les monographies.

L. L. B. : Il y a un autre niveau aussi, sans doute peu d'informations qui serait à explorer. Là on parle d'expérience du Département des Estampes parce qu'on est confrontés à des illustrateurs aussi dans les livres et on est face à des strates de catalogues qui n'ont pas fait sortir les illustrateurs. Ce n'était pas la tradition aux imprimés. Et là, je pense qu'il y aura un travail aussi à faire avec un autre niveau que peut être un travail sur les données. Les données des entrées, des accès à ces illustrateurs. Peut-être un travail à faire sur les contenus textuels des pages de titres pourquoi pas, pour encore faire ressortir ces informations. Et ça permettrait sans doute aussi de faciliter la recherche d'images, de cibler en tout cas les documents susceptibles d'en contenir.

Arnaud : C'est sûr que pour les recueils, c'est un exemple tout à fait pertinent pour cette première étape de la fouille d'images, d'identifier quelles sont les images à l'intérieur des recueils. Parce qu'elles ne sont pas décrites en fait. Elles sont décrites dans le catalogue général ... ?

L. L. B. : Non, non plus. Donc on les voit ou on voit par exemple apparaître des noms d'illustrateurs pour quelques ensembles d'imprimés dans les inventaires de graveurs par exemple, qui pointent vers des éditions de l'ancien département des imprimés. Mais finalement, c'est très très limité et on ne l'a pas eu en retour, on ne bénéficie pas en fait des accès qui auraient pu être sortis dans ces catalogues du département des imprimés.

Arnaud : Oui. Il y a encore deux personnes.

I. C. Carme : Comme je suis l'adjointe de la cheffe du service numérisation et à ce titre, je rejoins pas mal de choses qui ont été dites sur le fait que toutes ces collections de la BnF ne sont pas dans Gallica, que ce serait bien de les signaler par un iceberg digital ou par n'importe quel moyen visuel ou wiki. Je représente aussi le département de la conservation qui a des usages extrêmement variés. Donc je vais essayer de parler de retours de collègues qui n'ont peut-être pas été évoqués ici. Un peu en vrac, des retours sur des filtres de recherche qui ne sont pas toujours adaptés aux différents supports. On a par exemple de l'atelier de restauration des maquettes qui me parle du fait que les maquettes pliées sont typées images, mais que l'intitulé n'est pas forcément pertinent car ça se rapproche plus aux objets qui apparaissent ensuite comme maquettes en volume, ce genre de dédale et puis de changement de typologie qui peut porter à confusion. Le fait que les textes sont parfois difficiles à trouver. C'est vrai que j'ai fait, j'en ai fait l'expérience pour la recherche dont on m'a parlé, ma collègue. Si on cherche le traité de

peinture, par exemple... enfin la découvrabilité des collections, de toute manière, je rejoins les collègues n'est pas toujours optimale. Pour traiter de peinture avec ou sans guillemets, on n'a pas du tout les mêmes choses avec guillemets, on a même des Bottins qui sortent et on a des encyclopédies, c'est tout à fait étonnant quand même. En essayant de lire un peu plus, un peu plus loin dans les recherches, on ne comprend pas toujours comment ça se passe. C'est donc l'idée d'une recherche guidée pour être intéressante avec avec toujours le risque de diriger des recherches toujours vers les mêmes, les mêmes pratiques qui ne sont pas forcément adaptées pour tous les usages. Des collègues de Sablé (?), je pense que ça devait être des personnes qui s'occupaient de la numérisation d'images m'ont fait part de plusieurs souhaits sur les recherches images fixes, notamment le fait de pouvoir rechercher par date, les images fixes sont apparemment uniquement recherchables par l'année, ils aimeraient bien évidemment pouvoir les trier par mois par jour, comme c'est possible pour les périodiques, et puis aussi avoir la possibilité d'une recherche sous forme de calendrier comme c'est également possible pour les périodiques. Après moi j'avais une interrogation en tant que personne qui encadre un expert OCR, qui cherche, qui s'occupe des marchés ou de la chaîne d'océrisation en interne. Je m'interroge toujours sur la manière dont sont visibilisées les collections océrisées ou pas. Il me semble que ce n'est pas toujours, ce n'est pas toujours évident même pour nous quand on essaie de voir si un texte a fait l'objet d'une océrisation ou pas. Et je ne parle même pas du fait que comme toutes les campagnes de numérisation, le degré de précision, l'OCR est différent selon les marchés qu'on a passé, selon l'historique de développement des technologies. On a, il me semble un petit macaron, mais en fait visiblement ça ne marche pas à tous les coups, parfois on peut faire des recherches plein texte dans un dans un ouvrage qui n'est pas marqué comme ayant été océrisé. Et de même, ça serait bien, ça serait bien pour les personnes qui utilisent la recherche avancée, d'être conscientes du fait que quand on fait une recherche de termes par proximité qui est un usage qu'on utilise aussi ou une recherche texte intégral, on va avoir uniquement les résultats de collections qui ont fait l'objet d'un traitement complémentaire. Et ça, encore une fois, c'est un petit peu ce que disait L. A. tout à l'heure. C'est quelque chose dont les usagers n'ont pas vraiment conscience qu'au fur et à mesure et selon les recherches qu'on fait, en fait, on se limite dans des collections. Déjà de par le fait que Gallica ne contient pas toutes les collections de la BnF, par le fait qu'elle contient des collections qui ne sont pas uniquement celles de la BnF, mais aussi par des modes de recherche. On arrive finalement à avoir des résultats qui ne sont pas représentatifs, même de toutes les collections qu'on a numérisées. Et voilà, je pense que je vais m'arrêter là au cas où, parce que je crois qu'il y a d'autres personnes.

Arnaud : Alors peut-être une précision sur l'OCR. Alors effectivement, Jean-Philippe Moreux, collègue expert scientifique de Gallica, a travaillé sur des outils de visualisation qui nous permettent de voir ce qui a été ocrisé ou pas dans les monographies et dans la presse donc le taux d'ocrisé est assez bon, on est à 80 % d'OCR des monographies et 80 % des OCR de la presse avec l'ambition de Sébastien, avec la conversion rétrospective de l'OCR d'arriver à 100 % d'OCR d'ici cinq ans. On a quand même un objectif de couverture maximale de l'ocrisation des collections (de monographies et de presse). Concernant la recherche plein texte, en dans les documents non ocrisés, en fait, ça c'était dans l'étape des matières puisque l'OCR commence à partir de 2005 à la BnF et avant 2005, ce qui a été converti en format enfin transcrit, c'était les tables des matières et donc on peut faire de la recherche plein texte à l'intérieur des tables de matières et on sait par des collègues au département de l'accompagnement de la recherche au DCA que c'est une modalité de recherche intéressante pour constituer des corpus parce qu'on peut cibler des mots clés dans les tables des matières. Donc c'est un usage pertinent. Et donc du coup, c'est vrai que ça peut créer une confusion puisqu'on a l'impression de faire de la recherche plein texte dans un document, mais c'est seulement dans la table des matières qui a été convertie.

I. C. : J'entends pour cette question-là d'objectif 100 % des collections, mais tout à l'heure on parlait aussi du fait qu'on a 20 % des collections qui sont issues de bibliothèques partenaires. Or, ces collections ne sont pas toujours ocrisées.

Arnaud : Effectivement, il y a deux cas.

I. C. : Et donc il y a aussi, ça ajoute aussi une couche, même si nous on a aussi nos objectifs, il y a cette proportionnalité des collections qui reste invisible et juste, j'ai oublié quelque chose d'assez important. On a fait remonter le fait que quand on cherche une monographie ou des périodiques, enfin une monographie en plusieurs volumes ou des périodiques, l'affichage n'est pas un affichage chronologique et donc ça perturbe énormément la recherche, par exemple, de quelqu'un qui irait voir un numéro du Figaro ou les numéros du Figaro qui ont été publiés en janvier 1872 et qui auraient le fascicule de juin, puis de décembre, puis de... enfin voilà, ça pose problème.

L. A. : Il faut retrier systématiquement.

I. C. : Oui, mais ça ne devrait pas être le cas.

M. D. : Ça c'est pour l'affichage la pertinence. Tu as dit Arnaud, 80 % des mon[ographies] et combien pour la presse ?

Arnaud : Pareil, 80%.

M. D. : Tu es sûr ? Ça me semble énorme.

Arnaud : Oui. C'est les chiffres et les chiffres apparaissent.

Fanny : Ça Jean-Philippe peut les donner précisément, mais moi j'avais l'impression que 80% c'était la presse et les monographies. Peut-être qu'il y a une disparité importante entre la presse et les monographies.

M. D. : Je veux bien, si vous avez moyen d'avoir le détail. Parce que nous pour la collection de référence, c'est un vrai enjeu pour les années à venir.

Fanny : Je crois qu'il nous reste une personne.

E. P. : Oui, donc moi je suis E. P., je viens du département cartes et plans et donc je m'occupe à la fois des programmes de numérisation du département, je suis catalogueur également et aussi pour la recherche bibliographique via le service Sindbad et le service public. Donc évidemment, dans ces différentes fonctions, je suis amené à utiliser Gallica. Donc, comme ce qui vient d'être dit, j'ai plutôt tendance à passer par le catalogue général parce qu'il faut chercher par la côte, faire une recherche iconographique. Mais comme également on ne trouve pas la même chose dans le catalogue général et dans Gallica, et il est bien obligé de chercher dans les deux. Alors je vais plutôt parler de la recherche géographique, puisque c'est ce qui intéresse principalement les documents cartographiques. Nous, ce qu'on aimerait dans l'idéal, ce serait d'avoir une sorte de modèle de carte interactive qui permette de naviguer sur une mappemonde pour trouver des documents, un peu comme ce qui existait pendant un temps sur data BnF, ce serait un modèle d'accès donc également l'indexation principale pour les documents cartographiques, c'est l'accès géographique, la toponymie. Toutes ces choses, c'est très important pour nous. Et puis également, c'est aussi la possibilité d'avoir une fourchette qui permette de sélectionner l'échelle du document. Et comme on avait parlé d'échelle chronologique à l'instant, un modèle qui permet de préciser.

Arnaud : L'échelle, c'est la taille ?

E. P. : L'échelle c'est ça permet de préciser la précision de la carte, donc d'une 5/1 000^e à 1/1000.

Fanny : Et dans la recherche aujourd'hui, des cartes est-ce qu'il y a des difficultés pour trouver ces cartes, est-ce qu'il y a des choses qui marchent bien ? Sans parler des évolutions possibles aujourd'hui, qu'est-ce qui ne marche pas, pourquoi vous passez par le catalogue ?

E. P. : Par exemple pour l'indexation géographique. C'est plus facile de passer par le catalogue général. Également, on a un autre type de document, ce sont les séries cartographiques. Donc pour accéder à ces séries cartographiques, ce sont des cartes en plusieurs feuilles et l'outil, c'est le tableau d'assemblage. Et ce système de tableau d'assemblage n'est pas en ligne actuellement dans le catalogue général et on a une base de données qui s'appelle Cartomundi et qui offre le même, cette même modalité de carte interactive qui intègre les tableaux d'assemblage et permet de trouver la feuille correspondant au lieu que recherche l'utilisateur. Voilà, donc ce genre de choses est vraiment indispensable pour les documents cartographiques pour améliorer leur accessibilité.

Arnaud : Je crois qu'on l'a fait pour la carte de Cassini, non ?

E. P. : Oui, voilà. Alors ça a été fait dans la précipitation. Ça a été bricolé. Exactement. Et évidemment, on n'a pas qu'une seule série, on a des centaines de séries. Donc, le modèle n'est pas forcément transposable. Euh, qu'est-ce que je pourrais rajouter d'autre ? Il y a aussi un aspect corpus, peut-être pas tout à fait développé, c'est pouvoir faire des recherches à l'intérieur de corpus. Par exemple, on a différents corpus et dans chaque corpus, il y a un classement géographique qui permet de cibler tel ou tel document. Et bien souvent, dans l'affichage Gallica, on a un affichage qui est aplani en quelque sorte et on perd en granularité dans la recherche. C'est un peu dommage pour retrouver ces corpus et voir comment ils sont articulés à l'intérieur.

Fanny : Merci, est-ce qu'on a fini notre tour de parole ?

[Passage au temps 2 et aux explications d'Arnaud sur le moteur de recherche]

Temps 2 : Explications moteur

[Explications du moteur de recherche par Arnaud]

Arnaud : [...] Donc là aussi y a une subtilité entre deux formes de recherche exacte qui même pour nous parfois et pas toujours facile à appréhender.

I. C. Carme : Oui, les guillemets ne marchent pas toujours, puisque quand je cherche histoires des "traité de peinture" entre guillemets, il me met très bien des encyclopédies, un bouquin commercial. Je ne comprends pas comment ce mode de recherche fonctionne sachant que c'est aussi ce qu'on apprend à l'Université, quand on est étudiant.

Arnaud : [...] Voilà pour l'amour, le schéma fonctionnel synthétisé. Donc je peux répondre à vos questions.

Fanny : [reformulation]

L. L. B. : Oui, j'ai une question sur l'affinage de la recherche qui ne m'a jamais convaincu. J'ai plutôt l'impression que ça vient compliquer les résultats. Qu'est-ce qui se passe ?

Arnaud : En fait ? L'affinage. Il y a plusieurs types d'affinage. La recherche, l'affinage par... pas dans le document ?

L. L. B. : Oui

Arnaud : L'affichage en fait relance le même type de requête dans les résultats pour les documents dans les notices et en dans le plein texte ça veut dire qu'effectivement, on sait aussi par les enquêtes menées auprès des usagers notamment par Irène, des difficultés de compréhension entre la recherche dans les notices et la recherche par le plein texte. Normalement, quand vous faites la recherche, l'affinage, vous cherchez les termes à la fois dans les notices, dans le plein texte, ça relance une requête dans la liste de résultats. Là aussi, ce n'est pas toujours pertinent. Et normalement, si vous mettez & vous pouvez rechercher des termes et ça peut être aussi utilisé. J'ai testé et constaté que ça ne fonctionnait pas bien italien normalement, ça fonctionne aussi l'affinage dans le rapport de recherche. Quand vous exportez un ensemble de documents, là aussi il y a bien sûr un affinage possible. Si vous avez des exemples précis, là aussi, ça nous permet de voir nos collègues du DSI en disant tiens, il y a une anomalie sur telle ou telle requête. Est-ce qu'on essaye de comprendre ?

Fanny : Est-ce que d'une manière générale, vous utilisez ces facettes. Une fois qu'il y a eu des résultats.

[Plusieurs testeurs acquiescent].

M. D. : Oui, je me dis ce qui serait sympa ce serait de pouvoir faire une sélection. Par exemple, si je veux à la fois les images et les objets, au lieu de le faire en deux fois, si je peux le faire en une seule ce serait bien.

Fanny : [Explique la suite]

Temps 3 : Améliorations

Fanny : Allez-y

I. C. : Nous on a noté cinq points, on n'a pas vraiment eu le temps de les noter par priorité. Mais peut-être qu'une des priorités qui ressortait, c'était une page ou un outil de médiation tout de suite quand on arrive avec une infographie sur les types de collections qu'on trouve dans Gallica, le pourcentage de numérisé, les différents catalogues vers lesquels diriger des lecteurs en fonction des collections recherchées.

On aimait bien aussi l'idée de recherche assistée qui pourrait permettre de sélectionner les filtres les plus pertinents et avoir une recherche qui correspond mieux aux résultats... à notre idée... Entrer les filtres excluants a posteriori pourrait permettre aussi de réduire les résultats pour s'approcher des résultats... des résultats attendus. On s'est aussi posé la question de l'avenir de la recherche simple et plutôt que de la supprimer, on pensait peut-être que la combiner, combiner cette espèce, cette barre de recherche avec des filtres simples par des filtres géographiques, chronologiques, etc. Puis dans un second temps, après une première recherche, permettre des filtres plus précis comme on a dans la recherche avancée avec des facettes un peu plus ergonomiques pourrait permettre de perdre le moins d'utilisateurs possible. C'est une idée comme ça.

Et puis on s'est aussi interrogés en dernier lieu sur l'ordre des résultats, qui semble être un des facteurs de découragement les plus fréquents, que ce soit pour les collègues comme pour les utilisateurs. Et donc une réflexion sur la manière de trouver le bon ordre selon les recherches. Donc c'est aussi ce n'est pas une solution clé en main, mais ça peut être un ordre chronologique, ça peut être un ordre d'exactitude des termes selon la recherche, recherche, titre, recherche, auteur ou d'autres critères, mais qui est aujourd'hui le bon ordre du fait de la multiplicité des critères reste assez obscure, y compris pour nous soi-disant experts de Gallica. [Rires]

Fanny : Merci beaucoup.

I. C. : Je ne sais pas si j'ai oublié de parler de trucs ?

M. D. : Si, on avait parlé aussi de rendre plus clair la sélection de "je veux une expression exacte" ou au contraire, "je veux utiliser les expansions" et étendre ma recherche.

Fanny : Donc pour vous, elles n'ont pas vocation à disparaître ?

M. D. : Bah non, ça peut être intéressant, mais quand on est averti qu'elles sont là et qu'on souhaite les utiliser ou pas.

Arnaud : On se demandait est-ce qu'il ne faudrait pas changer de paradigme, peut-être d'abord entrer dans une recherche exacte ensuite proposer des expansions ?

[Les participants acquiescent].

Fanny : C'est-à-dire inverser les choses, aujourd'hui on propose beaucoup de résultats est-ce qu'il ne faudrait pas proposer moins de résultats et aller plus loin ?

M. D. : Oui, c'est ça. C'est sûr qu'avec le changement d'échelle 1 million/ 10 millions, on est dans la logique inverse. Voilà.

Arnaud : C'est sûr.

Fanny : Donc l'expansion orthographique pourrait être proposée en option. En tout cas, on est arrivés à cette mini conclusion. Merci. Peut-être que L. L. B. ?

L. L. B. : Avec ça, c'est effectivement la question de la troncature.

Fanny : Oui.

A. H. : Et donc on était arrivés à la même conclusion ou position pour l'expansion, effectivement, est-ce que ce ne serait pas plus judicieux, maintenant de par défaut, il n'y ait pas d'expansion mais qu'on puisse quand même l'activer si on le souhaite. Ce que parfois effectivement, c'est des recherches tellement spécifiques sur une personnalité, un événement etc qu'il suffit qu'il y ait un caractère qui soit mal reconnu pour qu'on ne retrouve pas le document ou l'article dans Gallica.

Fanny : Parenthèse pour l'expansion orthographique, il faut aussi avoir en tête que ça concerne en général le plein texte, donc ça concerne certains types de documents. C'est aussi pour ça que -à mon avis- on est à l'inverse, ça doit être proposé en option, parce que par défaut, elle ne concerne pas tous les documents.

A. H. : Par contre, ce qui m'avait donné pour l'exemple "citron" donc j'ai fait le titre, toutes les notices et tout et Citroën ne ressortait que dans toutes les notices. Et donc du coup la recherche était plus pertinente dans tout. Parce que du coup il y avait 600 résultats dans toutes les notices et 250 dans tout le reste...

Fanny : Il faut voir le paramétrage, oui oui. C'est des commandes obscures, des paramètres
[Rires] Non mais ce que ne vous a pas dit Arnaud c'est que ce tableau de bord pour paramétrer le moteur est assez tentaculaire. Donc aujourd'hui, il y a des choix qui ont été faits il y a dix ans. Il faudrait tester le cocher décocher pour voir quelle est l'incidence sur le moteur de recherche. On en est là. Donc en fait, c'est pour ça qu'on se nourrit de tous ces petits "Ah bah si ça ne marche pas alors que ça devrait marcher parce que j'ai compris comment fonctionnait le moteur de recherche", pour remonter ces informations au DSI pour mieux paramétrer avec l'existant. Est-ce que vous aviez d'autres idées ?

L. L. B. : On s'est interrogés sur la présence d'illustrations dans les années futures ? En attendant en tout cas, une intelligence artificielle qui viendra repérer les illustrations sur les pages. Est-ce qu'il est possible d'apporter... Comment dire ? Peut-être qu'elles sont importées, mais en tout cas, des données sont disponibles dans le catalogue mentionnant les illustrations. A ce moment-là, est-ce que le format ne pourrait pas être enrichi de ces informations matérielles et qui permettraient d'affiner les résultats ?

Fanny : C'est-à-dire par exemple savoir si le document contient des images.

L. L. B. : Voilà.

J. O. : Je peux poser une question à ce sujet parce que... Est-ce qu'on peut le faire avec des ressources extérieures ? Je pense toujours à mes monnaies parce que nous on est moissonné par des tonnes de portails où il y a des intégrations iconographiques qui sont parfaitement bien en place et il y a un lien entre nos notices et ces outils. Est-ce qu'on peut récupérer en partie cette indexation ?

Fanny : Je ne sais pas ?

Arnaud : En tout cas pour identifier les éditions illustrées, la stratégie de contournement c'est de chercher dans la notice ill. Dans la description. Mais ce n'est pas un critère de recherche. Donc après pour la question du moissonnage. En fait nous on est moissonnés par... dans les entrepôts OAI et c'est depuis des entrepôts OAI. (En fait, ce que je n'ai pas dit dans le schéma, c'est qu'on a des connecteurs qui -au niveau de l'indexation, vont récupérer les données du catalogue général et du catalogue archives et manuscrits- et qui créent un OAI à partir duquel les données sont converties en Dublincore dans le format, qui est appauvri). Alors, est-ce qu'on peut remonter en fait au processus de génération de l'OAI ? Ça, c'est une des pistes. Mais moi je ne suis pas assez expert, je me tourne vers X. L.. Ce serait d'utiliser le MET qui est un format

enrichi. Mais je pense que dans le cadre de la transition bibliographique, on aura sans doute d'autres types de données. C'est encore projeté dans cet après j'allais dire, qui devrait quand même donner des données mieux structurées le choix du Dublincore. C'est le choix de l'interopérabilité de Gallica avec le web sémantique.

X. L. : Sur cette question, sur le MET je pense que c'est Vincent Caron.

Arnaud : Parce qu'on avait eu la question à laquelle je n'avais pas su répondre c'était : En quoi ça allait impacter Gallica le changement de la transition bibliographique ? Là, c'est vrai que c'est un point qu'on pourrait éclaircir avec nos collègues, notamment dans la perspective d'enrichir les métadonnées de Gallica.

M. D. : Déjà, il est peut-être possible de faire évoluer le mapping des formats vers le Dublincore pour des données qui seraient manquantes, qui n'ont pas été prises en compte à un moment donné et qu'il serait peut être possible d'ajouter. Après plus on a, plus on en ajoute, et le format du corps étant pas très structuré, ça fait des zones description justement qui sont foisonnantes et difficilement exploitables en indexation.

Arnaud : C'est ça.

Fanny : Est-ce que vous avez d'autres choses ?

X. L. : On a abordé principalement deux pistes principalement une avec des déclinaisons. L'idée, c'est qu'on essaie d'éviter d'avoir deux interfaces de recherche de modes de recherche en ayant une recherche unique. Ce qui ne veut pas dire forcément l'une ou l'autre. C'est prendre un petit peu des deux, probablement. C'est-à-dire que l'alternative, soit on proposerait rapidement, dès la recherche de filtrer dans les modes de recherche de nouvelles idées par exemple recherche dans le texte, recherche plein texte, recherche dans le titre, etc, ou d'autres critères encore. Soit on garderait une recherche assez simple et ensuite, au niveau de la liste de résultats, on proposerait là des mécanismes de sélection, de filtrage, etc. Donc voilà, l'idée, c'est qu'on puisse avoir une modalité unifiée d'entrée dans la recherche et qu'on puisse proposer qu'on puisse se faire proposer en fait des paramètres à activer pour affiner, élargir etc. L'autre axe de travail, c'était d'être un peu dans la même idée d'être, d'être accompagné en tant qu'utilisateur. C'est qu'on puisse avoir des petites bulles d'aide contextuelle pour nous dire ce que ça va faire quand on va cliquer ici ou qu'on va rentrer une clé de recherche. Tout ce qu'évoquait tout à l'heure, M. D., c'est quoi la différence entre site ou bibliothèque ?

M. D. : Je crois qu'il y a des bulles en plus.

X. L. : On peut trouver mais encore faut-il qu'elles soient visibles.

Fanny : Normalement, il y en a sur tous les encarts [rires].

Marilyne : Mais quand on lit ???

Fanny : Ce n'est peut-être pas clair. [Rires]

X. L. : Là aussi il faut prioriser, il ne peut pas y en avoir partout, certaines peuvent s'afficher d'office ou d'autres pas.

Fanny : De toute façon, on ne peut pas... La recherche ne peut pas convenir à tout le monde, il y a tellement d'usages différents. La recherche avancée c'est 3 %, la recherche simple c'est 20% et le reste arrive sur le visualiseur. C'est aussi ça qu'il faut prendre en compte et effectivement dans la partie bibliothèque et site, cet encart-là est extrêmement demandé par les partenaires parce qu'ils veulent pouvoir rechercher dans les collections leurs collections ou les collections que la BnF en lien avec leurs collections. Et puis pour le grand public, ce n'est parfois pas très compris, il y aussi ça qui fait que... Il y a des strates de recherche qui sont complètement différentes.

X. L. : Il y a un point qu'on n'a pas encore abordé, puisque beaucoup de gens passent par Google, c'est la question du référencement. Quel type de métadonnées on expose pour les robots et comment on fait en sorte qu'ils arrivent ? C'est un gros sujet, je ne sais pas si vous avez des pistes dessus ?

Arnaud : Mais en tout cas ce qu'on observe c'est qu'on est bien référencés dans Google.

M. D. : Visiblement ça marche bien.

Arnaud : Visiblement, vous passez par Google. Donc pour nous c'est positif, ça veut dire qu'effectivement le référencement naturel de Gallica dans Google est performant. Après, quelles sont les stratégies qui sont en œuvre derrière, ça c'est Matthieu Gioux qui pourra vous éclairer. En tout cas ça fonctionne plutôt bien, les chiffres ce sont 41% des requêtes/ des accès qui passent par Google.

X. L. : Mais est-ce qu'en faisant une recherche Google, on arrive sur les bons documents c'est aussi ça ? On arrive sur un document Oui, mais est-ce que c'est le bon ?

Fanny : Pas forcément je pense, pas forcément en fait. Pour l'exemple dit tout à l'heure, si quelqu'un tape un nom d'auteur et un titre, j'imagine, qui va tomber sur le nom de l'édition qui

a été la plus cliquée. Bon après c'est plus un point d'entrée dans Gallica, mais fait, ce point d'entrée dans 40 % des cas ce n'est pas le moteur.

M. D. : Je pensais à un truc dont on n'a pas du tout parlé, c'est aussi l'accès par corpus qui est malheureusement très mal mis en valeur alors que c'est ultra intéressant parce que pour le coup on est vraiment loin de l'accès déjà avec une médiation très forte.

Fanny : J'espère que ça va changer et que la refonte sera à la hauteur de ça. L'idée, c'est de... dans les autres chantiers plutôt côté médiation, de mon côté à moi. Donc il y a deux grands chantiers en cours : la refonte technique, ergonomique et graphique de la page d'accueil donc il va y avoir un nouveau graphisme, une nouvelle ergonomie, une nouvelle typo... Tout va changer. N'ayez pas peur du changement, tout va bien se passer. C'est un gros changement, la page d'accueil elle date de 2017 est dans le briefing, le choix des prestataires, ce qui était demandé, c'est de mettre encore plus en valeur les corpus et d'avoir des entrées qui parlent à un plus grand public donc en parallèle, on a fait un travail sur l'arborescence. Le deuxième chantier c'est que ces corpus sont extrêmement riches, il y a à peu près 12 000 pages de sélections à tous niveaux confondus avec des corpus très spécifiques qui entrent dans le détail avec des listes de sélection hyper fortes et des choses diverses. Ce qu'on a travaillé... La piste sur laquelle on se dirige, c'est de mettre de côté les entrées par types de documents parce que ça représente une certaine départementalisation des départements des collections. Et de rentrer plutôt par des thématiques qui parleront à un public plus large je pense. Et quitte à retrouver dans ces thématiques là des choses qui rassemblent des documents, imaginons des manuscrits, on en aura forcément des choses sur la littérature dans les manuscrits. Mais avoir des grandes catégories, à peu près une quinzaine avec des facettes qui existent déjà (droit économie et politique, littératures restera littérature, on ne va pas réinventer l'eau chaude). Mais il y a des thématiques par exemple les grands événements historiques. On pourrait y avoir la Première Guerre mondiale, la Seconde Guerre mondiale, qui sont aujourd'hui enfouies dans des niveaux inférieurs et l'idée c'est que l'on puisse proposer justement des accès avec moins de clics et une liste de résultats qui arrive plus rapidement que ce qu'on a aujourd'hui.

M. D. : Et je ne pensais pas à ces corpus-là. [Rires]

Arnaud : Tu pensais aux codes projets ?

M. D. : Oui c'est ça. Parce qu'il y a plein de choses aussi qui ne sont pas dans les sélections.

Fanny : Oui, mais ça mais ça, qui l'utilise ???

M. D. : A part nous, pas beaucoup de gens.

Fanny : Mais qui ??

M. D. : Oui mais aussi parce que ce n'est pas connu et c'est dommage !

Arnaud : Ce n'est pas connu et ce n'est pas non plus documenté, moi j'ai une liste. En fait, il y a vraiment des centaines de projets, mais on ne sait pas forcément quel corpus il y a derrière et ni même quel département a été en instruction.

Fanny : Alors il y a les codes projet et il y a aussi l'entrée par thématique, via la recherche avancée ??

M. D. : C'est quoi ça ?

Fanny : Et ben moi j'ai trouvé que deux personnes qui l'utilisent au sein de la BnF : Arnaud Laborderie et une collègue à nous. Mais ils sont attachés à cette recherche par thématique. Il y a donc une recherche qu'on peut voir dans la recherche avancée qui s'appelle par thématique et qui ne concerne que les imprimés et les imprimés sont classés par thématique. Donc on a...

M. D. : Et ça c'est le cadre de classement ?

Fanny : C'est la clé et c'est Dewey.

Arnaud : C'est par thème et par sujet.

Fanny : Donc on mélange, tout tout tout est mélangé. Et donc vous pouvez retrouver dans arts, loisirs, sports tous les imprimés en sport.

T. L. : Je vais revenir là-dessus. C'est assez troublant parce que c'est ça rejoindre le problème. C'est-à-dire c'est la différence entre le fonctionnement du catalogue et le fonctionnement... Parce que, en fait, dans le catalogue, c'est parce dans le catalogue, ce n'est pas en Dewey, c'est Rameau quoi.

Fanny : Oui ! Je suis d'accord.

T. L. : Et en fait, ce serait intéressant qu'il y ait une... parce que la recherche par sujet, elle est intéressante et on la montre souvent, notamment quand on fait des ateliers doctorants, ce genre de choses, on la montre aux utilisateurs et donc ils se familiarisent souvent avec ça. Et puis ensuite, quand on est dans comment dire, le dans le moteur de recherche de Gallica, on retrouve autre chose.

Arnaud : Dans le moteur de recherche de Gallica, c'est une recherche par sujet.

T. L. : Oui par sujets mais ??

Arnaud : Ça renvoie à Rameau.

T. L. : Ah d'accord.

Fanny : Oui mais les sujets oui, les thématiques, non.

Arnaud : Oui mais les thématiques c'est...

Fanny : Dans la recherche avancée, à droite, en-dessous.

T. L. : Oui, d'accord.

Fanny : Justement, c'est les codes projets et il y a une recherche par thématique et là vous allez avoir une table comme une table des matières qui s'ouvre avec les grandes catégories Dewey à l'intérieur, vous allez avoir...

M. D. : Je pense que ça correspond au cadre de classement.

Arnaud : Oui, c'est le cadre de classement.

Fanny : Qui, je pense, n'est utilisé évidemment par personne à part...

M. D. : C'est dommage parce que Nathalie se casse la tête à les ajouter alors utilisons-les !
[Rire]

Fanny : Mais ça ne concerne que les imprimés. C'est un biais.

T. L. : Et ce n'est pas inintéressant.

Fanny : Dans une recherche avancée, on ne comprend pas que cette partie-là elle ne soit que dédiée aux imprimés.

Arnaud : En même temps, elle n'est pas toujours très exacte. Je ne sais pas comment ça été pris en charge, ça a été fait par les prestataires de numérisation et donc ce ne sont pas des bibliothécaires qui les ont indexés. Donc là, on m'a toujours dit d'être assez prudent sur cette question, sur ce classement thématique qui est, on va dire, qui donne voilà de grands ensembles, mais qui il peut y avoir aussi des erreurs.

M. D. : C'est comme le cœur de classement de la bibliographie nationale, c'est un peu le même esprit, c'est une liste très simplifiée de ces grands indices Dewey qui correspondent aux grandes familles de disciplines.

Fanny : Alors on a dépassé de 5 minutes donc il nous reste peut-être quatre minutes. Est-ce que vous voulez nous parler de vos évolutions/ améliorations ?

L. A. : Alors nous, ça rejoint des choses qui ont déjà été dit. La première chose, ce serait d'améliorer effectivement le mapping pour récupérer le plus possible de données issues des catalogues afin d'améliorer l'indexation sur Gallica et donc de rendre ces nouvelles zones qui seraient mieux mappées interrogeables dans le moteur de recherche.

E. P. : Oui, et on voulait aussi, côté cartes et plans, améliorer la géolocalisation des documents, notamment via les coordonnées géographiques.

Arnaud : Alors c'est des choses qu'on a déjà essayé de faire dans Gallica. Je ne sais pas si vous essayez parfois la recherche... Comment elle s'appelle ? Géographique/ géolocalisée dans Gallica ? En tout cas on pourra regarder ensemble.

E. P. : Oui oui.

Arnaud : En tout cas c'est produit automatiquement donc ce n'est pas toujours très précis.

E. P. : Non.

Arnaud : Mais justement on a un outil qui permet aux usagers de pouvoir repositionner les documents. Mais ce ne sont pas des documents cartographiques en fait, ce sont des documents qui sont dits à Saint Dié et donc l'algorithme va proposer une recherche par Saint Dié mais on ne peut pas savoir si c'est en centre-ville ou si c'est là. Donc il y a déjà un amorçage de recherche géographique dans Gallica.

S. L. : Oui, peut-être avoir un interrupteur un peu visible pour présenter par volume. Et non par série, c'est ce que disait L. A., quand elle affiche, elle veut des exemplaires en fait précis.

L. A. : Et en dernier lieu, améliorer la partie informations détaillées relatives à chaque document sur Gallica. On peut rendre cet encart un peu plus clair et plus visible. Notamment le lien vers le catalogue qui est toujours un peu noyé dans plein d'autres liens et il faut toujours...

S. L. : Ark...

L. A. : Oui, enfin on ne sait jamais quel Ark, sur lequel il faut cliquer.

XX [chuchote] : ceux qui commencent par Catalogue ou Archives et Manuscrits.

L. A. : Non, non, il y a juste Ark donc on ne peut pas savoir...

M. D. : Ah non, l'Ark c'est celui du catalogue de Gallica, sinon c'est catalogue.bnf nin nin nin ou archives et manuscrits blablabla. Il est plus haut, mais il peut y en avoir plusieurs parce qu'il y a la notice d'ensemble et y'a machin et y'a bidule. Donc des fois il y en a plusieurs...

L. A. : Oui, enfin, il y en a quand même souvent deux ou trois qui sont noyés...

M. D. : J'avoue, j'avoue. Et ils sont noyés dans la masse.

Fanny : Et puis dans cette fiche, à cet endroit-là, il y a aussi la conservation numérique avec la propriété du document, ça peut être BnF si le document a été numérisé à la BnF et le document peut appartenir à un partenaire, mais ça peut être tout peut être partenaire. Et cette partie-là est assez complexe. En fait, je pense que c'est des choses de la même manière qu'on pourrait aller vers une recherche plus simple, on pourrait avoir une version très simplifiée dans laquelle le chercheur a vraiment besoin de savoir qui conserve numériquement le fichier pourrait déplier quelque chose de plus "expert" en fait. Parce qu'effectivement là, on a 10 lignes, ...

M. D. : Oui, on pourrait séparer ce qui relève du biblio, de la provenance et pourquoi pas même des informations techniques sur la numérisation qui ne sont pas présentes actuellement mais qui pourraient...

Fanny : Oui, le grand public mais alors la majorité du temps, ils ne savent pas ce que ça veut dire. Il y en a beaucoup, je pense qu'il faut aller vers du moins avec la possibilité d'en rajouter.
Conclusion !

Arnaud : Oui, très rapidement. Merci pour votre participation. [Explication de la suite de l'enquête]

Irène : A chaud sur ce que vous avez dit, l'intérêt c'est que l'ensemble des améliorations que vous avez évoquées couvrent à peu près 6 thématiques différentes qui vont du plus général avec des points de principe qui sont : on produit du Dewey, on produit du Rameau, on produit du cadre de classement, comment est-ce que ça se généralise à l'ensemble des activités de la BnF. Je crois que vous avez eu un débat sur les dates aussi, c'est qu'il y a une forme de principe et de normalisation à l'échelle de la BnF. Donc ça c'est du très général. Jusqu'à des formes techniques qui sont plus dans les couches basses du paramétrage et du référencement et qu'elles sont des métadonnées extérieures présentées à des moteurs de recherche ou des robots, qui est

certes une question d'amélioration qui va certes toucher les usagers à un moment mais qui n'est pas active pour les usagers eux-mêmes. Et c'est à nous de définir ces éléments, j'ai mis avec un triangle vert ce qui est la définition de la BnF par rapport à des choses qui ne sont pas déterminées par l'utilisateur. Et vous avez identifié des fonctions, des choses déterminées sur lesquelles c'est l'utilisateur qui a la main et qui va avoir une interactivité pour définir son mode de recherche alors qu'il y a d'autres choses qui sont absolument transparentes pour lui, que c'est à nous d'implémenter d'une manière ou d'une autre. Dans ce que vous avez identifié comme étant à la main de l'utilisateur, il y avait bien évidemment pouvoir choisir ce que vous recherchez exactement, le périmètre dans lequel on recherche (OCR ou pas ?) ce qui nécessite de l'information et la médiation pour que l'utilisateur ait la main et soit autonome pour le déterminer, il faut à la fois des interfaces qui vous donnent le choix et puis à des éléments l'explicitation qui éclairent du choix de ces éléments, enfin qui explicitent les conséquences de ces choix. L'autre élément sur lequel vous disiez que c'est à l'utilisateur de choisir, ce serait la hiérarchisation. C'est quoi le bon ordre ? C'est quoi le bon document ? Avec des questions autour de la chronologie ou de l'ordre chronologique des sujets. Après ça, on passe du principe à la technique, avec des éléments sur les médiations des pages d'infographies, sur les autres ressources, des informations sur les corpus et les codes projets de l'aide contextuelle, j'ai ajouté la notion de correction. Ce que Arnaud disait sur la géolocalisation. Est-ce qu'on intègre le fait que les autres participants... ? A côté de la médiation, il y a qu'est-ce que va être la page de recherche ? Ce qui n'est pas la même chose que des évolutions sur les pages de résultats ? Il y a l'idée d'une recherche unifiée, l'intégration d'informations plus complètes, de techniques de recherche avec la troncature et donc la possibilité de choisir l'environnement de la recherche et peut-être le moissonnage extérieur. Et puis les résultats de l'appel de recherche, c'est la possibilité d'avoir des filtres excluants, combinés, assistés en fonction des résultats (ça c'est encore autre chose), de pouvoir choisir si on veut beaucoup ou pas beaucoup, par exemple avec des illustrations, de pouvoir afficher par volume et non pas par série. Et il y a un dernier champ d'amélioration c'était autour de qu'est-ce qu'on présente autour du document et donc notamment les informations détaillées ou l'Ark, le retour catalogue. Voilà, comme ça vous avez tout couvert.

T. L. : Juste une chose (et après on va peut-être partir [rires]). Je me posais la question des rebonds sur Gallica, est-ce qu'on peut faire des rebonds sur des choses qui ne sont pas dans le rebond ? Dans le catalogue général, vous avez dix autres documents qui sont intéressants, qui ne sont pas numérisés. [Rires]

X. L. : C'est décevant, on va aussi avoir plein de demandes de numérisation.

J. O. : Non, éventuellement pour aller dans les salles de lecture tout simplement ! [rires]

M. D. : “Relancer la recherche dans le catalogue” ?

Fanny : Je crois que c’est possible d’ajouter un lien externe. Oui, c’est possible d’avoir un document. Après faudrait voir ce que ça concerne...

Arnaud : Et comment l’automatiser surtout !

Fanny : Comment l’automatiser ? Est-ce que c’est obsolète parce que les documents finalement sont numérisés et on renvoie vers le catalogue mais il n’y a plus lieu d’être ? Enfin, je ne sais pas, c’est une bonne question ?

M. D. : Sinon il y a data !

Arnaud : Si, techniquement, c’est possible depuis peu on voulait renvoyer notamment les pages sur de la TEI, un des usages c’était de pouvoir renvoyer pour les documents qui existent en TEI et qui n’est pas disponible en Gallica mais pour lesquels il y a des attentes des chercheurs.

J. O. : Après, pour la numérisation en cours. Dans le catalogue général, on peut faire une recherche de documents numérisés ou non. On peut simplement fouiller que les documents non numérisés.

X. L. : Il y a déjà des renvois vers un wiki source actuellement, non ?

Arnaud : Effectivement. Il y a un lien qui a été fait dans le cadre d’un partenariat avec Wikisources dans la liste de résultats en fait ça apparaît, avec un petit w.

X. L. : Oui

Fanny : Donc il y a beaucoup d’informations sur cette partie. [...]

T. L. : Juste, ça fait écho d’ailleurs, à peut-être que je t’enverrai un message. Je suis frappé des fois de voir je pense en particulier pour des documents audiovisuels des fois la seule date qui apparaît, c’est la date d’entrée dans Gallica, ce qui est évidemment très problématique. Je pensais à des, des documentaires qui datent des années 60 et pour des raisons qui tiennent, je ne sais pas exactement à quoi ça tient et je serais content de savoir. Je pourrais t’envoyer un exemple très précis...

Fanny : Est-ce que dans le catalogue il y a une date ?

T. L. : Et je pense que non. Il faudrait vérifier mais...

Fanny : Ah bah dans ce cas là...

T. L. : Mais non, bien sûr, mais c'est pourquoi ? Comment ça se fait ? Parce que, en fait, ils ont été entrés pour être directement versés, enfin je vérifierai juste en vérifiant, mais pour moi, il y avait quelque chose de ??

Arnaud : N'hésitez pas à me renvoyer d'autres exemples de cas d'usage.

[...]

Transcription de l'atelier métier 4 : 16 mai 2024

Temps 1 : Témoignages

Fanny : On a commencé l'enregistrement et on va passer à la partie des témoignages.

Arnaud : Peut-être signaler qu'on a un collègue du DSI, Jean-Christophe Pautre, qui nous a rejoint pour ce dernier atelier, avec qui on a travaillé sur la partie plus fonctionnelle que j'aurais l'occasion de présenter tout à l'heure.

Fanny : Est-ce que du coup, il y a une personne qui veut commencer, nous parler de ses usages de Gallica, les manières dont il utilise Gallica, le moteur de recherche, etc ?

C. R. : Je m'appelle C. R., je suis au CCFr. Quelque chose qui me vient tout de suite à l'esprit, parce que j'en avais beaucoup parlé aussi avec M. C. à l'époque, pour faire des recherches pour aller à l'aide du CCFR. Je vais beaucoup voir ce qui se passe dans les autres bibliothèques, et ce qu'ils ont sur Gallica, en marque blanche, etc. Donc je prends l'habitude d'aller dans les pages partenaires et voir ce qu'ils ont sur Gallica, pour un exemple. Ce n'est pas toujours aisé, il faut que je descende tout en bas. Et je me demande s'il n'y a pas une autre solution, parce que ce n'est pas... Je vais tout en bas, en fait, il n'y a pas de partenaire, et après je tombe sur les pages. Là, je pioche, je fais mon marché.

Fanny : Et donc ce que tu interrogues, là, c'est la collection uniquement des partenaires ?

C. R. : Oui, parce que moi, je prends le point de vue du partenaire.

Fanny : Mais tu sais ce que tu cherches déjà ?

C. R. : Pas forcément, ou non. Mais en fait, je me sers de... De ce que le partenaire a dans Gallica, pour ses collections, parce qu'on a travaillé dessus, etc. Mais je trouve que ce n'est pas aisé, même qu'il y a beaucoup discuté avec M. C.. Peut-être que je me trompe, peut-être que je suis de fausse route, peut-être qu'il y a un moyen plus simple, mais c'est vrai que je galère un petit peu dans le... Je prends le champ, la liste de toutes les bibliothèques et je cherche, je cherche.

A. A. : Et tu ne passes jamais par la recherche avancée, par le nom de l'établissement ?

C. R. : C'est moins pratique. Oui, parce que ça, simultanément... Mais pour vraiment avoir un résultat global de ce que le partenaire veut, par exemple, pour faire un truc cohérent, avoir un

point de vue d'ensemble, j'essaie d'avoir la page où il y a tout du partenaire. Je m'y prends peut-être mal, je ne sais pas, mais je trouve que ce n'est pas direct.

A. A. : En effet, actuellement, en fait, on n'a pas la possibilité de combiner les requêtes sur les collections mentionnées et les collections intégrées. Ce qui fait que le seul moyen d'avoir une vue globale, c'est de passer par des pages de preuve.

C. R. : Tout en bas, mon petit. Je ne sais pas si...

D. D. S. : Dans le sitemap ? Dans le sitemap ? Dans tous les liens qu'il y a tout en bas ?

Fanny : Oui, oui, oui.

C. R. : Oui, tout en bas, là.

Fanny : Oui, tous les partenaires.

C. R. : Il y a une petite liste en colonne.

D. D. S. : OK. Donc, pour l'arborescence de tout ce qu'il y a dans le site, quoi.

C. R. : Voilà. Après, c'est les pages partenaires qui importent, que moi, je mets à jour grâce à Pierre-Emmanuel. Donc ça, leur rôle est important, également, c'était le travail de Maryse.

Fanny : Mais le moyen le plus direct, en soi, normalement, il faut... Je ne devrais pas répondre à la question. C'est la recherche avancée, de chercher dans la recherche avancée l'encart avec les partenaires et donc...

C. R. : Oui, il y a une liste qui apparaîtrait.

Fanny : C'est ça aussi. Bon, après, la liste est très longue.

C. R. : Voilà.

Fanny : Mais on a beaucoup de partenaires !

C. R. : Et en fait, la liste, elle est très longue et il faut...

Fanny : Oui.

C. R. : Chercher.

A. A. : Et tu ne peux pas croiser, effectivement, les collections mentionnées, les collections intégrées. Donc ça, je pense que c'est une distinction qui n'est pas du tout claire pour les utilisateurs, en général. À mon avis, il faudrait ne pas faire de distinction. Juste...

Arnaud : C'est vrai que dans le protocole, la première phase, c'est plutôt une phase de tour de table où chacun exprime un petit peu, voilà, ses usages, ses cas d'usage, ses exemples. Et on n'est pas supposé répondre les uns aux autres puisqu'on est plutôt, voilà, dans un...

Eléonore : C'est en temps de témoignages, quand Arnaud vous présentera le moteur de recherche, là on aura un temps de questions.

Arnaud : Là on était plutôt dans un recueil de pratiques et de besoins par rapport à vos propres usages.

A. C. : Moi je veux bien jouer le mauvais élève, le bibliothécaire réfractaire du XXe siècle. Quand un livre manque en place, ou parce qu'il a été numérisé, c'est une bonne raison pour ne pas communiquer le papier. J'utilise effectivement souvent Gallica en service public, un cas d'usage c'est le lecteur qui est très mécontent car il n'a pas pu avoir la version papier de l'ouvrage qu'il demandait car il est hors d'usage ou parce qu'il est ... ou parce qu'il a été numérisé car c'est une bonne raison de ne pas communiquer le papier. Et alors, à ce moment-là, mon rôle, c'est d'essayer de désamorcer la colère du lecteur en lui trouvant la version numérisée. Et alors, j'avoue que, quelquefois, je ne vais pas directement dans le moteur de recherche de Gallica, parce que j'ai beaucoup de bruit dans mes réponses. Donc, je fais ce qu'il ne faut pas faire, je fais ma recherche dans le catalogue général, parce que je le maîtrise mieux, en fait, je le fais souvent. Et puis, le lien vers la version numérisée dans Gallica me permet de retrouver l'ouvrage en question.

XX : Je suis d'accord avec ça, je fais pareil.

S. D. : Alors, je rebondis à ce que dit A. C.. Alors, moi, je suis encore moins bonne élève, parce qu'en service public, moi, j'interroge Google, et je mets le titre, et je mets Gallica, et il me le sort tout de suite.

V. L. : Alors, moi, je vais rebondir sur S. D.. Alors, on va peut-être dire nos noms, parce que je pense que notre collègue, elle a du mal à suivre. Donc, S. D. vient de parler, V. L. va parler. Ce sera plus simple pour une enquête, ça va sans doute être compliqué. Ben, moi, c'est pareil. En fin de compte, je ne passe pas par Gallica. Je vais le dire ouvertement, je fais beaucoup de recherche, je ne passe pas du tout par là. Je passe soit par Presse Locale Ancienne, qui me permet, surtout que je travaille plus sur la presse, donc je passe par Presse locale ancienne d'abord, en priorité pour moi. Quand c'est de la presse, surtout que c'est de la presse avant la Première Guerre mondiale, donc c'est parfait. Soit je passe par Google aussi, et je marque Notice

BnF. Et tout d'un coup, j'ai ce qu'il me faut, parce que oui, je peux... Alors, quand je cherche par nom d'auteur, je peux avoir énormément de bruit. Surtout par les auteurs, alors que ça peut sembler paradoxal, mais on dirait qu'il y a presque des rapprochements, je ne sais pas, qui fait que tout d'un coup, j'ai des listes, mais qui sont interminables, alors que j'ai demandé... Un auteur en particulier. Alors, quand je connais sa date magnifique, je l'ai, mais si on ne sait pas sa date, il suffit d'un homonyme, et là... fiouuuu ! Donc, voilà. Mais sinon, en recherche avancée, moi, ce que j'aime bien, c'est la proximité. Ça, j'ai adoré. Mais je pourrais dire que c'est un jeu de pomme tôt (?). Donc, je ne sais pas forcément, je le mets à combien de distance ? Trois ? Quatre ? Et en fin de compte, on ne sait pas. Soit on le fait à notre instinct, par rapport à ce qu'on souhaite trouver, mais c'est que subjectif. Voilà pour ma part.

Fanny : Juste pour agir à vos trois interventions, c'est dans le cas où vous savez ce que vous cherchez.

V. L. et S. D. : Tout à fait.

Fanny : Est-ce qu'il vous arrive, du coup, d'utiliser le moteur de recherche quand vous ne savez pas ce que vous cherchez ?

V. L. : Jamais !

S. D. : Ben oui !

V. L. : Moi, non. [rires]

Fanny : C'est bien ces ateliers !

V. L. : Moi, spécialement. Ou sinon, si je passe par le moteur de recherche, j'ai aucune idée, c'est pour l'icono. D'accord. Alors là, ouais aller j'y vais, je me fais plaisir. Mais bien sûr que la façon de rechercher, ça dépend de ce qu'on te demande.

S. D. : Mais bien sûr que la façon de chercher dépend de ce qu'on te demande. Si on demande à faire de la valorisation sur un partenaire, bien sûr qu'on va passer par la page partenaire, et qu'on va aller sur la partie de droite, où il y a toutes les notices sur lesquelles on pointe. Donc, bien sûr, on va faire comme ça.

Fanny : Oui.

S. D. : Si c'est un lecteur en salle qui te demande un ouvrage en particulier, j'aurais tendance que... Je suis désolée, mais j'utilise Google. Mais bien sûr, ça dépend de ta raison. Et on... Bien sûr.

Fanny : Non, mais du coup, est-ce que des fois, tu ne sais pas du tout ce que tu cherches ? Enfin, ce n'est pas lié à un partenaire, ce n'est pas lié à une demande précise du lecteur. Et dans ce cas-là, quel est ton parcours ? Est-ce que tu utilises la recherche simple ? Est-ce que tu vas directement en recherche avancée en te disant... Je vais choisir... Je vais choisir le sujet de la notice. Par exemple, imaginons que tu cherches quelque chose sur le train.

S. D. : Moi, je fais par type de sujet.

V. L. : Moi, spécialement, par type de document.

A. A. : Sur les collections iconographiques, moi, je passe clairement par la banque d'images.

V. L. : Moi aussi.

Fanny : On a eu tout ça.

V. L. : Mais en même temps, j'aime... Là, par contre, Gallica, j'aime bien... La banque d'images, j'ai l'impression qu'on me...qu'on me restreint ou qu'on m'oriente presque trop. Et je préfère passer par le côté... type de document. Et là, des fois, je peux trouver des collections insoupçonnées. Et j'aime bien. J'aime bien le côté exploration, en tout cas, parce que tu parlais du côté d'exploration au début. J'aime bien explorer de cette manière-là. Mais je lui dis, c'est plus dans l'aspect visuel.

S. D. : Après, il y a tout ce qu'on ne connaît pas. Je ne sais plus si ça y est toujours, cette recherche sur le code projet. Voilà, c'est pas mal. Les documents basques, les documents... Quand on fait de la valorisation sur les langues de France, après, il y a des...

A. A. : C'est la recherche thématique, qu'est-ce que tu évoques ?

S. D. : Alors, moi, ça s'appelait autrement. Alors, je ne sais pas si vous l'avez... En tout cas, je sais qu'il y a ça, mais... C'est la recherche thématique. Et après, il y a tout ce qui est vos sélections. Alors, je ne suis plus dans le moteur de recherche, mais tout ce qui est sélections, si un lecteur me dit, oui, je fais une recherche sur le sport, eh bien, je vais aller dans vos... dans vos collections aussi. Mais ce n'est plus votre moteur de recherche. Mais ça n'empêche qu'il y a des résultats là-dedans.

XX : Dans le même ordre des choses, moi, j'utilise beaucoup, véritablement, je ne sais pas comment ça s'appelle, tous les dossiers littéraires sur les écrivains.

Fanny : Les pages sélection ?

XX : C'est la sélection, voilà. Moi, je vais directement dans ces dossiers-là et je m'en sers dans mon dossier sélection. C'est ça.

V. L. : Alors, moi, j'y vais, c'est quand je veux découvrir quelque chose. Ou quand je n'ai pas assez d'assises en termes de connaissances, j'aime bien démarrer par là.

N. K. : Je peux rebondir un peu sur l'icône parce que du coup, moi, je fais beaucoup de recherches icône.

Fanny : Oui, je m'en doute. Et tu trouves beaucoup de choses.

N. K. : Et je trouve beaucoup de choses. Mais pas toujours comme je voudrais. [rire] Mais du coup, je vais pas mal passer par type de document et en général, commencer par l'image. Et ensuite, je vais aussi souvent beaucoup chercher dans la presse. Et là, je tombe... Enfin, un de mes... Une de mes difficultés, c'est que je tombe sur énormément de presse. Et en fait, je veux que la presse illustrée. Mais du coup, il y a un petit manque de filtre. Peut-être. Enfin, après, c'est un peu spécifique aussi. Voilà. Je vais parfois passer par la recherche avancée en réduisant les dates. Si je veux un truc sur une période précise. Voilà. Et après, il y a les questions de bruit plus pour ce qui va concerner, on va dire, des recherches qui ne sont pas icône ou pas que icône. Il y a des recherches de sources, voilà. Où je vais... Enfin... Je vais aussi avoir plutôt tendance à passer par la presse, même si au final, je pourrais certainement trouver d'autres sources aussi intéressantes et tout. Mais j'ai un cadre de référence, on va dire, sur la presse qui peut me manquer sur les autres sources. Genre, si c'est des monographies et tout, je ne vais pas forcément savoir bien évaluer où elles se placent, qui parlent, etc. Tandis que... Donc, ça va me paraître plus simple de passer par la presse. Je vais passer par la presse pour avoir déjà une espèce de présélection, en fait. Et après, je passe aussi par les sélections.

Fanny : Et pour l'iconographie, tu passes par la bande d'images ?

N. K. : Pas du tout. Je l'ai noté.

Fanny : Tu passes par Google ?

N. K. : Non. Je ne passe pas par Google pour aller sur Gallica.

V. L. : Ce qui est marrant, c'est quand tu passes par la bande d'images, tu prends le... Moi, je me suis amusée à faire une expo. Tu tapes ciel, un truc très bête. Euh... Tu refais la même chose dans Gallica. C'est assez rigolo en termes de résultats. Côté icono, tu ne retrouves pas les mêmes choses.

Fanny : En ayant sélectionné image ?

V. L. : Oui.

Fanny : C'est normal.

V. L. : Oui, c'est normal. Mais attends, je ne dis pas le contraire.

A. A. : Tu veux dire que c'est complémentaire ?

V. L. : Les deux sont complémentaires. Les deux sont complémentaires. Parce que tu ne trouveras pas les mêmes choses. Voilà. Alors que le corpus in fine est pourtant dans la même base.

A. A. : Oui, il y a une sous-base dans l'image.

V. L. : Voilà. Mais il y a deux bases.

Fanny : Après, il y a deux bases.

V. L. : Mais c'est ça qui est intéressant. En tout cas, si on reste sur la notion d'icono, quelqu'un qui va axer son travail là-dessus ne sait pas qu'il faut jongler sur les deux bases. Donc, est-ce qu'il faut faire un lien, à un moment donné, sur la banque d'images de Gallica, vraiment la banque d'images, pour donner un plus à un chercheur qui ne fait que de l'icono ? C'est une vraie question pour moi.

S. D. : Vous pouvez rappeler comment est alimentée la banque d'images ?

Fanny : La banque d'images, c'est une partie, c'est effectivement les images de Gallica, mais ils numérisent aussi à la demande. Et donc, ces numérisations n'intègrent pas forcément Gallica parce que souvent, il y a de la numérisation à la demande qui est sous-droit, donc qui reste dans la banque d'images. Donc, on va avoir des choses complémentaires, effectivement. Après, à voir si on peut les utiliser, suivant les usages. On ne peut pas forcément les utiliser, mais du coup, ça peut être complémentaire.

A. A. : Il y a aussi des numérisations, effectivement, qui ont été réalisées par le passé, donc une numérisation partielle de tout le monde. Aujourd'hui, ça n'existerait plus, mais la banque

d'images a été constituée au fil des années, et c'est différent pour tout le monde. Donc, il y a une différence entre la banque d'images et Gallica, en termes de production.

Arnaud : Une partie, en fait, de la banque d'images a été versée dans Gallica, je ne sais pas si vous l'avez connue ?

Fanny : Et du coup, N. K., je voulais réagir à ce que tu disais par rapport au bruit. Est-ce que tu peux... Qu'est-ce que tu entends par bruit ? Quels sont les documents ? Qu'est-ce qui ressort, en fait ?

N. K. : Je ne sais pas si c'est du bruit en tant que tel, ou si c'est juste que j'ai trop de réponses, et j'essaie de les réduire.

Fanny : Des réponses qui sont pertinentes ou pas pertinentes ?

N. K. : Je n'arrive pas à dire si elles sont pertinentes ou pas, donc du coup, j'essaie de les réduire.

V. L. : C'est intéressant.

N. K. : Après, du coup, je vais aussi faire ça en sélectionnant les dates sur les filtres qui sont sur le côté.

Arnaud : Pour réduire le bruit, tu utilises les filtres. Comment tu réduis le bruit ?

N. K. : Sur les filtres ?

Arnaud : De manière globale.

N. K. : Ce que j'utilise surtout, c'est les dates, mais du coup, c'est un peu limité, parce que c'est par siècle, donc c'est un peu... C'est les auteurs, quand j'arrive à voir un peu quels auteurs j'ai besoin. Et des fois, c'est un peu difficile. On ne peut pas sélectionner plusieurs auteurs, par exemple. Des fois, on voit bien qu'il y a le même auteur trois fois, avec soit les dates, soit pas les dates, soit machin, soit pas les bonnes majuscules. Et donc, il y a des trucs comme ça. Après, les autres, je ne les utilise pas trop.

V. L. : C'est limite à se demander s'il y a des problèmes de casse.

N. K. : Oui.

V. L. : Ouvertement. Pour les auteurs, c'est à se demander. Parce que je suis d'accord avec toi, quand tu vas retrouver un auteur trois fois, de manière aléatoire, tu te dis, attendez, je ne comprends plus.

Fanny : Parce qu'il n'est pas écrit de la même manière.

V. L. : Voilà. Mais où il y a une subtilité qui... Je ne sais pas, soit on a mis son deuxième prénom, et que l'autre est la première fois, on n'a pas mis son deuxième prénom. Et c'est normal, puisque là, on revient sur une notion toute bête, qui est la notion de notice.

Arnaud : Il n'y a pas d'autorité dans Gallica.

V. L. : Et il n'y a pas d'autorité. Ce n'est pas lié à des notices d'autorité. Eh bien, voilà. Ce n'est pas lié à des notices d'autorité.

S. D. : Ce n'est pas le cas dans le catalogue général.

V. L. : Exactement.

S. D. : Tu le retrouveras plus facilement dans le catalogue général. Il y aura moins de bruit.

V. L. : Merci. Eh bien, voilà.

Fanny : Pour vous rassurer, c'est ce qui ressort. Là, les ateliers qu'on a faits, on a très souvent, la notion, je passe par le catalogue général, parce que je sais ce que je veux, je veux tel auteur, je veux tel titre. Et ce sera plus simple dans le catalogue général de basculer en dessous. Et ça, ne soyez pas surpris, tout le monde a dit la même chose. Google. Tout le monde en a parlé. La banque d'images. Tout le monde en a parlé. Vous êtes dans les clous. C'est bon. [Rires]

A. A. : Nous avons oublié quand même Data.

Fanny : Il n'y a quasiment personne qui en a parlé.

S. D. : Est-ce que quelqu'un utilise Data pour qu'on ait le retour d'expérience ?

A. C. : En passant par Google, oui.

V. L. : Mais oui, en passant par Google ! Et là, oui. Mais en passant par Google.

N. K. : Moi, je ne comprends pas comment ça marche.

V. L. : Par contre, j'aime bien les infos qu'il y a dedans.

N. K. : Oui.

S. D. : C'est la structuration que j'aime bien. Mais autrement, en effet, c'est...

A. G. : Data.bnf en fait, ils ont pris les notices soit à peu près à la moitié des notices. Et donc, le problème, c'est avec l'autre moitié. Et puis, il y a NOEMI entre temps.

Fanny : Du coup, si on revient à la notion de bruit... Et la présentation des documents en liste de résultats ? Est-ce que vous avez des remarques ou des témoignages à avoir sur cette liste de résultats ?

A. C. : Peut-être sur le calcul de la pertinence... Qui est parfois un peu surprenant.

Fanny : Oui.

A. C. : On tape, par exemple, Jean-Jacques Rousseau. Et on a les images qui arrivent en premier. On le voit bien parce que Gallica est souvent utilisée pour les images. Donc, on veut privilégier les images. Donc, on veut privilégier les images dans les résultats, mais il se trouve que là, en fait, les images ne sont pas du tout représentatives des résultats qui arrivent. Il y a beaucoup plus de monographies que d'images. Et je dirais que là, effectivement, sur le calcul de la pertinence, ça serait peut-être bien de réagir un peu en fonction du contenu et mettre en vedette un contenu qui est représentatif de l'ensemble.

Fanny : Ça Arnaud, tu pourras l'expliquer tout à l'heure.

Arnaud : On profitera aussi de la présence de Jean-Christophe pour apporter des réflexions.

D. D. S. : Et sur le bruit, visuellement, il n'y a rien qui simplifie la visualisation. Il y a énormément d'informations. Il n'y a pas de hiérarchie de l'information. Les images ne sont souvent pas au bon format visuel. Elles n'ont pas la bonne taille. Ce n'est pas homogène. D'un point de vue du coup de design, moi ça me donne envie de m'enfoncer une aiguille dans les yeux du coup je passe plutôt par Pinterest ou Instagram parce que vous avez déjà fait une sélection, tout est au bon format et visuellement c'est agréable en termes de consultation.

Fanny : Et alors du coup pour parler de cette interface et de la manière dont c'est présenté si on parle de design etc ça c'est sur la liste de résultats ? Est-ce que vous avez quelque chose à apporter, à témoigner sur le visualiseur, une fois que vous avez cliqué sur ce document une fois que vous pensez que ce document est pertinent pour votre recherche quel est votre expérience sur ce document-là ?

A. C. : Déjà pour arriver au visualiseur, il y a souvent trop de clics on a besoin de faire 4 ou 5 clics pour arriver vraiment au document c'est quelquefois difficile à comprendre et je dirais que quand on a un document. Je ne sais pas comment mais ils arrivent alors peut-être qu'ils font des choix plus radicaux, qu'ils ont moins le souci de plaire à tout le monde mais on a en fait beaucoup plus facilement et donc plus rapidement l'accès au texte et il y a aussi alors ça c'est pas vraiment la visualisation mais un problème de temps de réponse alors là aussi la

comparaison est peut-être biaisée parce que quand j'ai fait mes tests aux Etats-Unis et en Australie c'était la nuit ils dormaient alors que Gallica tout le monde est réveillé, donc les serveurs ont du mal à fournir mais la comparaison n'est pas forcément à notre avantage en termes de temps de réponse et ça dans l'inconscient de l'internaute qui a la patience d'un enfant de 6 ans, 8 ans, etc ça peut le rendre mal disposé envers nous si il attend une seconde et demie au lieu de deux dixième de seconde

Fanny : Et s'il arrive sur un document qui n'est pas visible...

A. C. : Et si en plus il doit faire 4 clics au lieu d'un, ça met un peu de pression

Fanny : Et sur ce visualiseur de documents, est-ce que les informations qui sont fournies sur le document correspondent à la recherche que vous avez faite ?

N. K. : Il me semble que ça a bougé depuis quelques temps et que du coup sur le truc sur le côté on a l'histoire de la taille et du format du document qui est affiché et qu'à les plus loin il n'est plus affiché il faut cliquer dessus pour avoir les infos et ce truc là je le trouve un peu ... un peu gênant au final parce que du coup moi c'est les infos biographiques que j'aimerais trouver plus directement quand j'arrive sur le document. Après bon il y a des trucs genre... enfin... on va dire des problèmes techniques entre guillemets mais genre on arrive des fois sur la page de couverture, des fois sur la page de garde, des fois sur la page du document, et en fait des fois du coup c'est ... enfin genre quand on arrive sur la page de titre du document, on se rend pas forcément compte ... il faut avoir en tête qu'il y a des pages avant du coup il y a peut-être de l'illustration avant, des fois il y a des trucs qui vont nous intéresser, qui sont avant mais du coup c'est un peu caché par le fait d'arriver sur cette page de titre et puis après il y a des problèmes des fois le zoom quand on essaye de sélectionner une zone, le zoom il saute et alors du coup, il faut repartir sur la sélection... enfin bref mais ça c'est...

V. L. : Et après il y a toutes les strates de différentes numérisations aussi, dans Gallica, et ça c'est pas de sa faute, si je reprends de nouveau pour la presse on aura des systèmes où tout d'un coup on a un titre de presse et pour avoir l'intégralité des différentes années on a une année mais en fait il faut cliquer dedans et après ça se redéroule et là tout d'un coup on a tout l'ensemble des autres fascicules derrière et là c'est pas du tout et on aura à la fois une autre présentation pour un autre titre parce qu'il a été numérisé plus récemment donc après ça c'est de la num mais si on en est sur la notion de moteur de recherche en terme de présentation ce serait bien qu'il y ait une sorte d'homogénéité de présentation peu importe les dates de numérisation, peu importe en fin de compte notre histoire vraiment de documents numériques non natifs on va dire c'est

d'avoir une homogénéité... une homogénéisation de présentation... parce que c'est clair que pour... il y a des choses qui ne sont pas instinctives. Et puis il y a des astuces en même temps dans les recherches et là on arrive ou pas à les comprendre des fois il faut cliquer longtemps et tout d'un coup le sous-document ou dans l'éphéméride aussi, il y a des subtilités dans les éphémérides, ça aussi ce n'est pas simple, très complexe l'éphéméride ! Je ne vais pas rentrer dedans mais on n'a pas fini...

A. A. : Pour faire des recherches, vraiment efficaces, il faut connaître les formats de description de données, la structure de données, il faut savoir exactement quelles données tu vises et où elles se trouvent dans la structuration dans InterMarc et DublinCore.

XX : Un outil grand public en somme [rires]

E. B. : Si je peux ajouter un élément en termes d'usage de la recherche avancée de Gallica ça fait un moment qu'il y a une formation Gallica sur la recherche, sur la recherche avancée c'est pas anodin ! C'est pas quelque chose d'intuitif mais en effet il y a beaucoup de subtilités c'est à dire savoir à la fois déjouer les pièges c'est pas un dysfonctionnement c'est l'outil qui est comme ça qui s'enrichit continuellement mais capter et connaître les astuces ou savoir lire entre les lignes et ça pour un usager lambda c'est pas forcément évident, vous me direz que c'est le cas pour n'importe quelle interface de recherche il faut se l'approprier etc mais il me semble quand même qu'il faut quand même pour être efficace éventuellement avoir une formation imaginons qu'elle soit disponible pour le grand public en tout cas pour le public extérieur c'est ça que je veux dire par grand public pas un interne. Certainement mais parce que si on n'a pas en effet une connaissance en archives, en bibliothéconomie ni une connaissance globale des formats c'est un élément à prendre en compte parce que ça fait un moment que c'est comme ça et ça rend les choses pas forcément évidentes. Je mettrais mon témoignage, moi je ne fais pas de recherche régulièrement, je ne fais pas de service public, donc je ne suis pas amené à me poser la question, et c'est le moment où je l'ai fait, en effet fatalement par Google on arrive à remonter et c'est plus facile quelque part de démarrer par Google et ensuite réduire essayer d'affiner sa recherche plutôt que de passer d'emblée par l'interface de recherche avancée même si elle est complète même si elle a beaucoup de facilités pour autant on peut être un peu désappointé par le résultat.

Arnaud : En tout cas on peut se réjouir que Gallica soit bien référencée par Google.

E. B. : C'est déjà pas mal.

Fanny : Juste pour vous donner un élément, des derniers chiffres que l'on a, la recherche avancée, elle est utilisée par 3% des utilisateurs...

V. L. : C'est tout ?!

Fanny : L'arrivée via le catalogue c'est à peu près 3% des utilisateurs et à peu près 20% des utilisateurs qui utilisent la recherche simple ça nous fait 25%, et le reste c'est par Google. A mon avis, il y a aussi dans ce reste, il y a aussi des usages via le blog, parce que quand on arrive sur une image, on arrive sur le document dans Gallica, et là on ne compte que la navigation dans la base de données on va dire, il peut y avoir des sélections comme on parlait tout à l'heure donc arriver comme ça mais il y a quand même une minorité qui utilise la recherche avancée et la recherche simple.

A. C. : Ça c'est courant, c'est pour ça que l'on fait data.bnf.

D. D. S. : Après le principe de base de l'UX Design c'est qu'un utilisateur va passer moins de 4 secondes à décider si ce site va être un obstacle à sa mission ou à sa tâche ou à ce qu'il essaye de faire, je pense que du coup le choix est vite fait pour les utilisateurs actuellement...

V. L. : Ça fait peur "un obstacle" !? Waw.

D. D. S. : Le principe c'est qu'il y a tellement d'espace en ligne que les utilisateurs passeront toujours plus de temps sur d'autres sites ou d'autres solutions qui lui permettront d'atteindre son objectif plus facilement...

V. L. : D'accord, donc il y a une navigation très fluide, la plus fluide possible.

D. D. S. : C'est ça, et on ne lit pas tout ce qu'il y a sur un site, on scanne

Fanny : Il faut qu'il comprenne où il arrive aussi, et est-ce que l'utilisateur il comprend où il arrive aussi ?

V. L. : oui oui oui.

E. B. : C'est le principe de la file du supermarché, en général, c'est vrai qu'on saute assez vite quand on est dans une grande agglomération, c'est vrai que les gens ont un minimum de patience, ça se mêle avec l'UX design, ça va vite.

V. L. : Après moi il y a une chose aussi qui m'intéresse dans Gallica et je trouve ça intéressant. C'est le surlignage en jaune des noms d'auteurs parce que là les rebonds ça fonctionne bien. En

tout cas je trouve qu'on peut faire des recherches par rebond intéressantes. Mais je trouve presque que c'est sous-exploité c'est mon opinion, je pense que ça pourrait être, mieux exploité

A. A. : Oui, une recherche par nom d'auteur, ça peut te faire des rebonds.

V. L. : Exactement ! Parce que pour moi l'idéal, que ça soit pour l'iconographie ou pour tout, quand on est dans la presse c'est cool le rebond.

Fanny : Mais du coup, les rebonds qui existent aujourd'hui, les suggestions de poursuivre sa recherche est-ce que vous les utilisez ? Vous voyez où c'est ? Un même sujet, un même auteur... Est-ce que c'est des choses que vous utilisez ? Est-ce que c'est pertinent ?

D. D. S. : Oui, c'est hyper pertinent, c'est une des choses qui est bien ! Par contre il n'y a que deux suggestions, ensemble dans le terme des images, on arrive souvent à des boucles, où on tourne sur la même chose, mais c'est hyper pratique. Et c'est comme ça que fonctionne la majorité des sites pour faciliter la navigation, on clique d'un bout à l'autre

V. L. : Si je ne me trompe pas, il y a Gallica vous conseille aussi en haut ?

Fanny : Oui, on en parle peut-être après. Juste, pour que tout le monde voit, quand on est sur un document, on a en dessous de la notice une proposition de voir plus loin : les auteurs et il y a trois champs qui peuvent arriver ici c'est l'auteur, le sujet et la description.

A. A. : Le champ de description c'est le code projet.

Fanny : Ces trois champs, c'est pour ça que je demandais est-ce que c'est pertinent ? Est-ce que ça vous sert ? Est-ce que vous cliquez dessus ? Est-ce que vous ne cliquez jamais dessus ?

V. L. : Je clique. Pour moi hein, je clique. Parce qu'on fait compte, des fois ça nous emmène sur des choses sur lesquelles on n'aurait pas pensé... Mais en même temps, je ne sais pas, je voudrais bien avoir l'opinion de ceux qui l'utilisent... comment dire ? Je ne sais pas si le regard de l'utilisateur est vraiment attiré par ça ? Voilà, je me suis aperçue parce qu'à force de chercher mais instinctivement (si je suis honnête avec moi-même) mon regard n'a pas été attiré, par cette section-là...

D. D. S. : Ce n'est pas mis en valeur...

V. L. : Non ce n'est pas mis en valeur !

D. D. S. : Après c'est une image dans des textes, donc là ça peut attirer l'attention.

V. L. : Oui tout à fait mais ce n'est pas mis en valeur, je pense que c'est à force de chercher qu'on réalise à force de regarder, bon on regarde, on y va mais ça ne saute pas aux yeux.

Fanny : Mais a-t-on besoin de cette image dans du texte pour attirer l'attention ? Si ça ne tient qu'à ça, le fait que toi tu as plutôt été attirée par le fait qu'il y ait une image dans du texte donc ce rebond t'a attiré et du coup, si on va avoir le témoignage de V. L., si ça ne tient qu'à ça pour attirer le regard est-ce que c'est pertinent les images ?

D. D. S. : Après je crois qu'il y a une étude en 2023 sur les utilisateurs et une des premières choses qu'ils demandaient c'était des suggestions donc il me semble que c'est pertinent ?

V. L. : C'est marrant, tu avais dit tout à l'heure que tu jouais le côté rabat-joie, moi je vais jouer la vieille donc je suis un usager de plus de 50 ans. Ouvertement, trop d'images tue l'image. Il y a beaucoup d'images déjà sur la page d'accueil de Gallica, on est attirés par l'image donc à un moment donné avoir du texte c'est quand même bien. Moi je dirais plutôt, c'est comment mettre en valeur ce texte au niveau de sa casse, au niveau de la couleur, il y a plein de possibilités de mettre un texte en valeur mais trop d'images, fiouf. Je vais rebondir un peu, sur ce que disait A. C., on cherche du fond aussi, on ne cherche pas que de l'icône, on cherche du contenu et les contenus sont multiples. Et quand on recherche du texte, vraiment textuel, c'est intéressant aussi d'être attiré par une suggestion textuelle. Je pense que toute l'image peut aussi être au final avoir un côté plus dépressif qu'autre chose, je pense qu'il faut doser.

D. D. S. : Je parlais d'un point de vue iconographique, parce que c'est le côté qui m'a intéressé.

Arnaud : Il y a aussi d'autres rebonds qui sont relancés notamment avec la recherche par sujet, est-ce que c'est quelque chose que vous utilisez ? Parce que ça c'est un lien...

Fanny : Ce dont parlait Arnaud c'est dans la liste de résultats, il y a la possibilité de relancer la recherche et ça c'est ce qu'elle disait utiliser N. K., c'est-à-dire que tu choisis par auteur, tu relances ta recherche en ayant choisi un auteur.

V. L. : Oui, je vois très bien.

Fanny : En fait ça relance une autre recherche.

V. L. : Mais moi je ne le fais pas par sujet.

Arnaud : C'est proposé par sujet, quand tu regardes dans la notice, tu as les différents sujets qui sont indexés, ça te permet de relancer la recherche par sujet et donc éventuellement de rebondir.

V. L. : Tu le fais ça toi N. K. ?

N. K. : Je le fais par auteur et par type de document, des fois par bibliothèque ou par site de consultation.

Fanny : Et du coup d'une manière générale, est-ce que vous utilisez ces facettes ? Cette partie-là qu'on appelle, cette colonne de gauche, on appelle ça les facettes de recherche, est-ce qu'il y a des personnes qui l'utilisent ? Ou qui n'utilisent jamais qui repassent par la recherche ? Est-ce qu'il y a des personnes qui l'utilisent systématiquement ? Toi N. K. tu dis plutôt l'utiliser systématiquement, quasiment systématiquement parce que ta première recherche du coup n'est pas pertinente ?

N. K. : Parce qu'en général, c'est des recherches vraiment larges, plutôt que de passer par la recherche avancée.

V. L. : Tu réduis grâce à cette section ?

N. K. : Plutôt que de faire une recherche avancée en sélectionnant déjà tout d'abord, je vais faire ma recherche simple, et après je vais sélectionner dans les facettes.

Fanny : D'accord.

V. L. : La recherche avancée moi des fois, j'ai peur que ça me réduise, que ça ne m'offre pas assez de choses, mais c'est peut-être bête, c'est peut-être bête ? Je ne sais pas, mais des fois j'ai l'impression que je ne vais pas avoir ... peut-être trop restreint qui est en lien avec ma pensée mais justement qui ne me permet pas d'ouvrir un peu plus pas de beaucoup d'ouvrir mon angle au moins de 20° ou 30° et qui me permettrait justement de faire un rebond plus intéressant donc des fois c'est vrai que je commence toujours par la recherche simple aussi avant de passer par la recherche avancée.

Fanny : Et du coup est-ce que ça serait des recherches avancées ou plusieurs champs qui seraient combinés ou ?

V. L. : Des recherches avec des champs combinés mais tout d'un coup-là la notion de sujet apparaît, pour moi, mais je ne passe pas par le sujet en facette jamais.

Fanny : Est-ce qu'il y a d'autres personnes qui ont envie de témoigner ? N. N. ?

N. N. : Là je ne passe pas souvent par la recherche avancée parce que je travaille plutôt sur des collections personnelles et je trouve que c'est très, pour moi, la date de la mise en ligne...

V. L. : Ah ! Excellent !

N. N. : Par contre, quand je suis en service public, je passe toujours par le catalogue général.

A. C. : La recherche avancée, c'est là que quelques fonctions qui permettraient de réduire les problèmes de bruit par exemple en recherche par adjacent et qui est un peu difficile à utiliser et à expliquer aux lecteurs mais c'est efficace !

A. A. : Et le contournement par le catalogue général fonctionne bien sur les collections imprimées mais ne fonctionne pas du tout sur les documents qui sont intégrées dans BAM, avec le moteur de recherche, c'est un peu...

A. C. : il faut utiliser le CCFr manuscrits pour chercher.

A. A. : Ah, moi le CCFr Manuscrits, je l'utilise plutôt dans le backoffice de BAM dans le XML.

V. L. : Non mais là, c'est l'expert expert. Nous sommes d'humbles chercheurs.

A. C. : Alors en parlant de rebonds, on est parfois mis en difficulté et les lecteurs aussi encore plus que nous, sont volontaires par exemple de Gallica vers RetroNews quand on cherche dans la presse dans l'éphéméride, très souvent on se retrouve sans explication vers RetroNews c'est pas si grave que ça quand on est en Rez-de-Jardin puisque tout le monde a le droit à RetroNews mais les lecteurs qui sont à l'extérieur se trouvent parfois dans des situations très compliquées parce qu'il y a des titres de presse ou des anciens numéros pour lesquels ils ont quelques bonnes raisons de penser que c'est libre de droit ça l'est peut-être pas mais dans certains cas ça l'est quand même et ils sont basculés vers RetroNews ils sont bloqués ou ils sortent leurs cartes bancaires pour prendre la rerue à RetroNews

V. L. : Ou là tu fais du détournement.

A. C. : Et ils sont très en colère.

V. L. : Et là tu peux faire du détournement.

A. C. : En salle de lecture, on a parfois du mal à expliquer pourquoi certains titres... L'illustration bien sûr c'est parfois un peu particulier mais il y a d'autres titres de presse anciens y compris des titres du 19ème pour lesquels on est basculés vers RetroNews...

D. D. S. : Ça a peut-être un lien avec Google parce que dans les Google Trends la première chose qui est liée c'est RetroNews et Généanet...

Fanny : Non, c'est juste que le document a été intégré avec un lien avec RetroNews.

D. D. S. : Ok mais du coup ça justifie la Google Trends. D'accord.

Fanny : On est même des fois tombés sur des documents partenaires numérisés pendant quelques années puis la numérisation s'est arrêtée parce que nos partenaires n'avaient plus ce fonds et la suite ça a été numérisé par RetroNews et donc on a une notice RetroNews pour un document qui est aussi accessible sur Gallica. C'est un vaste sujet slash problème.

V. L. : Mais quand tu passes par Presse Locale Ancienne, au moins tu as les deux liens.

Fanny : Oui oui complètement et après normalement, alors je ne sais pas en tout cas dans mes exemples en tête les liens RetroNews sont toujours moins pertinents que les documents visibles dans Gallica je reprends l'exemple du même titre mais sur des années différentes le premier lien enfin le premier résultat qui apparaissait c'était ceux qui étaient numérisés par le partenaire et disponibles sur Gallica en deuxième occurrence il y avait le lien vers RetroNews mais oui c'est un sujet après ce sera de moins en moins un sujet puisque les documents arriveront un jour sur Gallica (bientôt !). M. P. dit oui !

A. C. : Dans quelques dizaines d'années oui.

Fanny : Non, non ! Ça a commencé en 2016, donc 2026 c'est dans deux ans, l'embargo c'est 10 ans. Mais oui oui je comprends que c'est confus.

V. L. : D'autant plus que la confusion s'explique quand RetroNews est présenté comme je cite "le site média de la BnF" donc là ça génère beaucoup de confusion mais là les contournements sont nombreux sont très nombreux et qu'ils fonctionnent.

N. K. : Tu fais quoi ?

V. L. : Je passe par presse locale ancienne.

N. K. : Pour trouver les trucs dans Gallica et ne pas avoir les trucs de RetroNews ?

V. L. : Voilà, ou tu peux aussi même via Google avoir les deux liens l'un derrière l'autre dans les recherches.

XX : Retronews d'abord ?

V. L. : Pas forcément ce n'est pas une question de RetroNews d'abord parce que des fois c'est même pas les mêmes sections qui ont été numérisées donc c'est pas du tout mais bon.

Fanny : Est-ce que vous qui faites du service public, est-ce qu'il y a des questions sur cette pertinence, l'ordre de pertinence des résultats, l'affichage des résultats, sur quelle est la

différence entre les livres qu'on voit en extrait ici on peut les lire, entre est-ce que je dois relancer une recherche dans Gallica Intramuros, ... sur toute cette explicitation de quels sont les documents, quelle est la base que j'interroge quand je suis dans Gallica, que je fais de recherche ? Est-ce que vous avez des questions sur ces choses-là ? Est-ce que les gens comprennent dans quoi ils recherchent ?

A. C. : Les gens cherchent des documents récents.

XX : Tu es sévère !

XX : Non, ce n'est pas vrai.

XX : Ça dépend de la salle !

XX : Ça dépend en Rez-de-Jardins.

A. C. : Si on regarde de quoi les gens se plaignent sur Sindbad, ils se plaignent de ne pas trouver parce qu'ils ont besoin des guillemets. Mais sinon, ils se plaignent que c'est vachement long à télécharger ceci cela. Des fois, ils signalent une erreur mais sinon, rien, a priori ça veut dire que les gens savent ce qu'ils cherchent

Fanny : Ou sont-ils partis avant de trouver ? Sont-ils partis avant d'avoir trouvé ?

A. C. : S'ils ne se plaignent pas c'est qu'a priori ils savent ce qu'ils cherchent et ils ont trouvé.

V. L. : C'est comme si on avait un biais cognitif, depuis le début, on part du principe qu'on sait ce qu'on cherche...

Fanny : C'est ce que je vous ai dit tout à l'heure, est-ce que des fois vous faites des recherches exploratoires ? Est-ce que l'utilisateur chez lui, sait-il ce qu'il cherche ?

Réactions variées : XX bah non ! Pas souvent ! Je ne crois pas ! ...

I. V. : Oui, des fois on a des questions, les gens ne savent pas ce qu'ils cherchent. Je veux un document, en latin, sur les arts...

V. L. : Et ça pour toi ce n'est pas savoir ce que tu cherches ?

I. V. : Bah non, c'est vraiment très large. La recherche ça peut être très bien thématique.

V. L. : Ce n'est pas ça que tu veux dire ?

Fanny : Si si, ça peut !

I. V. : Pour moi une recherche sur le sport, sur l'art, c'est à nous après de recentrer...

XX : Oui, on essaie de leur poser des questions.

I. V. : Mais pour moi c'est déjà une recherche. Mais des fois les modes de recherche sont très larges.

Fanny : Il y a une différence entre “je cherche les misérables de Victor Hugo” ou je cherche des documents traitant du sport.

V. L. : Mais c'est une recherche.

Fanny : Oui mais tu ne vas pas faire pareil.

V. L. : Oui mais tu ne vas pas naviguer dans Gallica pour voir ce qu'est Gallica et ne pas avoir de recherche ? Et là c'est différent pour moi, dans ta question je pensais que c'était des gens qui allaient dans Gallica juste pour regarder Gallica et là c'est différent pour moi. C'est évident qu'en salle on a toujours des recherches qui peuvent être larges ou recentrées.

Fanny : La frustration de ne pas trouver n'est pas la même quand on ne sait pas ce qu'on cherche et quand on sait précisément ce qu'on cherche.

XX : Moi j'ai aucun lecteur qui ne sait pas ce qu'il cherche mais bon...

D. D. S. : Parce qu'il y a une démarche déjà pour arriver en salle, il y a une démarche pour remplir une notice qu'il y a quelque chose qui marche ou pas ou quoi que ce soit mais si vous avez des gens en ligne qui se baladent qui tombent sur votre site par hasard qui arrivent sur la page principale c'est à partir de là qu'ils vont se dire oh bah je vais cliquer là ça a l'air intéressant du coup ça m'attire je vais cliquer là mais ils ont pas forcément d'idée fixe en tête et selon de nouveau l'étude de 2023 c'est quand même un cas assez fréquent.

V. L. : De toute façon imagine t'es un jeune chercheur, un étudiant en master mais tu vas faire tu vas fonctionner par rebond, tu vas explorer, sinon ton sujet il est déjà écrit et au revoir. Donc de toute façon il y aura pour moi cette recherche exploratoire qui correspond à une recherche par rebond parce qu'en fin de compte c'est des recherches par arborescence on part d'une idée j'ai mon sujet mais dans mon sujet j'ai des sous-sujets et en fin de compte j'ai ma carte j'ai tout d'un coup une map qui apparaît et qui va me faire rebondir au plein d'autres éléments ou plein d'autres choses mais, au départ, je sais pas du tout que j'allais arriver là-dessus je veux pas dire que c'est le hasard mais c'est la démarche de recherche donc d'exploration du creusement on va

dire je mine au niveau intellectuel qui me permet d'arriver à ça mais au départ je sais pas que je vais arriver à ça.

A. A. : Il y a la modalité effectivement rebond et puis la modalité qui a été évoquée par N. K. sa méthode c'est-à-dire faire une recherche large et puis on va appuyer les petites facettes.

V. L. : Mais c'est comme si tu crées, genre t'avais un grand grenier à blé et puis tout d'un coup tu crées dans ton grenier différents silos pardon pour la métaphore agricole.

A. C. : C'est la sérendipité.

V. L. : Oui si tu veux.

A. C. : Mais maintenant les lecteurs prennent Chat GPT ou autre outil et ça leur donne déjà une première liste de référence, ça leur donne encore du boulot mais ...

V. L. : Et je pourrais dire, et pourquoi pas ? Je veux dire à la rigueur ? Et pourquoi pas ? Je veux dire peu importe les points de départ, je veux dire faut bien démarrer par quelque chose.

A. C. : Après il cherche ses références.

Fanny : Pour réagir à ce qu'a dit A. A., sur le fait que la recherche de N. K., elle parle d'une recherche large et elle précise, est-ce qu'aujourd'hui l'inverse serait intéressant pour vous ? Partir, vous faites une recherche on vous propose quelques résultats et on vous propose d'étendre votre recherche ?

V. L. : Pour un esprit fermé c'est intéressant, moi je trouve ça intéressant.

Fanny : Mais c'est intéressant. Juste pour rappel le moteur de recherche tel qu'il existe aujourd'hui a à peu près 10 ans et il y avait 1 million de documents à cette époque-là, aujourd'hui on a 10 millions de documents, donc est-ce qu'on est dans la même logique ? C'est ça que je voulais interroger est-ce que inverser la logique c'est intéressant pour vous ou pas ? Est-ce que vous avez ... vous essayez déjà de faire ça ?

D. D. S. : C'est le principe des suggestions alors oui.

Arnaud : Disons que la logique c'est une recherche élargie qu'on va affiner, filtrer et finalement cibler, est-ce que la logique inverse serait d'avoir une recherche plutôt réduite qu'on pourrait élargir ?

E. B. : Passer du synthétique à l'analytique à l'inverse de l'analytique et synthétique par exemple. Du spécifique au général et non du général au spécifique.

Fanny : Je reprends mon idée “Les Misérables de Victor Hugo”, aujourd'hui si je tape ça dans Gallica, je vais avoir toutes les éditions...

N. K. : Tous les articles de presse qui en parlent...

Fanny : Tous les articles de presse qui en parlent, je vais peut-être avoir des dessins etc.

XX : en tant que novice, moi en tant que personne/ usager ponctuel, à titre perso, ça n'a rien à voir avec ça dans la présentation, déjà je bloque. Voilà. Je fais une recherche, je ne sais pas par où commencer. Où est-ce que je vais ?

D. D. S. : c'est hyper riche mais ça ne va pas du tout quoi.

Fanny : Juste Les Misérables, sans guillemet, sans rien, on a 142 000 résultats. Aujourd'hui, ma question, c'est, est-ce qu'on doit présenter ces 142 000 résultats ou est-ce que la personne qui a tapé Les Misérables, elle s'attend à avoir les 10 résultats les plus pertinents et si elle veut en voir d'autres, on lui dit, ben, voir d'autres, par exemple, la presse, Les Misérables dans la presse, par exemple, Les Misérables en images, etc ?

Réactions diverses.

S. D. : Bah non, tu as les facettes donc je ne vois pas l'intérêt de dire Les Misérables en images, sachant que là, la personne, elle va d'elle-même filtrer avec les facettes et que vous avez fait un Gallica vous conseille, pour, a priori, des documents qui sont les plus emblématiques.

Fanny : Alors, imaginons qu'il n'y a pas de Gallica, vous conseille, parce que...

V. L. : Non, mais non, vous ne pouvez pas faire comme ça.

Fanny : Il n'y est pas toujours, imaginons qu'il n'y est pas. Vas-y, V. L..

V. L. Piveteau (V. L. 2) : En fait, moi, mon expérience avec Gallica, c'est, pour une recherche large, j'attends de Gallica qu'il structure ma pensée. Le problème, c'est que l'arborescence, elle n'est pas du tout visible. Alors, effectivement, on passe les pages, jusqu'à ce que... On trouve un titre qui nous attire. Et là, c'est très intéressant. Et là, on aimerait avoir d'autres propositions, enfin, trouver des choses de cet ordre-là. Et là, on... Ben non, il faut s'atteler de patience, c'est vrai. Il y a quelque chose qui n'est pas du tout... Je trouve pas du tout... Il faut... Il faut faire l'effort soi-même, mais notre effort est récompensable.

V. L. : Moi, par exemple, si on parle vraiment moteur de recherche, je ne suis pas du tout technicienne ni rien. Tu as éteint Gallica vous conseille. Alors, première question, est-ce que,

dans cet exemple, quoi qu'il arrive, quand un usager ouvre, aura-t-il tout de suite Gallica vous conseille ? Juste ça, déjà.

Fanny : Oui. Oui, mais imaginons que ça soit... Je n'ai pas en tête une recherche sur laquelle je suis sûre qu'il n'y aura pas de Gallica vous conseille. Mais là, ce que je veux vous dire, par exemple, ce que j'aimerais interroger, c'est que, regardons les résultats. Vous voyez, la personne a tapé les misérables. Imaginons que c'est une personne qui vient pour la première fois dans Gallica et tape les misérables. Le premier document qu'il a, tu peux remonter... C'est une impression photo métallique des misérables. En deuxième, c'est un film. En troisième, c'est une photo.

V. L. : C'est sur ça que je voulais revenir.

Fanny : Là, il y a le Gallica vous conseille. Effectivement, le texte va arriver en premier. Le roman va arriver en premier. Mais s'il n'y est pas parce qu'on ne fait pas des Gallica vous conseille pour tout. Puisque des fois, on n'en a pas. Voilà. Ma question était, doit-on lui présenter... Ah, on est reparti sur 156 000 résultats. Mais doit-on lui représenter les 142 000 résultats ? Oui/ Non, parce que vous avez des usages différents, donc, vous aurez des avis différents. Ou doit-on lui présenter 10 résultats, les plus pertinents que sont le roman et quelques autres ? Et puis, on lui dit, explorez et allez-y.

D. D. S. : Oui, l'exploration. Enfin, normalement, la recherche, elle doit être pertinente. Et l'exploration, elle se fait à partir du visualisateur. Ce n'est pas dans le même espace. Parce que là, déjà, on a énormément d'infos. Il y a des infos partout. Il y a de la couleur rouge. Enfin, bref. Donc, le principe, c'est que si quelqu'un arrive sur le site avec une idée précise en tête, il veut le résultat le plus pertinent. Après, l'exploration et l'élargissement, elles peuvent se faire dans un autre espace. Quand il est plus calme, face à un seul document, avec pas 10 000 informations qui s'entremêlent.

Fanny : Alors, on n'a pas 4 heures, mais qu'est-ce que c'est la pertinence ?

V. L. : Eh bien, on n'a pas 4 heures, mais qu'est-ce qu'on a comme biais cognitifs ? On est parti du principe que les misérables, juste en partant du mot misérables, ce sera forcément Victor Hugo. Donc, dans le genre bien cognitif, c'est super. Alors que, c'est quoi les misérables ? Ça peut être totalement autre chose. Je cherche les misérables d'un point de vue sociologique, comme le 19e.

N. K. : Sur Gallica ?

D. D. S. : Faudrait demander aux utilisateurs.

Discussion vive et simultanée.

XX : On pourrait avoir le moteur de recherche qui classe automatiquement les résultats par type de document.

Fanny : Voilà, livre, image, manuscrit...

V. L. : Parce que tu me dis... C'est forcément... Pourquoi ça serait forcément Victor Hugo ? Suivant ton niveau, socio-économique, pourquoi ce ne serait pas la comédie musicale ? Tu t'en fous que ce soit Victor Hugo qui l'ait fait.

D. D. S. : Pourquoi, je dirais, sur la BnF et Gallica...

V. L. : Mais je n'en sais rien, tu tombes dessus ? Pourquoi ? Pourquoi pas ? Si je passe par Google...

D. D. S. : Dans tous les cas, je ne peux pas répondre à cette question puisque je ne suis pas l'utilisateur non spécialiste qui va utiliser ça.

V. L. : D'accord, mais les Misérables, pour moi, ce n'est pas forcément, tout de suite, forcément Victor Hugo. Pour moi, c'est un biais cognitif de fait. Une recherche, je tape les Misérables, pourquoi, j'attends forcément Victor Hugo ? Là, je pars du principe que je sais ce que je cherche. Mais si je ne sais pas ce que je cherche et je tape les Misérables ?

D. D. S. : Il faudrait faire une analyse de combien de personnes ont tapé les Misérables en pensant trouver la comédie musicale sur Gallica/ BnF et combien de personnes ne l'ont pas fait.

Fanny : Ou même autre chose.

V. L. : Mais pour moi, c'est un axe.

D. D. S. : Après, ça dépend aussi de comment vous le présentez.

A. C. : Si on met entre guillemets, ça donne quoi les Misérables ?

V. L. : Tiens, bonne question. Et sans article.

XX : Non mais vous pouvez dire que ça ne fonctionne pas les guillemets.

Fanny : En recherche simple c'est censé fonctionner.

XX : Ce qui m'aurait intéressé d'avoir dans le système à côté là...

Fanny : Les facettes...

XX : C'est livres, une presse, revue. Ça aurait été plutôt en termes d'histoire, en termes de sociologie, en termes de littérature, en termes de pays, en termes de à quoi c'est associé.

Fanny : Ça, c'est des types de documents qui correspondent aux départements des collections de la BnF. Là, on n'est vraiment pas très loin. Donc, l'utilisateur, il n'en sait rien de ça, mais bon, il comprend que s'il clique sur l'image, il va plutôt avoir des images. Après, qu'est-ce que c'est, une image ?

V. L. : Oh là là ! Mais on va faire de la philo là !

Fanny : Non, mais qu'est-ce que c'est ?

A. C. : Il n'y a pas de champs-sujets, mais il y a un thème.

S. D. : Mais ta question, c'était... Ta question, c'était... Je peux me tromper, mais c'était... On met un certain nombre de réponses par défaut qui sont pertinentes ou pas, qui, finalement, sont une sélection faite par je ne sais pas quoi, un algorithme, on n'en sait rien. Et après, si on a une page de n'importe quoi, de 10 résultats, on a quelque chose qui nous emmène plus loin, ailleurs, qui nous ouvre nos...

Fanny : Voilà, exactement. Voilà, c'était ça. Au lieu d'affiner les 150 000, alors là, on en a, maintenant, avec les guillemets, on a 26 000 résultats. Au lieu d'affiner dans ces 26 000 résultats, est-ce qu'on propose moins de résultats, mais on propose d'explorer, de continuer la recherche ? C'est une question complètement ouverte.

I. V. : Moi, j'ai une question. Il y a 26 700... Sur 1785. Ça veut dire quoi ?

Fanny : Pages.

I. V. : Ça veut dire que les 26 000 résultats, on les visualise vraiment ? À un moment donné, c'est pas arrêté ?

Fanny : C'est page par page.

I. V. : Oui, d'accord. Je pensais que ce n'était pas moi qu'on ne les voyait pas.

I. V. : Est-ce que c'est une amélioration qu'il y a eu ? Ça a toujours été ? On a toujours tous les résultats ?

Arnaud : Ah oui !

Fanny : C'est justement ça qui rejoint ma question. Est-ce qu'il faut tout montrer ?

Réactions diverses : Ça dépend/ Moi, je dis oui.

A. C. : Alors que dans Google, aller au-delà du 1000eme...

N. K. : Sur la recherche, quand on commence à taper le truc, on a des propositions qui s'affichent, des documents. Mais du coup... Enfin, on pourrait aussi envisager d'avoir des questions. Genre, est-ce que vous voulez dire Les Misérables de Victor Hugo ? Est-ce que vous voulez dire le sujet des misérables ?

V. L. : Tu veux dire des bulles ?

N. K. : Je ne sais pas si la forme, ce serait des bulles ou je n'en sais rien. Ça peut être...

V. L. : Non, mais est-ce que vous cherchez un sujet des misérables ? Il y a la recherche avancée, sujet des misérables ?

N. K. : Oui, mais du coup, le but, c'est que t'as déjà des gens qui sont en train de faire une recherche simple et que du coup, ces gens qui sont sur la recherche avancée, ils font la recherche simple. En fait, de fait, moi, je le travaille pour faire comme s'ils faisaient une recherche avancée.

Fanny : Attention, si on peut faire la recherche avancée avec sujet des Misérables, ça n'interroge que les notices dans lesquelles il y aura sujet des misérables. Tout à fait. Je ne vais pas mettre le doigt là où ça fait mal, mais...

S. D. : Toute la notice "Les misérables".

Arnaud : Par rapport à l'affichage des résultats, et notamment le fait que quand on fait de la recherche simple, on va chercher à la fois dans la notice et dans le plein texte. Donc, on a eu des retours d'utilisateurs, même des collègues en service public, qui disaient que ça créait des confusions pour les utilisateurs parce qu'ils étaient mélangés un peu, finalement, des résultats qui étaient issus du plein texte, d'autres issus des notices, et aussi une multiplicité de types de documents qui ne donnaient pas forcément une bonne lisibilité, finalement, des résultats. Est-ce que ça, c'est quelque chose que vous pensez comme une difficulté ou est-ce que c'est au contraire une richesse d'avoir cette diversité à la fois des types de documents et des sources, en fait, dans lesquelles le moteur de recherche va rechercher ?

A. A. : Je pense qu'il y a vraiment deux modalités de recherche différentes. Et donc, il y a des réponses, deux réponses, plusieurs réponses même différentes en fonction des modalités de recherche parce qu'il y a la modalité qui est plutôt envisagée par les professionnels ou par les chercheurs, et puis plutôt au grand public. Donc, je pense que pour le grand public, peu importe, ils veulent juste avoir le résultat le plus pertinent possible et le plus rapidement possible. Peu importe où on va chercher la donnée, finalement, si c'est l'OCR...

Arnaud : Là, je parlais en termes d'affichage, en fait.

A. A. : Ah oui, en termes d'affichage.

Fanny : Et alors, juste sur ce que tu viens de nous dire, par rapport à ce que tu disais, D. D. S., sur l'enquête, les premiers éléments du baromètre des publics de 2023, une bonne session sur Gallica, pour un peu plus de 30% des répondants, c'est trouver des documents qu'ils n'avaient pas... Même un peu plus de 30%, mais trouver des documents qu'ils ne pensaient pas trouver.

D. D. S. : C'est un parcours recherche ciblée et un parcours découverte ou exploration. Ce que tu as décrit au début. C'est de l'exploration, en fait, et c'est beaucoup plus intuitif et beaucoup moins normé. Et juste pour faire l'équilibre entre deux choses qui ont l'air d'être très opposées, en fait, la solution, je vais prêcher pour ma paroisse c'est une solution de design, en fait. En gros, la personne qui veut faire recherche devra toujours pouvoir le faire depuis l'exploration, et l'exploration doit toujours être accessible, avec un minimum de bruit recherche spécialisée pour les personnes qui sont là pour découvrir. C'est une histoire d'élégance dans le design, en fait.

Arnaud : On sait aussi, d'après les enquêtes d'usage, que les utilisateurs font les deux. C'est-à-dire qu'à la fois, ils cherchent un document précis, et ils veulent explorer par rebond dans le domaine de connaissances.

D. D. S. : C'est ça, un équilibre.

Arnaud : Ce n'est pas si évident si tu veux dissocier, finalement, les usages, le type d'usage. Souvent, c'est mixte, en fait. C'est des usages entrelacés.

V. L. : C'est marrant, parce que l'élégance dans le design, j'entends bien cette phrase, mais l'élégance dans le design, ce n'est pas l'élégance de la pensée. Pour moi, à un moment donné, soit je cherche, soit je ne cherche pas. Alors, si je ne peux faire que du design, ok, mais ça veut dire qu'on va orienter, tu m'as dit, orienter ma pensée vers des choix bien déterminés.

D. D. S. : Non, le principe de base, c'est de laisser la liberté à l'utilisateur d'atteindre son but et de le faciliter. Quand on vient dans une bibliothèque, on peut choisir ce qu'on veut faire, on peut avoir accès à plein de choses, mais on a les conditions physiques. Pour citer François XX, on a les conditions physiques de sa pensée, on a un espace dans lequel on se sent bien, et on peut aller au fond de quelque chose. Ou on peut juste rester sur un banc ou aller boire un café.

V. L. : Je connais très bien cette phrase.

D. D. S. : Et c'est la même chose que doit pouvoir offrir un espace numérique.

V. L. : On est d'accord.

D. D. S. : Mais ça n'a rien à voir avec l'atelier.

V. L. : Mais en même temps, si on reprend la finalité de la bibliothèque, qui est censée représenter quand même la Bibliothèque nationale de France et ses fonds, à un moment donné, il faut bien qu'ils soient visibles.

D. D. S. : Oui, bien sûr, mais pas overwhelming, pas chaotique. Il faut que ce soit visible de façon agréable et compréhensible.

V. L. : Ça, je suis parfaitement d'accord. C'est I. V. qui l'a dit, elle voit le truc, elle est perdue. Donc oui, c'est chaotique, on est bien d'accord. Peut-être parce qu'on a envie de présenter trop de choses. Ou tout de suite, voilà, je te donne le menu. Alors, comme les menus dans certains bars, il y en a des pages et des pages, tiens, tu te débrouilles. On en est peut-être là, on est d'accord. Mais en même temps, si on décide d'épurer, et c'est d'accord, on est dans une plus grande sobriété, mais alors dans ces conditions, comment on oriente le regard pour que les choses ne se perdent pas.

D. D. S. : Mais c'est une vraie question.

V. L. : C'est une vraie question.

Fanny : Je vais parler graphisme avant de partir. Peut-être que ce que je n'ai pas dit en introduction, mais vous le savez déjà, vu qu'on n'en a pas la dernière réunion de département, c'est qu'il y a aussi tout un travail qui est en train d'être fait sur le graphisme, l'ergonomie et la migration technique de la page d'accueil. Donc ça changera aussi en termes de typo, en termes de couleur, en termes d'esthétique. Alors, pour l'instant, on n'a pas dit... On sait comment va changer cette page-là, etc. Il y aura sûrement un tout petit peu de changement sur la typo dans

un premier temps, peut-être les couleurs. Peut-être qu'on pourra vous présenter, je ne sais pas, je n'en ai pas parlé à M. P., mais peut-être qu'à une prochaine réunion de département, je crois.

M. P. : C'est ce que je voulais proposer justement.

Fanny : On pourra présenter où on en est. Ça va être développé à partir du mois de juin, ça a été validé par la hiérarchie, etc. Et vous verrez que déjà, ça va dans ce sens-là. Simplification. Simplification, pas de l'esprit. Simplification du chemin.

M.-F. R. : J'avais une idée comme ça, c'est les applications générales, quand on les télécharge sur son smartphone, il y a des fois un tuto, c'est-à-dire que vous passez à autre chose, il présente comment l'utiliser. Et je me disais, est-ce qu'il est possible, par exemple, de faire une sorte de cache qui soit retirable ou pas retirable, un sort de mini-programme qui permettrait de guider la personne vers les choix ou des bulles. Je ne sais pas comment...

Fanny : Alors, il y a les tutos déjà qui existent. Les tutos d'Arnaud.

I. V. : Oui, il y a les tutos. Disons que vous avez les tutos à côté, mais un truc juste un peu simple. Un truc juste un peu caché.

Fanny : Oui du genre "Comment l'utiliser ? C'est votre première recherche ?"

V. L. : Comme pour les impôts. [Rires].

XX : C'est d'actualité, en fait.

Arnaud : les tutos ne présentent pas vraiment l'interface, c'est plus des scénarios.

Fanny : oui, c'est vrai, tu as raison, Arnaud. C'est plus des scénarios d'usage ici, mais peut-être qu'on pourrait avoir... Voilà, c'est votre première visite sur Gallica. On montre quelques éléments.

V. L. : C'est bien.

I. V. : Sur les côtés, si je ne vois pas le thème, ça ne me percute pas tout de suite. Mais si on met une pointe, si vous cherchez par thème une grosse flèche, là, je vais aller... Je ne vais pas perdre mon temps à regarder partout.

Fanny : Là, c'est la clarté des messages et de ce qu'on propose pour que ça soit...

A. G. : Un assistant virtuel en gros ?

I. V. : Oui, c'est ça.

V. L. : Juste un point, comme on est tous câblés différemment, en termes d'ergonomie, n'oubliez pas les dyslexies au niveau des casses, des typos. De tous les dys. Parce que ça joue énormément au niveau visuel. (Je sais de quoi je parle). Donc, ça joue... Beaucoup au niveau visuel. C'est important, puisque c'est pris de plus en plus en compte. Même beaucoup de sites prennent ça en compte, en termes de choix de couleurs, d'ergonomie et de casses.

A. C. : Vous avez prévu une version FAP justement ? C'est prévu ça ?

V. L. : Ce n'est même pas ça.

Fanny : Je ne sais pas, on va en parler à... On va en parler avec les développeurs du DSI. Là, dans les maquettes, on n'a pas réellement... Après, on est censé être en accessibilité à un niveau exemplaire. On doit être à un niveau 99%. Donc, il y a une première idée de graphisme, mais il va falloir que ça, ça s'adapte à l'accessibilité dans toutes ses formes. Il y aura peut-être des choix qui ont été faits côté graphisme, sur lesquels on va devoir revenir en termes de taille, en termes de... On a une exigence quand même assez forte.

S. D. : La refonte de la page d'accueil, ce n'est pas seulement graphique ? On est d'accord, c'est ergonomique, graphique ? On va naviguer différemment aussi ?

Fanny : On va naviguer différemment. Il y aura derrière un Drupal différent.

S. D. : Ah oui, c'est vrai, tu m'as dit ça.

Fanny : On fait tout en même temps. C'est-à-dire qu'en migrant techniquement, on est obligés de repenser l'ergonomie. Est-ce que l'ergonomie fonctionne toujours avec la version 10 de Drupal ? En faisant cette question sur l'ergonomie, on est obligé de repenser le graphisme. Là, la page d'accueil, elle a 6-7 ans.

S. D. : Donc, ça veut dire que vous avez revu des entrées qui sont sur le côté. Vous avez, enfin, tu sais, outils, voilà, outils et outils, partenaires, etc. Vous avez revu ça aussi ?

Fanny : Plus rien n'existe.

S. D. : D'accord. Est-ce que la place des partenaires, elle va être plus visible ? Puisque c'est la bibliothèque numérique de la BnF et des partenaires.

Fanny : Pas forcément sur la page d'accueil.

S. D. : D'accord.

Fanny : Mais ce qu'on a fait, c'est une création d'un menu. Il n'y a pas de menu actuellement et il y aura un menu. Et dans ce menu, il y a toute une partie sur les partenaires. Et justement, renvoie vers une page partenaire, vers la coopération nationale, la coopération numérique. Qu'est-ce que c'est ? Etc.

XX : Et vers le CCFr peut-être ?

Fanny : Mais c'est tous les textes qui existent déjà dans BnF.

S. D. : Vous faites des renvois sur bnf.fr. Donc, Gallica-Marque-Blanche aussi, vous faites un renvoi sur BNF.fr. C'est intéressant de savoir ça.

Fanny : Mais ça, on pourra rentrer dans le détail après. Mais tout ça, c'est ce qui existe aujourd'hui.

I. V. : Et est-ce que vous allez sur des pages plus profondes ? Et j'argumente.

Fanny : [sur le footer] Tout ça, ça va être balayé. Ça va être balayé pour intégrer un menu. Parce que ça n'a pas lieu d'être dans un footer. Et il ne restera dans le footer que les mentions légales, les conditions d'utilisation, les contacts.

S. D. : Sur la page partenaire, vous avez pensé à une refonte ?

Fanny : Non.

S. D. : Donc, il n'y aura pas de moteur de recherche intégré à la page partenaire ?

Fanny : Non. La refonte, elle concerne le blog et la page d'accueil.

S. D. : Parce que la page partenaire, en fait, c'est quand même hyper difficile de trouver l'institution, l'établissement.

Fanny : Pour l'instant, la refonte ne concerne que la page d'accueil et le blog. Ça ne concerne même pas les pages sélections. Mais on va quand même faire une migration technique. Et cette page-là, comme la recherche avancée, il y aura sûrement le graphisme qui va être revu parce qu'on a une autre typo qui va être mis en place. Sûrement des couleurs. Et pour l'instant, on n'y a pas travaillé.

S. D. : Est-ce que c'est compliqué d'intégrer un moteur de recherche dedans ?

A. A. : Alors, il y a une astuce très simple que tu as fait chez tous les partenaires. C'est le contrôle F.

S. D. : Sérieux ?

V. L. : Alors ça, c'est bien, mais tout le monde ne sait pas.

S. D. : Ah non, mais sérieux ? Alors, mieux de tard que jamais. N. K., tu savais ça ?

N. K. : Bah oui. Mais le contrôle F, ce n'est pas Gallica, c'est sur tout l'Internet.

I. V. : En fait, nous, ce qu'on fait, c'est qu'on biaise, on met dans tout Gallica.

N. K. : Mais sur Word aussi !

Fanny : Pour les partenaires, on va revenir dessus

V. L. : oui mais ça je le fais dans mon tableau Excel !

Fanny : Techniquement, il y a plusieurs pages qui existent aujourd'hui. Il y a un Drupal page d'accueil. Il y a un Drupal page blog. Il y a un Drupal page sélection. Il y a un Drupal page partenaire. Déjà, la refonte va avoir pour mission de rassembler le Drupal page d'accueil et blog pour qu'on ait une automatisation des biais de blog sur la page d'accueil, de manière automatique, les nouveaux articles de blog apparaissent sur la page d'accueil. Déjà, c'est la première chose. Ensuite, ça, c'est un autre Drupal sur lequel, pour l'instant, on n'a pas travaillé comme le Drupal des pages sélection. On n'y a pas travaillé. Assez logiquement, ce que je vous disais, c'est que les grands principes du graphisme qui vont être déployés sur la page d'accueil vont se retrouver sur cette page. Donc, sûrement de la couleur qui change, un type qui change, etc. Mais on ne les a pas fait maquetter. Donc, ça sera un intermédiaire pour l'instant. Si on a besoin dans un certain temps de refondre ces pages-là parce qu'il faut passer à un Drupal qui est un peu plus long, un Drupal 10, là aussi, il faudra le faire dans un second temps. On va dire que, dans un premier temps, on va tout mettre la même couleur, uniformiser tous les headers, tous les footers. Et dans un second temps, on verra si techniquement ou ergonomiquement, il y a besoin de faire des retouches. Ça, c'est pour page sélection et pour page partenaire.

V. L. : Quand tu dis des Drupal différents, c'est des sous ?

Fanny : Des back-office différents.

V. L. : Oui, des back-offices.

Fanny : Donc, il n'y a pas... Aujourd'hui, c'est pour ça qu'on a une page d'accueil. C'est une page d'accueil sur laquelle rien n'est automatisée. C'est une page de vitrine dans laquelle on met des informations même. Mais si demain, on publie un nouvel article sur le blog, la page

d'accueil ne comprend pas qu'il y a un nouvel article sur le blog parce qu'il n'y a pas de lien entre les deux. Je vous laisse là. Arnaud, tu leur expliques la suite. On aura l'occasion de reparler plus tard dans la réunion de département.

Temps 2 : Explications moteur

[Non enregistré]

Temps 3 : Améliorations

[Non enregistré]

Entretiens individuels

Les entretiens individuels ont fait l'objet d'échanges plus intimes que ateliers, certaines données personnelles ou confidentielles évoquées ne peuvent être divulguées c'est pourquoi nous n'avons pas intégré les transcriptions de ces échanges aux annexes. Elles peuvent toutefois être communiquées sur demande et après accord de la personne interrogée.

Protocole d'entretien

Introduction : Présenter le projet et le cadre de l'entretien (10 min)

Introduction : présentation du contexte et du cadre de travail

Exposer le cadre des travaux :

- Réflexion sur l'évolution du moteur > feuille de route
- Double approche système/usages
- Usagers : travaux commencés avec les collègues

Plusieurs formats, dont des entretiens semi-directifs avec des collègues qui ont un lien particulier et fort avec Gallica.

Temps 1 : Témoignage (qui je suis et d'où je parle) - 20 min

Témoignage : quel est votre lien/ rapport avec Gallica ?

Historique/ rétrospective/ évolution :

Vous avez participé à la fondation de Gallica, quelles sont pour vous les évolutions marquantes/majeures ? (et comment avez-vous accompagné ces évolutions ?)

- Le concept de bibliothèque numérique

Temps 2 : Les usages / ses usages / et fonctionnalités (30 min)

Fonctionnement de Gallica et fonctionnalités / son expérience

- Quels sont ses usages de Gallica ? En tant qu'utilisateur de Gallica les avantages et les inconvénients ? Quelles sont les limites, les difficultés ?

Quelles sont les fonctionnalités que vous préférez/ que vous utilisez le plus ?

Les quatre/six améliorations/évolutions proposées par les collègues ??

- Recherche simple / avancée / unifiée / progressive
- Ajouter du grain dans la recherche et l'affichage : œuvre/édition/exemplaire

- Documenter la plateforme, les collections, les outils et visualiser la collection (la recherche dans la collection)
- Paramétrage : paramétrer différemment par défaut pour tout le monde (liste de résultats) et de permettre à chacun de paramétrer l’affichage
- Fonctions expertes : troncature, recherche dans l’epub, recherche géographique, recherche par côte, code projet, etc.
- Bruit : expansions, etc.
- Pertinence et ordonnancement des résultats

Faire évoluer Gallica : Quelles sont les demandes d’amélioration de Gallica ? Quelles sont les demandes ? Quels points notables identifiez-vous ? Au cours des ateliers, plusieurs thèmes ont émergé (voir prises de notes d’Irène) qu’en pensez-vous ? Êtes-vous d’accord ou non avec ces points soulevés par les participants ? Avez-vous des choses à ajouter ?

Temps 3 : Gallica vision 2030

On assiste aujourd’hui à une transformation/mutation/ des modalités d’accès à l’information avec les IA génératives, les LLM, etc.

Comment voyez-vous l’avenir de Gallica confronté à ses bouleversements ?

Comment voyez-vous l’évolution de la bibliothèque numérique aujourd’hui et demain ?

Quelle vision de la bibliothèque numérique demain ?

- En termes de collections / usages / technologies / etc.

Conclusion :

- Reprendre 3 points saillants soulevés lors de l’entretien
- Calendrier et suite de l’enquête
- Remercier

Transcription atelier GMB : 27 juin 2024

Isabelle Mangou et Chloé Girard, responsables des GMB au sein du CNG, nous ont accompagné dans l'encadrement de cet atelier du fait de leur expertise sur la question des marque blanches.

Temps 1 : Témoignages

Eléonore : Alors, qui se lance ?

P. H. : Je peux commencer comme ça s'est fait ?

Eléonore : Très bien.

P. H. : Donc, P. H., INRAE, je gère le projet, gère la plateforme, donc Agate. Donc, le moteur de recherche, les points positifs, c'est la recherche plein de texte. Ça, ça marche bien, je suis assez contente. Les points négatifs, c'est comme je suis documentaliste, j'ai tendance à aimer les recherches booléennes, et alors là, ça ne va pas du tout. Mais c'est parce que c'est une formation professionnelle. Donc, moi, quand je dois faire des SRU, donc des requêtes pour faire remonter des corpus, je fais un bout de requête. Je copie la réponse, je copie l'URL, je fais un autre bout de requête, et puis après, j'essaye de mixer tout ça sans me tromper. C'est ce qui peut prendre un certain temps. Et quand on a des fonds un peu particuliers, on essaye de les assembler, de se retrouver sur un ensemble cohérent qui, au départ, n'était pas. Ça peut être très long. Donc, ça, c'est... C'est ma limite principale. Et on a présenté à Agate, via le moteur de recherche, en faisant chercher les collègues qui y étaient, donc documentalistes, bibliothécaires, enfin voilà, des bouquins dont je savais qu'ils existaient et comment on y arrivait. Et il y a des fois, ils n'ont pas su faire. Et c'est leur métier. Donc, l'interface, c'est pas forcément la... Je ne la trouve pas forcément hyper... Comment dire ? Ouverte au grand... Enfin, facile à décrypter. Je ne suis toujours pas sûre de savoir si je dois cocher regroupé par titre ou regroupé au volume au fascicule. Pourquoi est-ce que "recherche exacte" est tout en bas ? Enfin, voilà. Il y a des trucs qui me gênent, mais parce que c'est mon métier. J'ai l'habitude des bases bibliographiques hyper pointues où je peux faire des requêtes qui font 50 trucs. Et là, je me sens hyper limitée, en fait. Voilà. Autrement, il y a des choses qui marchent bien, effectivement. Recherche de termes à proximité, ça, ça marche bien. Pouvoir rechercher pour les collègues. Enfin, s'être occupé chez les collègues, des trucs qui marchent bien. Donc, c'est la collection d'une autre marque blanche. Je veux tel type de document. Ça, par contre, ça marche très bien. À condition

de ne pas vouloir une recherche précise. En exploratoire, je pense que ça marche très bien. En précis, ça devient compliqué. Voilà.

Eléonore : Et donc, vous utilisez surtout la recherche avancée. Est-ce que vous utilisez un petit peu la recherche simple ou pas du tout ?

P. H. : Jamais. C'est une déformation professionnelle, là.

Eléonore : Non mais c'est vrai que la recherche avancée, c'est assez spécifique.

P. H. : C'est un réflexe, en fait. Mais voilà. Et quand vous avez une collection, c'est ça, en fait. C'est la recherche avancée.

Eléonore : Et est-ce que vous avez un retour de vos usagers ?

P. H. : Alors, les usagers, ils me contactent quand ils ont trouvé quelque chose et que ça les intéresse. Donc, je ne sais pas à quel point il y en a qui ont cherché et pas trouvé. Finalement, moi, ceux qui m'ont contacté, c'est parce qu'ils avaient trouvé ce qui les intéressait. Voilà.

Eléonore : OK. Fanny, Arnaud, est-ce que vous avez des questions ?

Arnaud : Moi, j'ai noté, ce qu'on vous propose, c'est de faire déjà, de laisser la place aux témoignages, et de répondre à vos questions après. Donc, effectivement, j'expliquerai, notamment en présentant le fonctionnement du moteur, pourquoi on ne peut pas mettre la troncature et quelles sont les limites actuellement pour, effectivement, créer des requêtes comme Madame le souhaite.

Fanny : Et juste une petite précision. Juste une petite précision. Vos retours peuvent aussi concerner la visualisation du document, les outils qui sont proposés autour. Voilà. Une fois qu'on a cliqué, qu'on croit être le bon document, comment est-il présenté ? Ça, ça nous... Alors, ça dépasse un tout petit peu le moteur de recherche, puisque c'est plus de la recherche à proprement parler. Mais ça nous intéresse aussi pour faire évoluer... Pour évoluer ce visualiseur des documents. Est-ce qu'il y a trop d'icônes ? Est-ce que toutes les icônes vous sont utiles ? Est-ce que les informations... Est-ce qu'il y a des informations qui sont manquantes ? Est-ce qu'il y a des...

P. H. : Alors, c'est sur la navigation. Quand vous avez un fond image et que vous passez d'une image à l'autre, vous avez affiché votre image, vous l'avez regardée. Pour revenir au résultat, pour afficher la suivante, vous ne pouvez pas, en fait, faire défiler dans la liste les unes après

les autres. C'est un peu agaçant. Mais ça se fait. Mais voilà, en termes d'ergonomie, j'ai mon collègue iconographe que ça exaspère, par exemple.

Fanny : Donc, on est d'accord de la liste de résultats.

P. H. : C'est ça. Donc, quand vous avez sélectionné une, enfin, voilà, vous en avez affiché une, et vous dites, je veux la suivante, il faut revenir à la liste de résultats pour afficher la suivante. Ça, c'est dommage. Voilà. Mais c'est vraiment le truc qui agace ma collègue iconographe. Mais autrement, pour le reste, le fait qu'on ait... Les petits pop-ups qui vous montrent déjà une prévisualisation, moi, ça me convient. Enfin, voilà. Le fait qu'on arrive sur la page de titre et pas sur la couverture, je sais très bien. Enfin, voilà. Ça, en fait, je n'ai pas de limite là-dessus. Ça me semble bien. Ça me semble correspondre, finalement, à ce que je peux pratiquer et à ce que les personnes que je connais se pratiquent.

Eléonore : OK, très bien.

K. H. : Est-ce que... Est-ce que je peux juste rebondir là-dessus ? Parce que c'était un peu mon point le plus important à évoquer. C'est en fait, moi, le moteur de recherche, parce que je ne fais certainement pas les mêmes recherches que P. H.. Quand moi, je fais une recherche, souvent, je sais ce que je vais chercher. Je connais très, très bien notre fond. Donc, le mot-clé est très précis. Donc, je trouve tout de suite ce que je veux. Mais ce n'est pas la même démarche. Mais pour moi, c'est plutôt aussi la visualisation. C'est une visualisation qui est très compliquée. Alors, je veux dire, sur un journal, je zoome. Pour revenir à mon journal, il faut que je revienne en arrière. Donc, je ne peux pas dézoomer pour revenir à... C'est juste agaçant. On perd beaucoup de temps. Et pareil pour la visualisation de la liste des résultats. Je trouve aussi, on trouve une image... Bon, la visualisation en glissant sur l'image, c'est déjà pas mal. Mais moi, je trouve aussi qu'une fenêtre pop-up, qui est peut-être feuilletable, ça, ça serait un truc super.

Eléonore : Ok.

K. H. : Ouais. Mais ça, c'est plus une question d'ergonomie. Ça concerne moins le moteur de recherche en lui-même.

Fanny : Oui, oui. Mais on peut l'intégrer là-dedans.

Eléonore : C'est ça. Ça nous intéresse aussi parce que, justement, c'est comprendre un petit peu les résultats qui sont rapportés par le moteur. Et la visualisation des résultats est très importante. Donc, voilà. Ça nous intéresse aussi. Et... Par rapport à vos usages sur le moteur, est-ce que vous pourriez les développer un petit peu ?

K. H. : Alors, moi, je suis plutôt contente. Je trouve que les fonctionnalités sont claires. C'est assez intuitif, quand même. On peut, voilà, sélectionner le type de document. Je trouve aussi super bien la recherche par proximité. Ça, je fais la promo à chaque fois quand je fais une démonstration. Parce que je trouve que c'est quand même quelque chose qui n'existe pas partout. Et, par rapport au moteur de recherche, il y a une chose là, je ne sais pas non plus si ça rentre là-dedans. Nous, on a eu, avec Chloé, beaucoup d'échanges sur les légendes. Voilà. Il y a cette fonctionnalité de légende qui est géniale, mais qui est appliquée à Gallica qu'aux images. Ça veut dire, voilà, un album que je catalogue étant le type de document image, je peux intégrer une légende. Ça veut dire, dans le moteur de recherche, le titre de la légende, elle rentre dans le moteur de recherche. Par contre, si j'indexe une monographie, ce n'est pas le cas. On peut faire des commentaires, mais ces commentaires-là ne montent pas dans le moteur de recherche. Et ça, c'est vraiment dommage. Pour un exemple très précis, on a intégré un dictionnaire biographique des Basses-Pyrénées, où je fais donc des légendes qui ne sont que des commentaires. Et je fais des légendes, et là-dedans, elles sont bourrées de noms de notre région. Et ces noms-là ne remontent pas dans le moteur de recherche. Donc, il faut déjà tomber sur ce dictionnaire pour pouvoir découvrir tout le reste. Et ça, c'est vraiment dommage.

Eléonore : OK. Et donc, vous utilisez plutôt la recherche avancée ou est-ce que vous utilisez aussi la recherche simple ?

K. H. : Alors, j'utilise beaucoup la recherche simple. Et je trouve ça vraiment bien. Comme je vous ai dit tout à l'heure, souvent, je sais déjà ce que je cherche. L'église Saint-Martin à Pau. Alors, je sais que je vais avoir mon résultat. Mais beaucoup aussi, la recherche avancée. Et que je trouve très bien. Franchement, je trouve bien.

Arnaud : Est-ce que vous utilisez la recherche géographique ? Est-ce que vos documents sont géolocalisés ?

K. H. : Ah, alors, là, pas trop. Parce qu'elle est un peu cachée. Je tombe. L'autre jour, j'étais sur le site de nos collègues d'Agate, je crois, qui ont fait, je crois que c'était Agate, qui ont fait donc cette géolocalisation. Et je me suis dit, ça, c'est notre prochain projet parce que c'est génial. Mais c'est vrai, je n'ai pas le réflexe d'utiliser, peut-être parce que je n'en ai pas besoin, mais un outil super bien qui est un peu caché. Il y a juste le globe, là.

L. H. : Est-ce que je peux intervenir ?

Eléonore : Oui.

L. H. : Je suis à L. H. du Havre et je rebondis sur la géolocalisation. En fait, on a eu aussi un gros projet en Normandie sur cette géolocalisation des images, mais j'en ai déjà parlé à Isabelle Mangou. Ça ne remonte absolument pas dans la recherche du moteur parce que les données qui sont géolocalisées avec l'outil Arpenteur ne sont pas liées aux données du catalogue. Donc, on va réfléchir ensemble à comment envoyer ces données dans le dur du catalogue, en quelque sorte. Mais il faut savoir que quand on géolocalise avec la petite mappemonde et par arpenteur, en fait, eh bien, on a la surprise de se rendre compte que dans le moteur de recherche, ce n'est pas pris en compte. Donc, voilà. Il faut le savoir.

Arnaud : En effet, on est dans une base de... C'est-à-dire qu'on a une base de données qui est externe, en fait. C'est-à-dire que c'est à partir des données, des métadonnées du catalogue où on génère des coordonnées géographiques pour géolocaliser les documents. Mais ensuite, ce n'est pas réinjecté dans le catalogue général. Donc, effectivement, ce n'est pas indexé par le moteur de recherche. Mais on a bien noté le besoin.

L. H. : En fait, ce qui peut être intéressant, c'est que quand on a fait une géolocalisation vraiment avec des chercheurs, et on est vraiment sûrs de nous, on a vraiment des données qui sont disponibles, prévoir une petite procédure pour consolider cette connaissance, parce que c'est un travail assez considérable qui est réalisé, je pense que les autres collègues qui travaillent là-dessus sont d'accord, et c'est dommage qu'ils soient déjà pas accessibles, et aussi qu'ils risquent d'être perdus un jour, parce qu'ils ne sont pas aussi solidement enregistrés que les données du catalogue.

Isabelle : Alors, ce n'est pas parce que ce n'est pas dans le catalogue que c'est pas enregistré, de toute façon, mais effectivement, comme le disait Arnaud, c'est un petit module qui est effectivement rattaché à Gallica, et ces données ne sont pas ajoutées au catalogue. Après, pour ajouter des données au catalogue, c'est toujours la même chose, il nous faut des gens pour le faire, et on n'a pas suffisamment de force pour le faire. Donc, voilà. C'est aussi bête que ça. Mais tout ne doit pas forcément aller au catalogue, en fait. On peut quand même profiter, moi, je trouve, de ce nouvel outil, les cartes et les documents géolocalisés, même si ce n'est pas ajouté au catalogue. Mais je comprends bien la problématique.

L. H. : C'est que je m'interroge sur la pérennité. Bien sûr, c'est enregistré, mais est-ce que ça a la même pérennité ? Je m'interroge là-dessus. Et effectivement, j'ai bien conscience que déjà, l'équipe Gallica effectue des miracles avec ses moyens. Et d'ailleurs, j'en suis vraiment reconnaissante. Voilà, c'était juste, on en avait parlé, Isabelle, c'est un des pays, enfin, un moyen

de faire ça de manière automatique qui ne pèse pas, pour vos forces limitées. Voilà, ça serait super. Évidemment, ça ne veut pas de vous demander de le faire à la main. Ça devient ingérable pour vous, j'en ai bien conscience.

Fanny : Juste une question. Est-ce que les données géolocalisées, il n'y en a aucune qui remonte dans le moteur de recherche ? Ou elles sont... Ou les résultats ne sont pas pertinents quand on les interroge ?

L. H. : Ça ne remonte pas du tout.

Isabelle : Ce n'est pas dans la notice. Ce n'est pas dans la notice, en fait.

Arnaud : On pourra demander au DSI s'il est possible d'indexer, en fait, un référentiel qui serait constitué justement de cette base externe à Gallica. Donc, ça, ça fait partie des perspectives qui sont... qui peuvent être envisagées peut-être plus simplement que de reverser dans le catalogue général les données qui ont été produites par la géolocalisation. Ça, c'est une piste qu'on pourra explorer.

L. C. : Mais le problème n'est-il pas simplement que la géolocalisation est sur le mode collaboratif et dès l'instant où c'est pas contrôlé par un professionnel, on ne veut pas alimenter le catalogue avec des données non contrôlées ?

Isabelle : Non. Non, je ne crois pas. Non, c'est pas... C'est pas l'argument principal, en fait, puisque de toute façon, on a quand même joué sur la... Enfin, la survalidation, c'est-à-dire que lorsqu'un internaute décide quelque chose, on vérifie et on... C'est toujours au moins deux avis et puis en plus, on a un petit module pour valider les choix qui ont été faits. Donc, non, non, c'est... C'est surtout que le catalogue, c'est effectivement très lourd, que nous, en plus, la recherche géographique au niveau des cartes et plans, elle ne joue pas tout à fait sur les mêmes... sur les mêmes principes, qu'en plus, on va bientôt changer de catalogue. Enfin, je pense que là, on n'est pas forcément dans une perspective très, très faste pour ça, mais c'est sûr qu'il faut le garder en tête effectivement pour des recherches, pour de futures recherches et pour améliorer... Toujours pour améliorer les choses, c'est sûr.

P. H. : Que nous, ça remonte parce qu'on a mis les... en indexation, on a mis les lieux géographiques.

Isabelle : Oui. Oui, oui, oui.

P. H. : Pour les photos, ça a été un choix, on indexe le nom de la commune.

Eléonore : Donc, on reviendra sur les améliorations en temps 3. L. H., est-ce que vous pourriez nous parler un petit peu de vos usages, du moteur, est-ce que vous utilisez plutôt de la recherche simple, avancée, est-ce que ça fonctionne, etc ?

L. H. : Alors, de notre côté, il me semble que je n'ai jamais rencontré de difficultés, en fait, soit, je passe... En fait, j'utilise un peu tout, j'utilise à la fois la recherche simple, quand j'ai un mot vraiment très spécifique, je sais que ça va monter. Si c'est un peu plus complexe, je vais utiliser la recherche avancée et aussi nous, à la bibliothèque du Havre, donc sur Nutrisco, on a travaillé avec des pré-requêtes qui nous permettent aussi de retrouver des corpus facilement et de faire en sorte que nos usagers, qui ne savent pas faire des requêtes, trouvent facilement les images de la bombardée ou vraiment des questions qu'ils se posent souvent sans être un expert, mais de manière très intuitive. Alors, en fait, généralement, les difficultés qu'on rencontre, ce n'est pas, peut-être pas proprement parlé, les fonctionnalités du catalogue, mais parfois la pertinence des résultats, c'est la même difficulté qui existe dans Gallica. Par exemple, pour la presse, notre titre le plus important et le plus richement alimenté, en fait, le titre, ça va être pour l'instant le Petit Havre. Il arrive en page 3 ou 4, je crois, quand on arrive sur la page de la presse. On peut le confondre avec une autre notice où il n'y a pas beaucoup de journaux, donc on peut avoir l'impression qu'on n'a pas grand-chose sur ce titre alors qu'il est presque complet. Donc, effectivement, on peut trouver des stratégies de contournement. À défaut d'avoir la chance comme Gallica d'avoir un Nutrisco vous conseille, je pourrais rajouter une case dans, voilà, une case thématique où je mettrais le Petit Havre, mais effectivement, moi, je ne serais pas contre quelques pages du Nutrisco vous conseille. Par ailleurs, je travaille avec des personnes, enfin, voilà, des chercheurs locaux qui sont parfois âgés et... C'est la surprise parfois de les entendre me dire, ah, c'est bien, vous avez numérisé la presse, mais il n'y a que la page de titre. Donc, la visibilité pour tourner les pages, eh bien, ça ne semble pas être évident pour tout le monde. Enfin, je trouvais ça intéressant de rapporter un peu des usages non-experts, c'est-à-dire ni les miens, ni ceux de bibliothèques, de documentalistes, ni... de chercheurs expérimentés, mais voilà, parfois, ça ne semble pas toujours très visible pour des chercheurs, quand même pas complètement débutants, mais pas non plus experts, voilà.

Eléonore : Oui, c'est bien sûr des retours qui nous intéressent aussi, les vôtres, ceux de vos collègues et ceux de vos usagers, si vous les connaissez et si vous avez des retours qui concernent les usagers. On sait que ce n'est effectivement pas facile du coup, quand on n'est pas dans un milieu bibliophile de base, de se retrouver dans la recherche de Gallica, donc on l'entend parfaitement. Est-ce que vous avez des questions ?

L. C. : Je peux prendre la suite, si vous voulez, Eléonore.

Eléonore : Allez-y.

L. C. : Donc, L. C. pour les archives diplomatiques au ministère des Affaires étrangères. Donc, effectivement, sur les cas d'usage, alors, moi, en tant que bibliothécaire, mais ex-documentaliste et archiviste, j'ai été confronté effectivement à quelques difficultés ces dernières années en mode expert lorsque j'interrogeais la recherche avancée, mais avant toute chose, je voulais témoigner surtout sur le retour d'usagers et de collègues, parce que lorsqu'on a lancé la version 2 de notre bibliothèque diplomatique numérique, j'ai créé beaucoup d'ateliers en interne, et j'ai commencé à former les usagers de façon à pouvoir leur mettre le pied à l'étrier sur la recherche avancée, parce que je m'apercevais en salle de lecture, les contacts que j'avais avec les chercheurs et le grand public qui fréquentent également nos salles, c'est que le formulaire de recherche avancée leur semble très complexe. Il y a beaucoup trop de choses, alors que moi, ça ne m'avait jamais effleuré l'esprit, en fait, jusqu'à présent, parce qu'étant dans une démarche plus professionnelle, bon ben voilà, j'ai un cadre de saisie. Je le maîtrise plus ou moins bien, mais en tout cas, ça ne m'effraie pas. Clairement, pour eux, ça les effrayait. Ils lançaient plutôt des recherches simples. Et là où j'ai été étonné, finalement, c'est que aussi des collègues bibliothécaires lancent prioritairement des recherches simples et ne vont pas sur la recherche avancée pour affiner leurs recherches. Ça m'a plutôt étonné. En tout cas, c'est un retour d'expérience. Après, dès l'instant où on leur présente le formulaire de recherche avancée, ça leur semble assez intuitif et ils n'ont pas de difficultés majeures. De mon côté, j'ai été confronté à quelques difficultés. Alors, je vous les donne. Comme P. H., effectivement, la possibilité de forger des requêtes extrêmement complexes pour avoir le résultat le plus pertinent ça vient à me manquer quelques fois j'aimerais bien par exemple que sur les données du catalogue, on peut interroger sur plus de 10 occurrences articulées avec des opérateurs booléens, ça c'est quelque chose qui me freine quelque fois vers une recherche ... avec un beau résultat de requête que je peux mettre du coup dans le Drupal pour faciliter l'accès à des corpus pour nos usagers. Là aussi je rejoins L. H. du coup on passe aussi par le Drupal pour faciliter l'accès à certains corpus. Les troncatures, alors là c'est un vrai sujet, les troncatures, les guillemets, la recherche exacte avec un résultat qui m'a semblé très curieux. Alors c'est vrai que la requête était particulièrement tordue et je suis arrivé à un résultat qui n'était pas cohérent. Donc j'avais fait remonter à P. H. Martin et puis à Isabelle Mangou, ça concernait en fait Metternich qui était aux affaires étrangères en Autriche-Hongrie au 19e siècle. Et donc la recherche avec guillemets et la recherche exacte m'ont donné des résultats incohérents. Donc ça aussi je pourrais vous le

recommander. Remontez du coup cette requête de manière à avoir votre expertise et me dire en quoi j'ai fauté sur cette requête pour avoir le résultat le plus cohérent. Dans l'ancien formulaire de recherche avancée j'aimais bien que la langue soit dissociée des données bibliographiques parce que c'était un module à part avant d'être encapsulé dans le menu déroulant de la recherche catalogue et texte intégral. On ne peut pas faire de requêtes sur plusieurs collections numériques, ou plusieurs établissements de conservation. J'aurais bien aimé aussi avoir la possibilité de faire une requête mixée à ce niveau-là. Ce que j'ai aussi identifié comme difficulté c'était sur la recherche géographique. Alors là pour le coup alors c'est vrai qu'elle est un petit peu (Alors ça c'est pareil, je suis d'accord avec K. H.), dans le formulaire elle est un peu basse, c'est une nouveauté, donc elle s'est logée là où elle pouvait. Il faut vraiment faire de la médiation pour la faire comprendre. Moi par exemple, intuitivement, j'ai eu du mal à comprendre que si je ne sélectionnais pas le rayon de la recherche, je ne pouvais pas lancer la recherche. Donc le simple fait de saisir la ville, même si j'avais sélectionné ensuite le nom qui m'était proposé, si je ne sélectionnais pas le rayon de la recherche, ça ne marchait pas. Et puis, qu'est-ce que j'avais d'autre ? C'est au niveau du rapport de recherche. Là c'est un usage post-formulaire de recherche avancée. Je suis extrêmement gêné parce que je génère beaucoup des rapports de recherche. Mais dès l'instant où je tape sur de la presse et revue, j'ai tous les fascicules qui remontent. Et il est impossible de faire un export en tableau Excel qui soit propre, voire même faire un export tout court. Donc ça c'est vraiment un frein. Comme P. H., je sectionne, je divise, et puis après je recompose. Et c'est quand même assez chronophage. Donc je crois avoir tout dit. Mais effectivement, la médiation c'est vraiment indispensable. Si on ne fait pas cet effort, même vis-à-vis de collègues qui sont censés être des professionnels, s'ils sont dans le monde des archives, ils ont une expertise sur les services informatisés des archives, mais pas sur Gallica. Et il faut leur donner quelques clés pour interroger. Il faut leur donner quelques clés pour interroger de manière plus savante le formulaire de recherche avancée. Et puis, comme L. H. au Havre, j'organise prochainement des ateliers en interne de géolocalisation. Donc le premier aura lieu la semaine prochaine. J'ai été aidé par Enora justement dans le montage de ces ateliers de géolocalisation. Donc je pourrai faire du retour, même si on a testé déjà en interne la géolocalisation. Ce qui m'a gêné, mais je n'ai pas réussi à reproduire ce matin, cette... cet exemple, c'est que j'avais interrogé Isabelle, il y a quelques mois, justement, sur la géolocalisation, qui ne devait porter normalement que sur des images et sur des cartes. Or, quelquefois, je voyais le petit globe où la possibilité de géolocaliser des articles qui n'étaient ni des cartes ni des images. Mais je n'ai pas réussi à le reproduire. Donc peut-être que vous avez

fait un travail ces derniers temps sur la géolocalisation pour exclure les autres types de documents qui sont les cartes et les images. C'est tout pour moi.

Eléonore : C'est déjà très bien et très complet. Merci beaucoup. Arnaud, tu voulais réagir sur la recherche géographique ?

Arnaud : Je vais répondre sur deux points, là, précisés par L. C., qui ne concernent pas le moteur de recherche, juste pour y penser. Donc sur... Concernant la recherche géographique, effectivement, il y a une amélioration d'interface, on va griser le bouton, en fait, pour que ce soit actif. Et aussi, les utilisateurs... ne remplissent pas le champ de rayon pour qu'on comprenne qu'on est dans un parcours. Donc ça devrait normalement être plus explicite. Le bouton sera actif que si on a rempli le rayon. Et par ailleurs, pour le rapport de recherche, effectivement, les résultats sont déployés en mode dégroupé puisque l'objectif du rapport, c'est de pouvoir exporter les données et notamment tous les identifiants. Et en fait, dans ce cas-là, on a une limite d'export qui est à 50 000 résultats et donc la difficulté que vous avez avec les titres de presse qui ont énormément de fascicules, c'est que vous êtes au-delà des 50 000 résultats et que ça vous oblige en effet à segmenter votre requête par lot.

L. C. : Oui, mais tu vois Arnaud, par exemple, si l'export se faisait uniquement sur le titre de la presse ou du périodique et non pas sur l'ensemble des fascicules, ce serait beaucoup plus simple, ça réglerait le problème.

Arnaud : Oui, mais ça, c'est en fait, dans l'état actuel de l'architecture technique, ce n'est pas possible parce qu'en fait, on a des clés de regroupement qui s'appliquent à posteriori au niveau de l'affichage. Et donc, cette fonctionnalité, en fait, elle n'est pas appliquée sur le rapport de recherche. Le rapport de recherche utilise des API et du coup, ce n'est pas au niveau de l'interface de Gallica et de l'affichage des résultats qu'opère la clé de regroupement. Donc, ce n'est pas possible dans le rapport de recherche. Donc, ce qui serait envisageable, c'est d'augmenter la limite d'export. Il faut voir si, effectivement, on peut, avec les nouveaux serveurs, aller au-delà. Je l'ai noté.

Eléonore : Les améliorations, ce sera en temps 3. On va essayer de continuer pour tenir le timing. Donc, est-ce que quelqu'un veut prendre la suite ? Est-ce que quelqu'un veut livrer ses usages ?

L. S. : Bonjour, moi, Provence-Alpes-Côte-d'Azur, Musée du sport, L. S., je souscris à ce qui a été dit. J'ajouterais, on a mentionné les recherches iconographiques. J'ai l'impression qu'on

manque un petit peu d'outils en termes d'exploration iconographique. Je ne parle pas forcément sur nos fonds, mais les fonds de la BnF et de ses partenaires, c'est sûr que les fonds d'agences de presse pour l'histoire du sport, c'est encore plus iconographique majeur. J'ai l'impression de manquer d'outils assez simples pour de la recherche iconographique, de type des orientations, paysages, portraits, et le choix entre couleur et noir et blanc. Ce sont des choses qui ont été développées pour GallicaPix, que, parfois, j'utilise parce qu'il y a un corpus sport qui est indexé. Mais c'est vrai que ça, ça serait très confortable, et la possibilité de télécharger des corpus d'images en masse. Je trouve la fonctionnalité de téléchargement d'images extrêmement chronophage. On est sur quatre clics pour une image, sauf s'il y a un bouton magique, dans ce cas-là, prévenez-moi, il n'y a pas de possibilité simple de télécharger. Il y a encore plus d'images. Ça, ça me semble... Pour moi, c'est majeur, en tout cas, et dans l'utilisation, et dans celle de mes collègues, et dans celle de nos usagers. Je me demandais aussi, à la fois, il y a cette tension, évidemment, entre... On réclame une interface plus simple pour le moteur de recherche avancé, mais on réclame toujours plus de critères. Et je me demandais s'il avait été évoqué la possibilité de... Enfin, de diviser la recherche avancée en deux, une recherche avancée et une recherche experte, comme on l'a parfois sur certaines bases, qui peut-être ajoutent de la complexité, mais c'est vrai qu'il y a, entre les tensions, eu... Enfin, voilà, les... Je pense que c'est difficile d'articuler ce type de retour. Qu'est-ce que... Attendez, c'est-à-dire... Comme graphique... Peut-être également, là, c'est le côté musée, on a peu de finesse en termes de typologie d'objets ou de documents. À première vue, en tout cas, la recherche image, c'est un corpus extrêmement disparate entre les affiches, les cartes, les fonds photographiques, les dessins... Voilà, là-dessus, je... Ensuite, dans une logique musée, évidemment, nous, les collections qu'on a chargées, sur Gallica, nos collègues sont étonnés qu'on ne puisse pas faire de recherche par dimension, par exemple. Mais... Mais... Mais ça, c'est lié à notre... À notre domaine, et à... Donc voilà, mais en tout cas, c'est vrai qu'il y a un peu plus de finesse en termes de typologie, notamment de... De collection iconographique, ça me semblerait bien vrai.

Eléonore : Ok, très bien. Donc effectivement, comme le disait Fanny dans le chat, peut-être qu'on pourra répondre à ce qui est fouilles d'images après.

Fanny : Ou peut-être j'en dis deux mots maintenant ? Donc la recherche par l'image, c'est un vrai questionnement, et c'est à l'ordre du jour dans notre service. On a accueilli depuis le mois de mars une chargée de projets qui travaillait sur le projet qui s'appelle Gallica image, Gallica fouilles d'images. On vous en avait déjà parlé, je crois, quand on s'était vus à l'INRAE la dernière fois. Le projet est en cours. Normalement, on va tester les premiers prototypes avec les

fonds images. Donc le but étant de chercher par l'image, et non pas par le texte, des images dans Gallica. Par l'image, ça veut dire pouvoir interroger les corps puisqu'ils sont des images, mais aussi les images, les visuels dans la presse, dans les imprimés, etc. Les premiers prototypes sont en train d'être faits. Les premiers prototypes sont en train d'être faits. Et pour l'instant, on interrogera les images, donc les corpus images. En grande majorité, je crois qu'on va commencer par la BnF, par les estampes. Mais en tout cas, on est conscient de tout ça, que la recherche par l'image, c'est une recherche que font beaucoup les utilisateurs. On pourra en reparler aussi sur les résultats du baromètre des publics 2023 qu'a fait la BnF. Mais c'est une recherche qui est importante. Il y a beaucoup d'utilisateurs qui cherchent de l'image et uniquement de l'image dans Gallica. Et j'imagine que dans les marques blanches, c'est la même chose. Donc on a une problématique là à laquelle il faut répondre pour que l'image soit plus facilement accessible, les images. Ça amène beaucoup de questions et on pourra rentrer dans les détails plus longuement, peut-être plus tard, peut-être à l'automne, quand on se verra. Qu'est-ce qu'une image ? Qu'est-ce qu'on segmente dans une page, etc. Mais c'est en cours de réflexion, gestation, etc.

Arnaud : Et peut-être juste pour le bouton magique, en fait, il y a des plugins. Un plugin IIIF, je vais mettre dans le chat le lien pour pouvoir le télécharger et le mettre sur Firefox ou sur Chrome, qui vous permet de télécharger les images. Alors une par une. Mais une par une, sans passer par les quatre clics du téléchargement. Donc ça, ça peut faciliter votre exploration et vos récupérations des images.

L. S. : Ça, je l'utilise. Oui, mais ça, je l'utilise. Je sais qu'il me semble qu'il y a une possibilité, il y a un script pour récupérer en masse les images.

Arnaud : Absolument.

L. S. : Mais ça, je ne l'ai pas...

Arnaud : Là, c'est vrai que nous, on a eu des restrictions parce qu'on a fait appel à... On a subi beaucoup d'attaques, en fait, de serveurs et notamment de téléchargements en masse. De téléchargements en masse de nos données. Et donc, c'est vrai qu'il y a eu des restrictions sur l'usage de ces API. Et on est en train de mettre en place un serveur spécifique API pour améliorer le service. Puisque tout était sur la... Comment dire ? Les téléchargements massifs de données, en fait, impactaient le service de Gallica lui-même. Donc ça a été l'objet de restrictions.

L. S. : Merci.

Eléonore : Est-ce que vous voulez ajouter des choses ?

L. S. : Non, dans un premier temps, c'est bon pour moi.

Eléonore : C'était déjà plutôt complet sur la recherche d'images. Est-ce que quelqu'un veut réagir à ce qui a été dit ou nous livrer son témoignage ?

C. S. : Oui, bonjour. C. S., je m'occupe de Memonum, la marque blanche de la bibliothèque de Montpellier, du fonds patrimonial. Je rejoins beaucoup ce qui a été dit. Effectivement, sur l'iconographie, on travaille beaucoup là-dessus. Et c'est vrai que ce n'est pas toujours évident d'interroger plus par typologie, effectivement, les images. Ce serait vachement bien. Nous, on utilise quand même beaucoup la recherche simple et la possibilité d'affiner, qui finalement permet parfois de contourner la complexité de la recherche avancée. Je rejoins aussi là-dessus les collègues sur... Moi, je n'avais pas compris non plus. Là, on change géographiquement. On change de graphique avec le rayon de la recherche. Le format qui déstabilise parfois nos usagers. Ils ne comprennent pas trop exactement à quoi ça va correspondre. Donc, voilà, ce sont des points assez importants que je voulais souligner. Peut-être aussi dans le rapport de recherche, quand on affiche des extraits en mode texte, ça fonctionne bien, mais ça ne fonctionne pas très bien en mode images. Elles ne ressortent pas. Là, j'ai réessayé hier et je n'avais pas les visuels, en fait. Sur le rapport de recherche. Pour le reste, quand on est passé en marque-blanc, c'est assez récent. On a pu faire des ateliers pas mal communiquer. Et le public avait plutôt un bon retour quand même sur le système de recherche. Une recherche assez simple, on va dire. Mais pas trop redondante, effectivement. J'approuve relativement ce qui a été dit en amont.

Eléonore : Merci. Fanny, Arnaud, vous avez des questions ? Pas particulièrement. D'accord. Très bien. Peut-être, est-ce qu'il y a des fonctionnalités qui ne sont pas comprises ? Ils ont l'air d'être plutôt satisfaits. Est-ce qu'il y a des choses qui bloquent, quand même ?

C. S. : Je reviens sur la recherche avancée, en fait. Là, le format, la recherche géographique, parfois aussi l'incompréhension par rapport à la collection. La collection thème, les gens ne comprennent pas forcément à quoi ça correspond derrière. Donc, ils ne vont pas l'utiliser, en fait, pour filtrer un peu plus leur recherche.

Eléonore : Ok.

C. S. : Voilà.

Fanny : Peut-être un chiffre à vous donner de ce qu'on en sait. La recherche simple, donc au niveau de Gallica, la recherche simple, c'est à peu près 20% des usages. Et 3% des usages, enfin des usagers, passent par la recherche avancée.

C. S. : D'accord.

Fanny : Donc, c'est minime. La majorité des personnes arrivent sur un document via Google Images. Donc, sur les 3%, on est dans un public spécialisé ou expert. On appelle expert, qui a compris effectivement qu'il pouvait passer par la recherche avancée et qu'il avait d'autres critères possibles pour que la recherche soit simple. En majorité, les gens utilisent la recherche simple ou arrivent sur Gallica via Google.

L. S. : Ça veut dire que, pardon, 80% des gens arrivent sur Gallica via Google Images ?

Fanny : Non, pas 80% parce qu'il y en a qui ne consultent pas forcément de documents. Donc là, on a mélangé un peu tout. Il y en a qui arrivent par le catalogue général de la BnF. Mais on va dire que les gros pôles, il me semble que c'est 40%.

Arnaud : Oui, c'est ça. Ça dépend de la manière dont on compte, mais c'est entre 40% et 50% qui arrivent directement par Google.

Fanny : Via un moteur de recherche. 20% via la recherche simple. Et 3% via la recherche avancée. Après, c'est un peu plus disparate pour le reste.

Arnaud : On a à peu près 15% par les sites afférents, c'est-à-dire notamment par les sites de généalogie pour une part importante du public qui fait des recherches généalogiques.

Fanny : Les réseaux sociaux, etc.

Eléonore : Effectivement, parler de la recherche dans Gallica, c'est déjà s'adresser d'une certaine manière à une niche, surtout si on parle de recherche avancée. Nos retours sont très intéressants là-dessus parce que c'est vrai qu'on voit que pour certains, ça semble évident de passer par la recherche avancée. Pour les usagers, ça ne l'est pas forcément. Les retours d'experts sont très intéressants. Les retours d'usagers également. Est-ce que quelqu'un veut enchaîner ?

G. T. : Bonjour. G. T.. Je suis avec A. et D. pour Laborar. La marque blanche toute fraîche née à Cambrai puisqu'on est en ligne depuis seulement trois mois. Et je rebondis sur ce qui vient d'être dit parce qu'effectivement, ça correspond majoritairement à notre usage en mode recherche simple vu que notre marque blanche adresse avant tout le grand public. Et donc vraiment, on donne de nombreux ateliers de présentation de la navigation et surtout de la

recherche depuis le lancement de la marque blanche. Et on instille vraiment les usagers à passer par la recherche simple et ensuite à affiner avec les facettes. Ce qui effectivement permet de contourner les difficultés de lecture de la grille de recherche avancée. Alors bien sûr, il y a des cas, mais encore une fois des cas de niche où on utilise la recherche avancée. Tant nous-mêmes que parfois un public effectivement qui a des recherches très spécifiques à faire. Mais voilà, on a un cas un petit peu particulier parce qu'Isabelle pourra confirmer mais on a peu de métadonnées en dehors des titres et du texte océrisé sur la presse. Donc du coup, la recherche simple qui porte sur une recherche tout intégrale fonctionne assez bien pour nous. J'avais une petite question sur la recherche géographique parce que de mon côté, en fait, le bouton recherche géographique est toujours grisé que je précise un rayon de recherche ou pas. Et je me pose vraiment la question, mais c'est sans doute lié à l'architecture technique, de pourquoi ce bouton lancer la recherche géographique est-il présent en supplément du lancement de la recherche classique ? C'est un petit peu complexe à lire du coup pour un utilisateur. Et les autres remarques que j'avais encore à faire, oui. Sur le texte océrisé qui est très utile pour nous parce qu'on a un fond de presse conséquent. On a donc déjà, il faut vraiment guider l'utilisateur qui veut faire cette recherche-là pour qu'il choisisse bien les bons champs, c'est-à-dire texte intégral en fait dans le premier bloc catalogue et texte océrisé. Mais en même temps, on a un deuxième bloc de recherche avancée qui s'intitule par proximité et qui est systématiquement présent. Je ne sais pas si ce serait de l'ordre du faisable d'avoir une modalité où effectivement on indique qu'on va d'abord faire une recherche océrisée et alors seulement il y a des critères supplémentaires, notamment la recherche par proximité qui apparaissent. Ça améliorerait peut-être la lisibilité de la liste. Les autres remarques également que j'avais sont sur justement qu'on l'utilise beaucoup nous, c'est-à-dire la recherche simple et ensuite on affine par facette. Ergonomiquement, j'apprécierais beaucoup qu'on puisse modifier la recherche sans devoir revenir en arrière jusqu'à la grille initiale de recherche. En tout cas, je n'ai pas réussi à modifier ma recherche autrement jusqu'à aujourd'hui. Et puis quand on fait une recherche en texte intégral, alors c'est très sympa parce qu'on a des aperçus en fait, des extraits de texte qui vont surligner le terme qui a été recherché. En revanche, quand on veut saisir un terme supplémentaire dans affiner la recherche, là les aperçus disparaissent. Donc on nous représente des documents, mais on n'est pas absolument sûr que ces documents comportent le terme complémentaire qu'on vient de rajouter dans affiner la recherche. Et ça, c'est un petit peu frustrant. Voilà, je pense que j'ai à peu près tout dit.

Eléonore : Ok, c'est très bien. Effectivement, ça complète bien ce qui a été dit, comme quoi la recherche est vraiment multiple dans Gallica. Arnaud, Fanny, vous avez des choses à ajouter ? Non, pas particulièrement. Ok, donc on peut passer au retour suivant. Oui ?

Arnaud : Non, sur le highlight, en fait, c'est parce qu'il y a plusieurs highlights qui sont dans le texte. Et donc, c'est la plus forte occurrence du terme qui ressort en premier. Et donc, c'est pour ça que vous ne voyez pas le terme que vous avez recherché, qui pourtant figure dans le texte, mais qui n'est pas visible en premier.

G. T. : D'accord. Mais l'aperçu, du coup, initial, qui comportait effectivement le terme initial surligné, il disparaît complètement, en fait, quand on a rentré un terme supplémentaire dans affiner la recherche. On ne voit plus rien. Il n'y a même pas le terme précédent qui, effectivement, avait peut-être une occurrence supérieure, mais on n'a plus d'aperçu du tout. En tout cas, moi, c'est comme ça que ça se matérialise.

Arnaud : D'accord. Écoutez, on regardera ça précisément, parce que normalement, ce n'est pas le comportement normal.

G. T. : D'accord.

Eléonore : Alors, qui veut prendre la suite ?

B. D. : Moi, je veux bien, si vous voulez, pour l'INSEE.

Eleonore : Allez-y, oui.

B. D. : Donc, moi, je m'occupe de la BNSP, c'est la Bibliothèque numérique de la statistique publique, donc de l'INSEE. Et puis, tous les services statistiques ministériels. Donc, je commence par les points positifs. Donc, la recherche plein texte, y compris sur les PDF, parce qu'on a beaucoup de publications en format PDF sur la BNSP. Donc, ça, c'est quelque chose qu'on attendait depuis l'ouverture de la BNSP en 2021. Donc, c'est chose faite. Et du coup, ça permet aussi la recherche par proximité en intégrant tous les documents PDF. Donc, ça, c'est très bien. Et puis, on a aussi une recherche par auteur, aussi, qui est apparue cette année. Donc, ça, c'est les points positifs. Ensuite, donc, moi, j'ai peu de retours des usagers extérieurs à l'INSEE. Moi, c'est surtout des collègues proches, en fait, qui font des recherches. Je m'interrogeais sur la qualité mode de texte qu'on a quand on a la liste des échos, en haut, on a par défaut pertinence. Donc, déjà, par défaut pertinence, je trouve que c'est... Moi, j'aurais préféré avoir par défaut, par date décroissante pour avoir les derniers documents. Là, on a par pertinence, mais en fait, je n'ai jamais compris ce qu'était la pertinence. Et donc, dans ces tris,

on a la qualité mode texte. Et bon, là, c'est pareil. Je ne sais pas trop à quoi, enfin, comment derrière la requête se fait. Mais en tout cas, j'ai tous les derniers documents. Donc, les derniers documents. C'est les derniers téléchargés qui apparaissent en dernier dans cette liste. Donc, comme si la qualité des derniers documents était moindre que les anciens documents. Je ne sais pas comment c'est trié. Ensuite, on a la date d'édition, en fait, quand on fait un tri par date d'édition, c'est par année et on n'a pas les mois. Et donc, à l'INSEE, c'est un petit peu embêtant parce qu'on a des publications mensuelles, typiquement l'indice des prix qu'on fait tous les mois. Et si on fait un tri, on a l'indice des prix un petit peu mélangé sur les 12 mois. Ensuite, oui, en fait, quand on clique sur informations détaillées, quand on a les échos et que je clique sur ne voir les résultats de cet auteur, il n'y a rien qui s'affiche. Et quand je regarde la requête, il me semble que dans la requête, il y a des mots en trop "Auteur du texte" qui est en trop dans la requête. C'est-à-dire qu'on a l'Institut national de la statistique et des études économiques, Auteur du texte entre parenthèses France. Il faut que j'enlève "auteur du texte" pour que la requête fonctionne.

Chloé : Je me permets une toute petite incise, c'est Chloé. Je me permets juste de dire que ce bug-là est en cours de correction. Il va être mis en prod prochainement.

B. D. : Merci. Bonne nouvelle.

Eléonore : Déjà une réponse.

B. D. : Et ensuite, oui. On n'a pas le nombre de pages sur les fichiers PDF dans la notice, en fait. C'est quelque chose qui nous manque. En plus, on les télécharge quand on envoie à la BnF. Voilà, en fait, j'ai terminé.

Eléonore : Oui. Donc surtout des problèmes d'incompréhension de certaines fonctionnalités. Et sur la liste de résultats qui n'est pas forcément suffisamment fine, il manque du grain pour les dates, les mois, etc. Est-ce qu'il y a des recherches qui sont incohérentes ? Est-ce que vous avez des cas précis, en fait ?

B. D. : Non, les recherches ne sont pas incohérentes. C'est juste le tri, en fait, qui n'est pas bien fait puisque le tri se fait par année. Et on aurait aimé l'avoir par mois, année. Je parle de la date d'édition.

Eléonore : Oui.

Arnaud : Je reviendrai sur la question de la pertinence et de l'ordonnancement des résultats tout à l'heure. Concernant le tri, je ne sais pas si c'est possible et envisageable d'avoir un calendrier

comme on l'a pour la presse. Pour, effectivement, les tableaux qui sont sous forme de fascicules. Ça, c'est peut-être à envisager. En tout cas, le problème du tri par date, c'est que du coup, le critère de pertinence ne s'applique pas. Il peut y avoir beaucoup de bruit si on tri par date. Mais j'y reviendrai.

T. B. : Je veux bien prendre la suite de B. D.. T. B. directeur de la bibliothèque A. D. de l'INSEE. Donc, naturellement, je me retrouve complètement dans ce qu'a dit B. D. et dans beaucoup de choses qui ont été dites jusqu'à présent. Recherche plein texte, oui, j'en suis satisfait. Pour l'iconographie, c'est pareil, j'ai quelques réserves, je trouve qu'effectivement, on perd un petit peu de temps. Il pourrait y avoir peut-être des choses un petit peu plus pratiques. Mais dans l'ensemble, j'en suis assez content. Moi, j'ai des recherches extrêmement simples. Mais je pense qu'elles concernent beaucoup de nos lecteurs et statisticiens qui ne sont pas des experts en recherche documentaire. Donc, je n'ai pas trop honte de dire que la plupart du temps, ce que j'ai besoin de chercher, c'est tout simplement un document que je connais très bien. Donc, je connais soit le titre, soit l'auteur. Et là, effectivement, que ce soit en recherche simple ou avancée, je ne sais pas comment faire. Mais je me retrouve avec des résultats qui sont noyés dans un ensemble. Donc, je ne comprends pas la pertinence. Je n'arrive pas à les reclasser. Donc, on a des stratégies de contournement. Sur notre BNSP, on a la possibilité d'explorer par thème, région, pays, période. Mais également, alors, producteur. Donc là, il faut entendre un petit peu auteur-collectivité. Et puis, par auteur. Donc là, auteur-personne physique. On met l'auteur, le prénom et le nom et le prénom. Et on est bien sûr de tomber uniquement là-dessus. Et pas sur plein d'autres choses. Effectivement, quand je cherche dans Gallica, alors il y a des fois des possibilités d'explorer aussi avec des outils. Un que j'aime beaucoup qui a été évoqué, c'est le calendrier pour la presse quotidienne que j'utilise beaucoup. Mais voilà. Sinon, si je veux faire des recherches basiques, c'est un petit peu idiot. Mais je vais sur Google ou quand c'est un article, je vais sur Openedition. Et là, après, je retrouve. Ça me renvoie simplement. Je retrouve le lien pour Gallica. Mais je n'arrive pas à le faire sur Gallica. Et sur la BDSP, j'y arrive parce que maintenant, on a ces outils-là. Au niveau de l'ergonomie, après les documents qu'on a trouvés, il y a deux petites choses. Moi aussi, je me suis fait avoir plusieurs reprises au début sur la visibilité de tourner les pages. Et quand ça fait longtemps que je n'y suis pas allé, c'est vrai que spontanément, je me refais avoir. C'est vrai qu'il y a un petit peu quelque chose. Et il y a quelque chose du même ordre pour enregistrer ou imprimer les résultats. Quand je recherche un article qui se trouve dans un volume d'un périodique, j'arrive bien tout de suite, mon résultat est sur la page qui s'affiche. Et si je fais enregistrer, on dit enregistrer du début. Et à chaque

fois, je me fais avoir. Je mets du début. Et puis alors, il faut que j'aie vu combien de pages. Parce que bon, là, ce n'est pas renseigné. Donc, je dois compter mes pages. Je reviens. Je marque du début oui jusqu'à la fin de la page visibilité. Et naturellement, je me retrouve avec le début du volume lui-même et pas mon article. Donc, voilà. Donc, là, quand je le fais une fois de suite, après, j'ai compris. Mais ça, ce n'est pas très pratique.

Eléonore : Alors, effectivement, vous disiez que c'est idiot de passer par Google. Mais alors, pas du tout. Puisque, comme l'a rappelé Fanny, c'est le cas de très nombreuses personnes, y compris les conservateurs à la BnF. C'est important de le rappeler. Et puis, si ça fonctionne, en fait, ça peut être une stratégie tout à fait valide. On n'a pas forcément besoin de passer par la recherche avancée. Une des sociologues de la BnF le rappelle, les usages numériques sont vraiment propres à chacun. Et donc, toutes les stratégies sont bonnes tant qu'on arrive au résultat. En tout cas, c'est ce qu'on espère justement, c'est de pouvoir arriver au résultat. Est-ce que, Fanny et Arnaud, vous avez des questions ? Pas particulièrement. OK. Alors, on n'a pas entendu C. L. et E. L., je crois.

C. L. : Du coup, C. L., École nationale des ponts et chaussées. Alors, moi, je suis à la fois sur la partie build-up numérique et je suis archiviste. Donc, du coup, je vais avoir une petite réflexion déjà, en tant qu'archiviste, c'est que je trouve que la cote n'est pas assez visible quand on arrive sur les résultats de recherche. Mais ça, c'est vraiment un réflexe d'archiviste. On a beaucoup de documents qui sont uniques en tant qu'archives. Et donc, du coup, la cote, c'est vraiment le truc qu'on considère quand même comme vraiment un marqueur. Alors, je sais que c'est moins le cas forcément côté bibliothèque. Mais en tant qu'archiviste, c'est un petit manque que j'avais trouvé. Sinon, en termes d'usage, de recherche, etc. À titre personnel, j'utilise à la fois la recherche simple et la recherche avancée. Ça dépend de ce que je cherche. Je sais que des fois, avec la recherche simple, si je tape tel terme, je vais retrouver les documents qu'on a dans les premiers résultats. Mais ça dépend vraiment de ce que je cherche derrière. Mes collègues, c'est un petit peu pareil. Après, au niveau des retours utilisateurs, je vais un petit peu revenir sur ce que les autres ont déjà dit. C'est-à-dire qu'on a beaucoup d'utilisateurs qui trouvent que c'est compliqué, en fait, de s'orienter dans Gallica de manière générale et dans le moteur de recherche en particulier. Aussi, il faut que ça soit un petit peu compliqué au niveau de la navigation. Par exemple, quand on tombe sur des périodiques, selon sur quoi on clique, on va arriver sur la vue où on a d'abord, par exemple, par année, pour après rentrer au niveau des dates. Des fois, ça va être une vue en liste. Et en fait, au départ, on ne comprend pas très bien pourquoi, quand on clique à un endroit, on tombe sur la vue en liste. Et pourquoi, quand on

clique sur un autre, on tombe sur un autre, on tombe sur cette vue par date. Donc, il y a peut-être un petit truc là juste dans l'affichage à améliorer. Alors sinon, j'ai quand même un cas très spécifique d'une recherche où je ne comprends pas concrètement les résultats. En fait, deux recherches qui n'en renvoient pas les mêmes résultats, alors que je pense qu'elles devraient renvoyer les mêmes. Donc là, c'est assez spécifique à nous, je ne sais pas si c'est le cas pour les autres marques blanches. Mais en fait, si je fais une recherche vraiment dans notre marque blanche à nous, dans Héritage, si je fais une recherche toutes les collections avec l'ensemble des documents, c'est-à-dire vraiment renvoyer tout ce que nous, on a envoyé dans Gallica, j'ai 17 158 résultats. Si je fais, pas par titre, mais vraiment par fascicule. Alors que si je fais la recherche en allant mettre en fait comme collection, notre collection de base en fait, qui est ENPC 00, avec trois O, pardon, derrière, je vais avoir 17 172 résultats. J'en ai plus si je mets en fait notre code projet principal que si je fais une recherche dans l'ensemble de nos collections. Alors si ça avait été l'inverse, je me serais dit c'est que des fois, on a oublié de mettre le code projet ENPC 000. Mais là, en fait, j'ai l'impression que c'est plus en fait au niveau du moteur de recherche. Peut-être que je me trompe ? Mais du coup, j'ai l'impression que c'est peut-être au niveau du moteur de recherche où il y a des documents qui me sont sortis quand je mets ENPC 000 qui ne ressortent pas dans... À vrai dire, je ne sais pas trop si c'est au niveau de l'indexation ou du moteur de recherche, mais en tout cas, il y a un petit truc derrière peut-être à voir où serait l'origine. Sinon, ce que je m'étais notée, c'était que la recherche par proximité, juste avant l'atelier, en fait, je sens que j'ai re-regardé la recherche avancée en regardant notamment ce que je n'utilisais pas, dont la recherche par proximité, j'ai effectivement trouvé qu'elle était très fonctionnelle, mais j'avoue que je n'avais pas encore l'usage. Et pareil pour la recherche géographique, j'ai mis aussi un petit peu de temps à comprendre qu'il fallait mettre quelque chose dans la partie rayon pour que la recherche fonctionne. Et du coup, tout de suite, je me suis posé un petit peu des questions. C'est-à-dire que nous, quand on remplit les données pour les notices du catalogue général, il y a à la fois le champ indexation, on peut mettre un sujet des lieux, et les coordonnées GPS. Et en fait, j'avais une petite question, peut-être que ce serait juste à clarifier au niveau de l'affichage, qu'est-ce qui pourrait être recherché, ou est-ce que quand on met une indexation géographique, ça ressort quand on fait une recherche comme ça par localisation ou pas ? Oui, c'est juste que j'avais relevé à ce qu'avait dit L. S. sur l'éventualité, en fait, de distinguer recherche experte et recherche avancée, on va dire plus simple. Je ne sais pas non plus si c'est forcément la chose la moins complexe à mettre en place pour la suite, mais en tout cas, ça peut être une bonne piste pour effectivement distinguer des usages experts de l'usage d'utilisateurs qui sont un petit peu perdu quand ils vont aller sur les outils. Et pour finir sur un

petit retour aux usagers, ça nous arrive en fait assez souvent, parce que nous, en fait, les usagers, ils nous envoient très facilement des mails pour faire des recherches, etc. Et donc, ça arrive qu'ils disent, on va chercher un document que je n'ai pas trouvé dans Gallica. Sauf que nous, on sait que le document est là. Donc, en fait, on fait la requête pour eux, on leur envoie après. Mais en fait, ça veut dire que concrètement, eux, ils cherchaient un document, ils ne l'ont pas trouvé. Donc, c'est pareil qu'au niveau du moteur de recherche, tout n'est pas clair quand on est moins expert.

Eléonore : Il y a un retour qu'on a eu, également. Donc, oui, c'est vrai que certains restent bloqués, notamment sur la recherche simple. Ils n'arrivent pas à modifier leur recherche, trouver des stratégies de contournement par manque de culture numérique, de connaissance de la collection, par problème lié à Gallica aussi. Enfin, c'est assez complexe, mais ça, on l'entend bien. Et on essaie de travailler dessus.

Fanny : Juste rapidement, avant qu'on passe à la dernière personne. Si je comprends bien, C. L., il y a plus de résultats quand on utilise la recherche avancée en choisissant les collections de l'École des Ponts et Chaussées que quand on choisit le site de consultation.

C. L. : En fait, du coup, c'est une recherche que je dois faire manuellement. En fait, je dois faire la requête moi-même. Si je fais la recherche avec vraiment la collection, en fait, c'est le code projet derrière, le principal, le ENPC 000, j'ai un nombre de résultats qui est plus important.

Fanny : Oui, mais là, le problème, le problème, il est facilement visible. C'est que, je viens de le faire, 5 327 résultats avec le code projet collection de l'École nationale des Ponts et Chaussées. Et il y a dedans quelques documents qui ne sont pas de l'École des Ponts et Chaussées qui ont été indexés avec ce code projet. Par exemple, j'ai Numelio.

C. L. : Mais du coup, toutes les collections, est-ce qu'on a quand même... Nous, dans l'héritage, on a, juste en passage, la part bibliothèque au site où on a l'École nationale des Ponts et Chaussées qui est en gras parce que c'est nous. Mais on a quand même, en fait, les autres institutions où... Enfin, si on mettait toutes les collections, ça comprenait, en fait, les collections du coût des partenaires et des indexations.

Chloé : Je voudrais juste dire un petit mot. Est-ce que ce n'est pas simplement, C. L., qu'un certain nombre de documents qui portent le code projet générique de l'ENPC ne portent que ce code projet générique et ne sont pas inclus dans une des collections ? Il doit y avoir un certain nombre de documents qui n'ont que ce code générique. Je pense que ça serait la raison pour

laquelle il y en aurait au moins quand on sélectionne... Parce que quand on sélectionne toutes les collections, on ne sélectionne que les collections, au final. Mais on regardera ça ensemble.

Fanny : Oui, mais je ne pense pas que ça soit... Je pense que ce n'est pas une question du moteur de recherche.

Chloé : Je ne pense pas. On regardera ça, C. L., ensemble.

Fanny : On doit pouvoir corriger ça sans... Mais je ne pense pas que ce soit le moteur de recherche.

Chloé : Oui, oui, je pense que c'est différent.

L. C. : Fanny, excuse-moi, mais effectivement, je rebondis sur ce qu'a dit C. L.. Lorsqu'on fait une requête dans le formulaire de recherche avancée, en ce qui me concerne, sur ma marque blanche, par tous les établissements, tu as un gap avec une même requête sur toutes les collections numériques. Normalement, on devrait tomber sur le même résultat de 500 en l'occurrence. Et ce n'est pas normal. C'est-à-dire que tu fais tous les établissements, tu trouves 13 350. Tu passes à la corbeille. Et d'ailleurs, j'ai oublié de parler de la corbeille pour nettoyer ton formulaire de recherche avancée. Et tu refais la requête sur toutes les collections numériques. Et tu as donc un gap de 500. Et normalement, on devrait en fait être à l'équilibre. On a fait tous les établissements ou toutes les collections numériques. Donc il y a un truc en fait.

Chloé : Non, parce que c'est la même chose. S'il y a des documents qui ne portent que le code générique et qui ne portent pas de second code projet, ils ne rentreront pas dans toutes les collections. Puisque les collections, c'est ce qui est défini par les codes projets secondaires. Donc on peut avoir...

L. C. : Tu veux dire, Chloé, que s'il y a eu un loupé et que seul le code générique, M-A-E-D-I-G-E-N-0 pour moi, a été adossé sans sous-collection en fait.

Chloé : Voilà. Et bien, il ne rentrera pas dans cette recherche par collection.

L. C. : Ah, ok. Donc en fait...

Chloé : Les collections sont définies par ces codes projets secondaires.

L. C. : D'accord, ok. Très bien, j'ai compris. Mais alors comment je vais pouvoir tester les 519 ?

Chloé : A retester. On peut revoir ça ensemble. Mais je pense que c'est vraiment la réponse à cette requête. Les collections, en fait, c'est ça. Ce sont ces codes projets secondaires.

L. C. : Je suis d'accord. Alors que moi, j'ai intégré le code d'entrée dans les collections. Alors moi, ce qui m'intéresserait, c'est de pouvoir du coup récupérer les 519 qui n'ont pas de code projet dit secondaire, indexé, parce que c'est un défaut de DPI, parce que j'ai tout indexé moi-même. Donc je sais...

Chloé : On va pouvoir le faire en faisant code projet générique, moins code projet secondaire, troisième, quatrième, etc. Enfin, sauf, etc. Puis on pourrait les faire sortir comme ça. Mais on reparlera de ça ensemble.

L. C. : Ok. Et alors du coup, Corbeil, effectivement, à l'usage, je dois toujours insister auprès des collègues et des usagers, pour que lorsqu'ils lancent une requête, le résultat ne leur convient pas, ils reviennent en arrière et puis ils changent certaines occurrences sur leur formulaire. Eh bien, ils oublient qu'il faut avoir préalablement, souvent, nettoyer l'affichage, donc effacer la recherche préalable pour pouvoir repartir sur un bon pied parce qu'effectivement, ils ne pensent pas à cette corbeille pour nettoyer le formulaire.

P. H. : Et... Et une gomme serait peut-être plus logique qu'une poubelle.

L. C. : Ah non ! Oh non !

P. H. : Et la poubelle, je n'aime pas, mais voilà.

L. C. : Non, en plus de ça... Bah, moi, je pense que le terme "effacer la recherche" qui apparaît simplement dans une petite fenêtre, si elle était comme "lancer la recherche", "effacer la recherche", déjà, ce serait mieux. On a vraiment un pavé explicite. Parce que sinon, s'ils ne passent pas la main sur l'icône, ça n'affiche pas "effacer la recherche".

Eléonore : Et c'est un retour qu'on n'avait pas encore eu, ça, donc très bien, on le prend. Merci.

C. S. : Ça peut aussi être peut-être juste réinitialiser.

L. C. : Ouais.

C. S. : Voilà, un terme peut-être plus générique, plus clair pour tout le monde.

Eléonore : Et donc... Oui ?

L. S. : Pardon, c'était juste pour ajouter, j'ai oublié de vous dire à quel point j'aimais la recherche par proximité. Et qui, pour moi, c'est mon point de vue, enfin, c'est mon point d'entrée principal

dans les collections. Et notamment, je l'utilise pour contourner les soucis dont on a parlé de recherche exacte ou de guillemets. Souvent, j'ai tendance à faire une pression exacte à une distance maximale de un mot. Enfin, plein de choses, ça me permet de contourner. Peut-être simplement qu'on ajoute, en termes de recherche, dans la recherche par proximité, j'imagine que l'opérateur par défaut, c'est « et ». Je cherche à la fois cette proximité et cette deuxième. Mais en termes de résultats, je ne retrouve pas le « et ». En tout cas, ce ne sont pas les résultats les plus pertinents les résultats qui s'affichent en premier souvent ne contiennent qu'un des deux termes. Je ne sais pas. Mais en tout cas, pour moi, c'est vraiment un point d'entrée majeur. Dans les collections et séries, je ne sais pas si mes camarades ont le même usage ? Et j'aurais tendance à le mettre plus en avant. Moi, c'est vrai que je l'utilise beaucoup pour répondre à des usagers. Je conseille souvent la recherche par proximité et elle est plus visible alors qu'elle me semble très utile.

Eléonore : Ok, merci. Peut-être Élodie ?

E. L. : Bonjour, je suis É. L.. Je travaille à la bibliothèque numérique Numba. C'est la bibliothèque du CIRAD, dont on a un grand milieu tropical à Montpellier. Personnellement, je n'ai pas trop de soucis sur la recherche pour trouver des documents. On a surtout des ouvrages, quelques iconographies, un petit peu, mais pas énormément. J'utilise autant la recherche simple que la recherche avancée. Pas de difficultés particulières. Si ce n'est que je sais qu'une collègue, par défaut aussi, utilise beaucoup la recherche booléenne, qui ne fonctionne pas très bien. Et je voulais remercier le collègue L. C. qui vient de me faire découvrir la poubelle. Je ne savais pas du tout qu'elle existait. C'est vrai que je viens d'aller voir et au survol, il y a “effacer la recherche”. Je trouve que ce bandeau gris en bas, j'ai tendance à l'oublier. Avec trop de coupé-partie, trop de volume au fascicule, j'ai tendance à ne pas du tout l'utiliser. Peut-être pas très visible. Au niveau des retours utilisateurs, on en a très peu au CIRAD. On a quelques retours sur des documents qui n'ont pas été trouvés, justement. Donc, un peu comme les collègues, on fait la requête à la place. Les chercheurs. Et puis, on... On fournit le document. Je pense que ce qui peut un peu impressionner aussi, c'est tous les critères qui sont possibles d'interroger dans la recherche avancée. Bon, au contraire, moi, je trouve que c'est assez... Ça permet de faire beaucoup de filtres. Donc, voilà. Je pense que c'est un peu les utilisateurs qui ne sont pas expérimentés, qui n'en ont pas l'habitude. Donc, voilà.

Eléonore : C'est aussi ce que constatent les collègues qui aident à la BnF, à la stratégie et à la recherche. Justement, ils disent un manque de culture numérique, de compréhension de ce qu'est Gallica, le catalogue, etc. Donc, oui, c'est des choses qu'on comprend. Peut-être, quels

sont vos cas d'usage ? Comment vous effectuez vos recherches ? Donc, recherches simples et avancées. Est-ce qu'il y a des choses qui sont pertinentes ? Ça fonctionne plutôt bien, j'ai cru comprendre ?

E. L. : Oui, oui, ça fonctionne plutôt bien. Je n'utilise pas tous les champs. C'est vrai que la recherche par proximité, je ne l'ai jamais utilisée. Il faudrait que je regarde de plus près. Après, les autres champs, oui. Date de mise en ligne aussi, c'est très utile pour nous pour faire de la vérification sur ce qui est entré en marque blanche ou pas. La recherche géographique aussi, c'est vrai que j'ai un peu du mal, donc je ne l'utilise pas. Sinon, tous les autres champs, pas de souci.

Eléonore : Ok, très bien. Donc, Arnaud, tu me dis, mais peut-être qu'on va passer aux questions slash explications du moteur rapidement.

Chloé : Je voulais vous couper deux petites secondes justes pour dire à L. S., L. S., l'opérateur entre les deux, quand tu as deux, quand tu as deux champs proximité, c'est bien un "et". Et tu disais, si je ne me trompe pas, que ça avait l'air d'un "ou", c'est ça ?

L. S. : Les résultats étaient... J'ai l'impression... Dans la recherche par proximité. Quand tu as deux... Quand il y a deux recherches par proximité différentes, j'ai l'impression que les résultats que j'obtiens à première vue, c'est un "ou".

Chloé : Ok, mais c'est intéressant parce que dans la requête, c'est bien un "et". Donc, il faut regarder ça de plus près. Merci, L. S..

L. H. : Est-ce que je peux juste faire une toute petite incise ? J'ai un repentir je voulais parler juste de deux choses. Quand on fait la recherche dans Bambou, tout simplement rechercher, en fait, dans les autres sites d'Internet, c'est le gros bouton qui sert à faire valider. Et parfois, c'est un peu frustrant parce qu'il faut cliquer sur la loupe. Ce n'est pas hyper intuitif, ça, parce que dans tous les autres sites, c'est le bouton. Parfois, on se dit, ah non, j'ai encore cliqué sur le mauvais bouton. Ça ne doit pas être la seule à qui ça arrive. Et l'autre point, je l'ai mis dans le chat, c'est justement la recherche avancée. Elle n'est pas utilisée par tout le monde. Et en fait, la recherche par institution, elle n'apparaît que dans la recherche avancée. Et en fait, moi, j'aimerais bien qu'elle puisse apparaître dans la liste de résultats comme sur Gallica, surtout quand on a une bibliothèque numérique qui accueille beaucoup d'institutions comme chez nous. Voilà, je me dis, comme ça existe sur Gallica, ce n'est peut-être pas si compliqué de le mettre aussi, de le proposer aussi sur les marques blanches.

Eléonore : C'est noté.

Fanny : C'est-à-dire sur les facettes sur la gauche, c'est ça ?

L. H. : Oui.

Eléonore : Comme sur Gallica, oui.

L. H. : Si on veut juste les documents de la bibliothèque, on peut l'avoir. Si on veut ceux des archives des musées ou bien de la BnF, voilà, ça permet de trier de manière intuitive. Alors qu'honnêtement, si on n'est pas un usager expert, on ne sait même pas qu'on peut le faire. Enfin, la possibilité de le faire n'est pas apparente.

Arnaud : Oui, merci à vous toutes et tous pour vos remarques et votre témoignage qui relève, comme on l'avait vu dans les ateliers précédents, à la fois du moteur de recherche, de son fonctionnement du point de vue plus technique, mais aussi au niveau des interfaces et des IHM. Et puis, enfin, au niveau, on pourrait dire plus de la médiation, du mode d'emploi et de la formation. Là, je vais revenir, juste pour vous donner quelques éléments, quelques clés de compréhension du moteur de recherche en tant que brique logicielle et puis ensuite répondre à vos questions.

Temps 2 : Explications du moteur

[Explications d'Arnaud]

Temps 3 : Améliorations

Eléonore : Dans ce cas-là, je vous propose qu'on passe au temps 3. Les améliorations. Comme Chloé vous l'a écrit dans le mail, on vous propose de nous donner une liste de trois améliorations qui vous semblent être prioritaires, particulièrement importantes, qu'elles concernent vos collections, etc., spécifiquement, ou qu'elles soient plus générales. Et, comme pour les retours pendant le temps 1, ça peut être spécifique au moteur ou alors plus dédié à la médiation autour des différentes fonctionnalités, l'interface en machine, etc., des améliorations qui permettraient de rendre la recherche en délicat et dans les marques blanches peut-être plus simple, j'espère, etc. Vous avez déjà donné quelques idées dans vos différentes prises de parole. Maintenant, on essaie de synthétiser pour trouver les retours les plus importants.

P. H. : Nous, la liste au Père Noël, ce serait la navigation plus facile dans les résultats affichés, de passer du résultat 1 au résultat 2 en appuyant sur une flèche et pas revenant à la liste des résultats pour passer au 2. Plus on perd de temps, plus on est susceptible de perdre des gens.

Eléonore : La fonctionnalité est la plus importante. Est-ce qu'il y en aurait une autre ?

P. H. : Honnêtement, c'est un problème qu'on a tous. On fonctionne tellement dans nos outils qu'on est sur du jargon métier, par bibliothèque, par collection ou thème. Kézako, comment on fait pour rendre ça intelligible ?

Fanny : Je suis bien d'accord.

P. H. : Mais ça, on a tous le même problème. On passe notre temps à se dire non, ça c'est du jargon. Et quand on passe de bibliothécaire à documentaliste et archiviste, il n'y a pas le même et on ne se croise pas.

Fanny : Oui, et de la même manière, je sais que vous ne pouvez pas échanger grand-chose, mais les sigles sont parfois incompréhensibles par les utilisateurs. J'ai un utilisateur qui est arrivé déjà là, c'est 3%. Mais sur le titre et la manière de présenter des choses, on pourrait corriger assez facilement. On vous a dit tout à l'heure, les témoignages et les évolutions, c'est sur le moteur de recherche, mais ça peut aussi empiéter sur le reste parce qu'il y a des choses auxquelles nous, on ne pense pas. On a la tête dans le guidon et quand on prend du recul, ça peut être pertinent de changer certains paramètres ou certains mots que l'on présente. Mais en termes de médiation, je suis tout à fait d'accord, il y a des choses qu'on pourrait revoir sur la page d'accueil et sur la manière de présenter la recherche.

Eléonore : J'ai oublié de préciser, mais ça peut aussi être des fonctionnalités, des évolutions qui n'existent pas encore. Par exemple, imaginez des fonctionnalités d'IA, ou alors des fonctionnalités que vous retrouvez sur d'autres sites, notamment les sites e-commerce, et qui vous plairaient bien dans Gallica. Pour donner un exemple, la fourchette de prix qui est affichée sur beaucoup de sites e-commerce, a été assez demandée en étant transformée en fourchette de date pour pouvoir faire une recherche qui soit un petit peu plus précis. Ce genre de fonctionnalités-là peuvent également faire partie de votre liste au Père Noël.

Arnaud : Agate, j'avais noté aussi des difficultés sur les requêtes, notamment avec les booléens, et sans doute une attente sur la troncature qui va peut-être être citée par d'autres participants, d'autres partenaires. Le principe des expansions, de multiplier les requêtes par des appels au dictionnaire pour ramener les variations orthographiques, pose des difficultés pour la troncature

puisque ça multiplierait d'autant plus les requêtes. Là, on a une limite technique dans la forme actuelle du moteur de recherche. On a ce choix des expansions et du rappel très large qui nous contraint au niveau des troncatures, comme je le disais, un milliard de termes dans l'index. On a des problèmes de performance pour pouvoir tout calculer avec cette fonctionnalité-là.

P. H. : Les expansions me dérangent. Pour ce que j'ai à faire sur Agate, pour gérer mes collections, mes affichages, ça me dérange. Mais parce que j'ai un type de travail qui est particulier, ça doit être utile à un lecteur, alors que blanche, ça me dérange. C'est parce que c'est un usage extrêmement pointu. Le dernier truc pour Noël, c'est vraiment... je pinaille. La recherche de base n'est pas en recherche exacte quand on est sur le moteur de Gallica, sauf pour la recherche géographique. Si on n'affiche pas exactement le terme attendu par la recherche géographique, on a zéro. Le temps de réponse est assez long. Après ça, je sais qu'il y a des jours où c'est un peu plus long. Mais si on met juste Saint-Martin-Vésubie par exemple, qui a encore pris une lave torrentielle. Ce n'est pas la première, ce ne sera pas la dernière. Si on ne met pas Saint-Martin-Vésubie et qu'on met juste Saint Martin Vésubie parce qu'on ne sait plus où sont les... On ne trouve pas. Donc soit c'est exact, soit ça n'est pas.

K. H. : Voilà, plus qu'on y est, moi aussi j'ai un vœu. Peut-être que ça a été abordé pendant mon absence. Moi j'imagine bien une liste des résultats que Gallica nous sort, qu'on puisse cocher des notices qui nous intéressent, les mettre dans un panier et de les faire ressortir plus tard. Par exemple, je fais une recherche sur un Pyrénéiste. Je trouve ce livre, cette image, tout ça, et au milieu je trouve plein d'autres choses qui ne m'intéressent pas. Voilà, de pouvoir les cocher, de les mettre de côté et d'alimenter une liste personnalisée.

Fanny : Est-ce que vous connaissez le compte personnel ?

K. H. : Moi je ne l'utilise pas. Ça existe là-dedans ?

Fanny : Oui, c'est ça qu'on peut faire.

K. H. : Très bien.

Fanny : Mais moi non plus je ne l'utilise pas.

P. H. : On est toujours dans la profession et on ne pense plus à des choses particulières.

Isabelle : Il n'est pas accessible sur les marques blanches.

Fanny : C'est ce que j'allais dire. Il n'est pas sur les marques blanches mais sur Gallica, c'est exactement ce à quoi il sert, c'est-à-dire rassembler des recherches favorites, des termes favoris, d'avoir une liste de résultats favorites, etc.

K. H. : Je rajoute sur ma liste ça pour les marques blanches.

Fanny : Je ne sais pas cela dit si on peut exactement, par exemple, cocher une liste de 10 résultats. On n'en a que 8 qui nous intéressent. Je ne sais pas si on peut exactement cocher les 8 ou s'il faut le faire individuellement. Mais en tout cas, on peut avoir une historique des recherches, tout ça, via le compte personnel. Moi je ne l'utilise pas non plus.

Eléonore : L. S. ?

L. S. : Non, je disais, je crois qu'on est d'accord qu'on ne peut pas cocher, par exemple, 10 résultats sur 12. Ce n'est pas possible. On peut sauvegarder les documents, mais un par un, à partir du visualiseur, c'est ça ? Pour l'espace personnel.

Fanny : Il faut que je me penche dessus, mais j'ai l'impression qu'on ne peut pas cocher. Non.

Isabelle : J'aurais dit que oui moi.

Fanny : Mais j'ai un doute en le disant. J'ai l'impression que oui, mais...

L. S. : Il me semble qu'on peut le faire à partir du catalogue, mais pas de Gallica.

Fanny : Ça, c'est sûr, le catalogue.

Fanny : Dans tous les cas, effectivement, sélectionner ce qui semble être la liste de résultats qui nous intéresse, enfin, les résultats qui nous intéressent, ça, ça pourrait être une évolution intéressante.

Arnaud : C'est possible au niveau du rapport de recherche. C'est-à-dire qu'une fois que vous avez exporté la liste de résultats dans le rapport de recherche, là, vous pouvez sélectionner les documents qui vous intéressent, et vous pouvez le faire. Dans l'espace personnel, il y a eu des évolutions fonctionnelles récemment. Je n'ai pas eu l'occasion de me reconnecter depuis. Mais dans l'espace personnel, c'est plutôt un panier. C'est-à-dire que vous mettez en favori les documents qui vous intéressent ou les requêtes. Vous pouvez sauvegarder vos requêtes et les rejouer depuis l'espace personnel. Mais il n'y a pas d'export de la liste de résultats dans le...

Fanny : Oui, mais je ne sais pas si, par exemple, quand tu as fait une requête, tu peux sélectionner la liste des résultats qui t'intéresse et n'avoir que la liste des résultats qui apparaît.

Isabelle : Non, je suis dessus et non. Ce n'est pas cliquable.

Fanny : Par contre, on peut le faire dans le catalogue,

K. H. : Donc mon vœu reste valable.

Fanny : Oui, il reste valable, K. H..

Eléonore : Y a-t-il des demandes à ajouter ?

Arnaud : On a eu des retours, on a eu beaucoup de retours des collègues en interne justement à réfléchir à une convergence des fonctionnalités entre le catalogue général et Gallica. Donc, ça aussi, c'est des choses qu'on a en tête.

L. H. : Dans la liste au Père Noël, je me suis déjà exprimée, mais vraiment, honnêtement, je ne dois pas être la seule, mais j'y tiens à ma facette institution parce que quand on veut vraiment voir ce qu'il y a chez nous pour un usager ou bien... C'est une manière aussi de valoriser le fait que... Enfin, parce qu'aux yeux des usagers, une marque blanche, son contenu, ce n'est pas forcément si évident que ça. Donc, le fait qu'il existe une facette, ça permet de comprendre qu'il y a des fonds de plusieurs institutions. On peut se dire, ah bah tiens, je ne savais pas, en fait, qu'il y a aussi des documents d'une autre bibliothèque ou d'un centre de recherche. Ça fait aussi une médiation sur le contenu des marques blanches.

G. T. : Alors, puisque c'est le moment des vœux au Père Noël, moi, j'ajouterais bien quelque chose, effectivement, comme on a déjà la chance d'avoir pour contenu textuel la recherche par proximité, est-ce qu'on pourrait rêver, évidemment à moyen ou long terme, d'une recherche par proximité de contenu pictural ? J'entends par là qu'on puisse interroger

le moteur de recherche en langage naturel pour lui dire, peux-tu trouver toutes les images qui représentent un blason ?

Fanny : On y travaille !

G. T. : Oui, j'imagine bien.

Fanny : Voilà, donc ça, c'est le projet Gallica Fouille d'Images, c'est d'utiliser l'intelligence artificielle, d'entraîner l'intelligence artificielle sur des corpus pour faire reconnaître des formes et que ces formes puissent être interrogées. Par exemple, un parapluie, que ce soit dans un manuscrit, je vais prendre une licorne, ou un soleil, que ce soit dans un manuscrit médiéval, ou dans une photographie de l'agence Rol, etc. Mais, évidemment, on va y aller étape par étape et

du coup, le plus simple au début, c'est quand même de commencer sur un corpus qui est déjà un corpus d'images. Donc, de photographies, d'affiches, enfin, un type de document qui est déjà de l'image. Ça commencera par-là, et oui, normalement, l'idée c'est de l'étendre à la presse, aux imprimés, etc. Alors, après, la recherche, entre guillemets, par proximité, bah oui, en fait, on pourrait la faire en se disant qu'il va, voilà, je reprends mon idée du soleil, on sélectionne un manuscrit médiéval dans lequel il y a un soleil, on pourrait très bien avoir une proposition de tous les autres soleils dans les autres manuscrits médiévaux. Mais ça, ça va nécessiter d'entraîner l'IA à obtenir des résultats pertinents. Et de faire, on va dire, strates par strates, d'abord sur des images, ensuite sur de la presse, sûrement, ensuite sur des monographies, etc. Mais oui, c'est le projet. Mais peut-être qu'on pourra vous en dire plus à l'automne sur ce travail-là.

G. T. : On est impatients de voir ça !

Fanny : Oui, nous aussi. Nous-même, nous cherchons beaucoup d'images.

L. H. : Et c'est une question, en fait, est-ce que vous avez des projets un peu plus fantaisies, je dirais, pour le grand public ? Ou on aurait, par exemple, un avatar avec lequel on peut discuter de bibliothécaires qui feraient des recherches pour vous, en intelligence artificielle ?

Fanny : Alors ça, on en a parlé aussi, Arnaud, dans le premier atelier qu'on a fait. On a un chercheur, un post-doctorant qui réfléchissait à toutes ces choses-là et notamment à désambiguïser certains termes pour qu'on comprenne quelle catégorie il regroupait. Par exemple, le nom de ville. Je prends l'exemple toujours de Orange. Est-ce que Orange est un agrume ? Est-ce que c'est une ville ? Est-ce que c'est une couleur ? Etc. Et dans les options qui avaient été réfléchies, Arnaud, tu pourras en dire plus, peut-être, parce que je n'ai pas tout suivi. Il y avait la notion de chatbot via l'intelligence artificielle et d'autres bibliothèques numériques ou d'autres institutions publiques qui ont des collections en ligne. On aura déjà mis en place des choses.

Arnaud : Oui, oui, tout à fait. Peut-être que vous l'avez remarqué dans Gallica, dans votre interface, quand il y a une recherche qui est ambiguë, il y a un bloc en haut à gauche qui vous dit voulez-vous dire telle chose ou telle chose et qui permet de, quand vous tapez par exemple Victor Hugo, voulez-vous Victor Hugo comme auteur ou comme sujet, qui permet déjà, dans ce bloc de désambiguïstation, de pouvoir mapper ce que vous avez tapé avec ce qui est indexé par le moteur de recherche. Là, le prototype sur lequel travaille le post doc, c'est de pouvoir améliorer avec un modèle de langue cette proposition de désambiguïstation. On attend un prototype qui va être présenté à l'automne. On pourra revenir vers vous pour vous en dire plus

sur comment ça peut se déployer, enfin, déjà comment ça peut fonctionner et comment ça peut se déployer. Après, dans les perspectives ouvertes par l'IA, alors là, on est vraiment dans un temps à plus long terme. Ce ne sont pas des choses qui sont en expérimental, enfin, en prototypage. Donc, il y a la possibilité effectivement d'utiliser des modèles de langue, soit pour faire de l'indexation, soit pour faire effectivement de la... proposer des résultats de recherche. Alors, peut-être que vous connaissez le prototype qui a été mis en place par la bibliothèque du Luxembourg. On pourra partager avec vous des liens si vous voulez le tester. Et donc, ils ont déployé un chatbot, enfin un chat GPT, une application de chat GPT sur un corpus documentaire donné de presse. Et donc, quand on interroge donc le chat, la boîte de dialogue, les résultats qui sont proposés puisent dans le corpus qui a été sélectionné et source en fait la réponse en fonction des titres dans lesquels l'algorithme est allé chercher les résultats. Et donc là, on a un usage tout à fait intéressant de de... qui va permettre d'identifier en fait les contenus à l'intérieur même des textes en fait, enfin des résultats à l'intérieur même des textes. Et dans les perspectives ouvertes aussi par les modèles de langue, on a donc Fanny en parlait, la similarité visuelle, donc la recherche par similarité de forme pour trouver les blasons ou les images similaires. On a aussi de la similarité de contenu de texte donc avec des algorithmes qui vont pouvoir comparer des notices, qui vont pouvoir comparer des résumés ou même des contenus même des livres et pouvoir faire de la recommandation par similarité de contenu. Donc ça, ce sont des perspectives qui sont ouvertes par ces évolutions technologiques et sur lesquelles on va se pencher aussi pour voir comment on pourrait prototyper ou utiliser ce type de nouveaux outils mais on est vraiment dans une démarche beaucoup plus prospective.

Fanny : Et juste par rapport à ChatGPT, tout à l'heure, je ne sais plus qui disait que pour des requêtes, on ne pouvait pas mettre plus de 10 termes, etc. Je ne sais plus qui disait ça. Bon, je ne devrais pas le dire, ce n'est pas très corporate, mais ChatGPT peut faire les requêtes dans ces cas-là.

P. H. : Ah, pas mal !

L. S. : Est-ce que tu peux donner un exemple ?

Fanny : J'ai la formulation, je vous l'enverrai. Je l'enverrai à Chloé et Isabelle qui vont la transférer. En fait, il faut cibler les documents que l'on veut... Alors, c'est dans le cas où on connaît les documents et on veut faire une requête qui rassemble des documents. Il faut cibler les documents, on lui donne la liste juste des identifiants ARK et il fait la formule s'il peut y en avoir, il fait la requête, s'il peut y en avoir 100, il va faire la requête.

Isabelle : Ah, d'accord. Parce qu'on a dit 10, mais en fait c'est plutôt au bout de 8 que ça plante.

Fanny : Bon, je ne devrais pas le dire, mais je le tiens d'I. D.. Donc, je crois que j'ai un peu de légitimité. Dis-moi L. C..

L. C. : Oui, Fanny. Alors, effectivement, tu mets le doigt sur le nerf de la guerre, la médiation, la formation. Parce que moi, j'ai été confronté avec mes collègues. Là, effectivement, dans tous nos échanges, on aperçoit qu'on apprend les uns des autres. Le coup de ChatGPT, jamais ça ne m'aurait traversé l'esprit.

Fanny : Moi non plus. C'est ma prédécesseure qui m'a donné l'astuce.

L. C. : Et puis, je retiens en fait trois choses aussi, tu vois. Par rapport à tous les tutos, parce qu'il y a beaucoup de choses sur le blog Gallica pour nous apprendre comment utiliser les formulaires de recherche avancée. Mais je pense que sur la géolocalisation, j'ai été confrontée à un défaut de tuto parce qu'il y avait assez peu de choses finalement qui étaient disponibles en ligne. J'ai bénéficié de l'expérience de nos collègues du Havre pour élaborer le mien. J'en ai élaboré un moi-même que j'ai renvoyé à Enora juste avant qu'elle ne parte. Et puis, j'ai dû interroger la médiation Gallica pour dire, voilà, je suis un petit peu ennuyé, donnez-moi des pistes. Là, je pense qu'il y a quelque chose à faire pour bien géolocaliser parce que tant qu'à faire les choses, il faut les faire correctement. Pour la recherche pas proximité, je pense qu'il faudrait quand même bien que ce soit clair pour tout le monde que ça ne concerne que les documents qui ont été osérés, donc des imprimés parce que c'est dans la petite bulle mais je pense qu'un certain nombre d'entre nous ont dû se planter là-dessus. Je m'apercevais, en faisant une petite recherche aussi dans le visualiseur, il y a deux choses que je n'utilise jamais à savoir la mise en scène, donc les petites icônes de droite par exemple une page de titre, on a l'icône de droite mise en scène et puis marque-page, j'ai essayé de regarder dans l'aide si je trouvais une explication, je n'en ai pas trouvé mais j'ai fait ça à la va-vite donc peut-être que ça existe mais je n'ai pas trouvé je ne sais pas à quoi ça sert et puis enfin pour répondre à un problème que C. L. soulevait et sur lequel j'ai rebondi entre une recherche tous les établissements ou toutes les collections numériques en fait je pense avoir le résultat lorsqu'on fait la recherche tous les établissements, ça exclut notre établissement à nous donc c'est pas tous les établissements parce que c'est tout sauf nous

Chloé : c'est tout sauf celui qui est en gras

L. C. : ben voilà, écoute c'est pas du tout easy-friendly !

Chloé : ah non non c'est pas du tout easy-friendly et c'est pas du tout explicite tout sauf le partenaire principal

L. C. : et là je retourne à peu près effectivement sur mes pas à trois items

Eléonore : merci.

Fanny : juste peut-être pour répondre à L. C. alors mise en scène c'est un outil qu'on a développé peut-être que vous vous souvenez de cet outil qui s'appelait enfin qu'on avait appelé stories slash à la loupe qui est en fait un outil qui permet de faire de la médiation sur le document en zoomant sur certaines parties ou en proposant comment dire comme un diaporama de documents sur lesquels on peut annoter des choses cet outil là il avait été créé par une société extérieure à la BnF et on a fait redévelopper notre propre version par la BnF par les développeurs du DSI et effectivement on l'a mis en ligne il y a eu un article dans le blog de Gallica mais on n'a pas dans l'aide à la recherche par exemple expliqué concrètement comment ça fonctionnait et ça ce chantier là on va l'ouvrir à l'automne parce que en plus de ce chantier sur le moteur de recherche je suis en train de travailler à la refonte de la palette d'accueil de Gallica de l'identité de Gallica et peut-être que ça on vous en reparlera à l'automne quand ce sera au point de sortir et dans cette optique là je suis en train de faire un peu le tri justement des informations qui sont données aux utilisateurs et les informations d'aide à la recherche c'est ce qui ressort dans toutes nos enquêtes baromètre fait par la BnF il faut expliciter la recherche donc il y aura un menu sur la page d'accueil dans lequel il y aura « aide à la recherche » avec déjà des tutoriels visuels des vidéos qui ont été dont Arnaud avait piloté la réalisation et qui s'adresse plutôt à un public on va dire étudiant/ chercheur et puis toute une explication je sais pas encore trop comment si on fera ça sous forme vidéo ou sous forme d'articles qui expliqueront pas à pas la recherche avancée etc c'est des choses qui existent déjà qu'on va réactualiser et donc du coup inclure dedans Gallicarama c'est une fonctionnalité qui a un peu moins d'un an. Je crois que j'ai répondu à tout ? Ah et marque page, il est lié au compte personnel.

Eléonore : L'heure tourne s'il y a plus de choses à ajouter sur la liste au Père Noël, je vais peut-être redonner la main à Arnaud pour la conclusion histoire de vous parler de la suite de l'enquête un petit peu et de ce qu'on va faire du matériau collecté aujourd'hui. [...] Oui L. S. ?

L. S. : moi c'était une question d'aperçu dans la liste des résultats alors c'est de nouveau lié à l'iconographie donc arrêtez-moi tout de suite mais d'avoir un affichage de la page du résultat alors je sais que c'est possible dans le rapport de recherche on a une image de la partie de la

page où apparaît le mot recherché mais parfois c'est utile d'avoir un aperçu de la page entière typiquement je cherche un nom de sportif j'ai envie de retrouver dans la presse des pages où il est photographié et non uniquement des pages de résultats sportifs où il arrive 53ème du Tour du Périgord c'est tout simple mais un aperçu de la page un aperçu visuel de la page de résultats.

Fanny : donc un aperçu mais qui est en lien avec ta recherche ? Tu connais sûrement la vue mosaïque donc ce n'est pas ça que tu aimerais c'est la page en question qui reprend le terme mais s'il y a plusieurs pages ?

L. S. : s'il y a plusieurs pages on est vraiment au format vignette, c'est juste pour pouvoir visualiser d'un coup si c'est une page qui contient de l'iconographie ou une page qui contient uniquement du texte typiquement je cherche... ?

Fanny : je comprends, c'est plutôt dans la presse.

L. S. : c'est plutôt dans la presse, dans les imprimés contenant de l'iconographie.

Fanny : ok

L. S. : Je ne sais pas si ça rendrait quelque chose de très ergonomique en terme d'affichage je ne pense pas on pourrait peut-être le limiter à un certain nombre de résultats

Fanny : On pourrait aussi imaginer qu'on l'a au survol plutôt parce que... ouais, ça me semble difficile à la place de la vignette en fait mais peut-être au survol, je ne sais pas... il faut y réfléchir.

L. S. : C'est des idées qui sont très certainement inférieures aux fonctionnalités que vous allez proposer dans le futur sur la recherche iconographique, donc vous allez faire bien mieux.

Fanny : De toute façon, comme toute liste au Père Noël, tous les cadeaux ne seront pas... [rires]
Non mais, on ne pourra pas c'est certain qu'on ne pourra pas tout faire il y a des choses qui regroupent vraiment ce qui a été dit en interne par nos collègues ce qui sera sûrement dit par les Gallicanautes un peu moins experts et on se focalisera sur ces points-là et après oui il y a du point de détail ou des choses qui pourraient être améliorées et il y a des choses qui ne concernent pas forcément le moteur de recherche donc qui peuvent être améliorées plus vite ou de façon différente enfin voilà même si c'est l'ergonomie même si c'est le visualiseur c'est utile aussi pour nous, parce qu'on n'a pas en fait beaucoup de temps où on se pose et on réfléchit à comment sont présentés les documents qu'est-ce qui est présenté etc, ... Arnaud ?

Arnaud : [conclusion]

Atelier Jeunes chercheurs : 12 juin 2024

Introduction

Support de présentation projeté lors de l'introduction et l'explication du projet aux participants.



Figure 2 : Support de présentation atelier Jeunes Chercheurs



Figure 3 : Introduction, support de présentation atelier Jeunes Chercheurs

L'enquête

SUR LA RECHERCHE DOCUMENTAIRE DANS GALLICA



Enquête interne auprès des collègues de la BnF : ateliers, entretiens, questionnaire



Enquête externe auprès des usagers : ateliers (ULille, Gallicanautes)



Production d'une synthèse des résultats pour construire une feuille de route sur les évolutions de Gallica

3

Figure 4 : Présentation de l'enquête, support de présentation atelier Jeunes Chercheurs

L'atelier

Jeunes chercheurs - ULille



3 objectifs

Temps 1 : Prendre en main la recherche dans Gallica sur un sujet réel (mise en situation)

Temps 2 : Collecter vos retours lors d'un tour de table (focus groupe)

Temps 3 : Laisser libre court à votre créativité pour proposer améliorations et évolutions de la recherche dans Gallica (discussion par groupe, tour de table)

4

Figure 5 : Présentation de l'atelier, support de présentation atelier Jeunes Chercheurs



Temps 1

EXPLORATION / PRISE EN MAIN

Durée 50 minutes

Choisir un thème parmi la liste proposée et faire des recherches sur Gallica pour trouver des documents pertinents par rapport à votre thème.

Vous avez deux objectifs :

- trouver entre 10 et 15 documents pertinents pour votre thème
- et trouver le document indiqué dans la colonne "Document à chercher" en rapport avec votre thème.

Pendant que vous explorez Gallica, remplissez votre guide d'exploration (papier ou numérique).

Il n'y a ni bonne ni mauvaise réponse !

5

Figure 6 : Temps 1, support de présentation atelier Jeunes Chercheurs

Temps 2

TÉMOIGNAGES

Durée 50 minutes

Retour d'expérience sur votre navigation et vos recherches dans Gallica.

Est-ce que c'était facile de trouver des documents en lien avec votre thème ?

- Est-ce que vous avez réussi à trouver le document indiqué ?
- Si oui, comment ? Quelle stratégie ? Quelles fonctionnalités ? Comment contourner les problèmes ?
 - Si non, quelles difficultés ? Quels sont les éléments bloquants ?

Quels sont vos échecs/ vos réussites ?

Qu'est-ce qui vous a plu ou pas ?



6

Figure 7 : Temps 2, support de présentation atelier Jeunes Chercheurs

Pause

RETOUR DANS 10 MINUTES

7

Figure 8 : Pause, support de présentation atelier Jeunes Chercheurs

Temps 3

PISTES D'ÉVOLUTION ET D'AMÉLIORATION

Durée 30 minutes

10 minutes : par groupe de 2 ou 3 dressez une liste de 3 à 5 propositions qui amélioreraient la recherche dans Gallica.

Sur les post-it à disposition, une amélioration par post-it. Il peut s'agir d'améliorer une fonctionnalité existante ou d'en ajouter de nouvelles.

20 minutes : restitution et discussion autour de vos propositions.

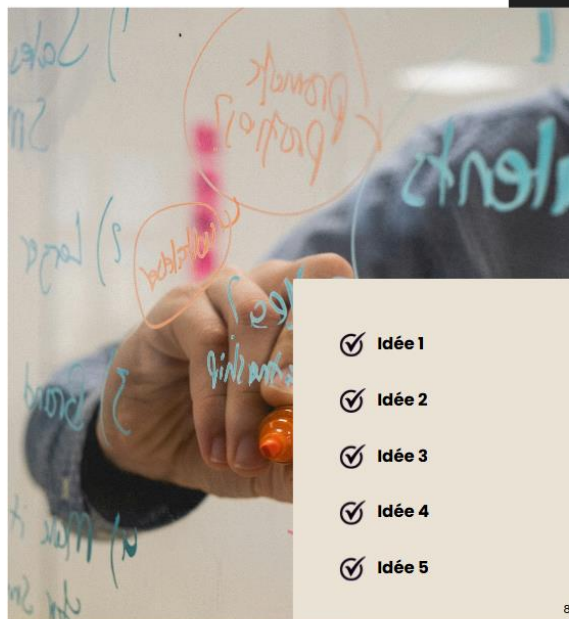


Figure 9 : Temps 3, support de présentation atelier Jeunes Chercheurs

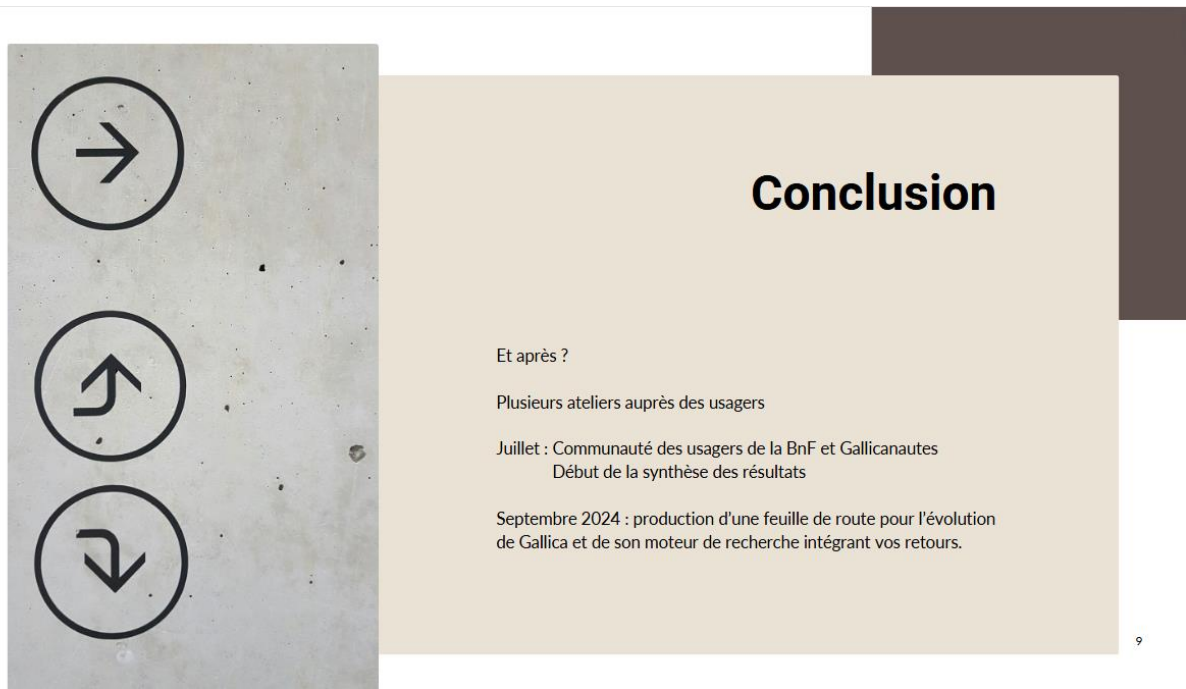


Figure 10 : Conclusion, support de présentation atelier Jeunes Chercheurs



Figure 11 : Remerciements, support de présentation atelier Jeunes Chercheurs

Partie 1 : Exploration

Transcription des feuilles d'observation des participants.

J. B. - Boléro, Ravel [Observation par Eléonore]

Tape son sujet dans la recherche simple “Boléro, ravel”, son document apparaît en premier en quelques secondes. Elle demande alors comment elle fait si elle a déjà réussi ?

Elle explore la liste de résultats, trouve beaucoup d'enregistrements mais voudrait d'autres documents. Elle passe beaucoup de temps sur la liste de résultats à faire défiler, ouvre certains documents dans de nouveaux onglets. Elle utilise le filtre “Presse et revue” pour trouver des documents qui ne soient pas des enregistrements sonores.

Elle utilise les filtres sur la gauche de la liste de résultats.

Elle trouve des documents liés à son thème et qui l'intéressent, elle les conserve dans de nouveaux onglets et constitue ainsi son corpus.

Elle cherche “Ravel” dans la RS pour trouver d'autres documents liés. Elle regarde la liste de résultats longtemps, très longtemps. Ouvre de tout, des documents pertinents et d'autres moins. Elle n'explore pas ce qui est en dehors de la recherche simple, la liste de résultats et le visualiseur de documents.

Elle s'arrête pour discuter, fait une pause sans réellement savoir quoi faire. Commence à remplir son guide d'exploration.

Elle recommence sa recherche, lit et survole les résultats dans la liste, va voir la recherche avancée pour le guide, arrive sur le blog, utilise quelques filtres mais dit ne pas les comprendre, elle trouve de nouveaux documents liés. Termine de remplir le guide. Elle a terminé et elle discute avec ses voisines.

N. K. - Les tartes sucrées et salées [Observation par Eléonore]

Chercher le thème dans la RS. Elle a beaucoup de bruit, fait défiler sa liste de résultats, ouvre les documents, cherche dans le visualiseur de document les informations qui l'intéressent, elle lit/ survole quelques documents.

Elle retourne dans la liste de résultats, relance une recherche, ouvre de nombreux documents dans plusieurs onglets (+ de 10 onglets ouverts simultanément). Elle survole les documents, en lit une partie, ouvre de nouveaux documents et ferme quelques onglets progressivement.

Elle cherche l'auteur du livre qu'elle cherche. En tapant son nom dans la recherche simple, elle trouve un autre livre de l'auteur mais pas celui qu'elle cherche.

Elle recommence sa recherche sur les tartes, elle ouvre de tout (plus ou moins pertinent pour le thème). Change de mots-clés dans sa recherche "tartes sucrées et salées", "tartes sucrées", "tartes salées", "tarte", "tartes"...

Elle ouvre le guide et commence à le remplir : retourne à la page d'accueil et explore, descend jusqu'au pied de page.

Retourne à la RS : cherche à filtrer "Tout Gallica" mais ne sait pas quel type de document elle cherche donc elle laisse "tout Gallica". Elle utilise la recherche avancée, essaie d'utiliser quelques filtres (par proximité, par date de publication, ...). Elle prend le temps de lire toutes les bulles d'aide contextuelle. Elle prend le temps de lire et de comprendre la recherche avancée. Elle tente la recherche par date de publication, ça ne fonctionne pas, elle réessaie "1977".

Elle trouve Gallicarama, l'ouvre et explore. Elle arrive sur le blog et fait défiler les articles, elle trouve un billet en rapport avec la cuisine et le lit.

Elle retourne à sa recherche et regarde toutes les fonctionnalités de la RA qui pourraient l'aider. Elle va voir dans le pied de page si elle trouve quelque chose pour l'aider.

Discute avec sa voisine "Je ne comprends pas..." "Moi non plus" en parlant du guide d'exploration.

K. G. : J'accuse... ! [Observation par Eléonore]

Prépare un document Word pour y entrer les références de son corpus.

Cherche son sujet dans la RS "J'accuse... !" Elle ouvre la liste de résultats, arrive sur le rapport de recherche et essaie de comprendre. Retourne à la RS, regarde Gallica vous conseille et ouvre plusieurs documents liés à son sujet. Elle arrive sur l'Aurore, l'ouvre dans le visualiseur de document, cherche la date du numéro avec le calendrier.

Elle cherche “procès de J’accuse... !” dans la RS, n’obtient aucun résultat, elle est déçue. Elle arrive sur le blog, trouve le tutoriel pour rechercher la presse dans Gallica, le lit attentivement, puis retourne dans la recherche simple. Elle retente sa recherche “J’accuse...!” dans la recherche simple mais ne trouve toujours pas ce qu’elle veut. Elle reste bloquée sur sa recherche longtemps, regarde la liste de résultats et ne sait pas quoi faire.

Elle retourne sur le document qu’elle a trouvé. Elle cherche à ouvrir sa recherche pour trouver autre chose mais n’y arrive pas.

Elle prend le guide pour le remplir, repart sur la page d’accueil, relance une recherche dans la RS “J’accuse... !”, regarde les résultats et soupire, elle arrive sur Gallicarama et repart. Elle trouve les sélections A à Z et cherche son sujet. Elle est bloquée par un temps de chargement très long.

Cherche à nouveau “J’accuse... !” dans la RS, regarde à nouveau Gallica vous conseille. Ouvre à nouveau la presse, regarde les fonctions d’export, retourne sur Gallicarama, recherche “j’accuse” dans l’OCR.

Elle cherche dans la RS “procès de J’accuse sorti du tribunal”, n’a pas de résultats, ... Regarde son téléphone et baille, termine son guide.

A. T. P. – Nintendo [Observation par Eric Kergosien]

Recherche les mots-clés, parcourt les métadonnées. Fait défiler les documents pour choisir.

Retourne sur les thématiques.

Retourne sur la page d’accueil. Passage par la collection. Retourne sur les thèmes. Retourne sur une collection “les consoles à cartouches” et navigation dans la collection, entre dans les documents.

Pour remplir le guide d’exploration : RA, arrivée sur la liste de documents. Retour sur l’interface. Recherche par le navigateur web. Il recherche des mots-clés puis des facettes. Regarde la rubrique “à découvrir” sur la gauche. Retourne à une banque d’image par la collection.

Résumé : usage de mots-clés puis usage des thématiques puis retour aux mots-clés.

usage de la collection “consoles à cartouches”

utilisation par moment des facettes (même légère)

stockage des documents dans un document annexe, copie les liens des pages.

E. T. - Les Jo de Paris, 1924 [Observation par Eric Kergosien]

Recherche simple “Jo Paris”, lecture des métadonnées des documents, ouverture de documents images.

Recherche de mots-clés ouvrage

Recherche dans les thématiques, parcours les documents par les métadonnées.

Retour sur les thématiques, retour sur la liste de documents, ouverture d’un document.

Retour sur la liste de documents, recherche thématique (longue). Entrée dans un document image puis retour à la liste de documents.

Création d’un compte, parcours disponible.

Retour sur la liste de documents.

Parcours interface accueil “les actualités”.

Retour à la liste de documents puis image puis retour à la recherche (France, etc)

Résumé : aucun stockage de document

regard rapide sur les documents

usage des mots-clés similaires

beaucoup de navigation entre mots-clés, liste de documents, documents et thèmes

analyse longue des thématiques

création d’un compte personnel

J. V. - Le jardin du Luxembourg [Observation par Eric Kergosien]

Barre de recherche “Jardin du Luxembourg”, défilement des documents, parcours les filtres à gauche.

Recherche mots-clé ouvrage.

Ouverture d'un document, lecture du début, ouverture d'un autre document, usage des filtres à gauche, parcours les métadonnées et ouverture d'un document. Parcours des documents images. Stockage dans la fonction "favoris" du navigateur.

Quand les favoris ont été choisis, création d'un fichier, liste de documents, sélectionné avec les métadonnées.

Résumé : Beaucoup de recherche dans les documents

Usage des filtres

peu de mots-clés

Travail intéressant de recherche d'abord puis de sélection et de stockage des documents.

N. H. - Les recettes traditionnelles dijonnaises. [Observation par Eléonore]

Ouvre Gallica, cherche son sujet dans la recherche simple "recettes traditionnelles dijonnaises", explore la liste de résultats, arrive sur le rapport de recherche, regarde les possibilités d'export du RR.

Cherche longuement (très) dans les sélections, regarde chaque entrée une à une pendant plusieurs minutes. Trouve "Écrits Gourmands", s'y attarde mais ne trouve pas ce qu'elle veut.

Essaie d'améliorer sa RS, ouvre la RA pour affiner mais la ferme directement, cherche dans le pied de page, retourne dans les sélections et ouvre le blog.

Retourne dans la RS "recettes traditionnelles dijonnaises" mais n'a aucun résultat. Réduit sa recherche "recettes dijon". Retourne dans les sélections, "Écrits gourmands" et ouvre les entrées une à une, explore les listes de résultats et trouve "recettes provençales" et autres régions.

Essaie de passer au document à chercher avec titre + auteur dans la recherche simple "5 menus humoristiques Maurice Piat", ne fonctionne pas, elle ne trouve pas et passe au remplissage du guide d'exploration (arrête d'explorer pour remplir le guide).

Pour le guide, elle fait le tour de la PA, des thématiques, du blog, retourne sur "Écrits gourmands" et m'appelle "Eléonore, je ne trouve pas, comment je fais pour trouver ?", j'indique

qu'elle est au bon endroit, et qu'il faut chercher dans les entrées sous "Ecrits gourmands" notamment par région. Elle essaie mais ne trouve toujours pas.

Utilise son téléphone, elle commence à décrocher.

Elle réessaie de chercher avec les sélections de A à Z. Retrouve la cuisine provençale.

Elle termine de remplir le guide d'exploration. Fin du temps 1.

Guide d'exploration complété individuellement par les participants lors du temps 1.

Atelier sur la recherche dans Gallica



Atelier Université de Lille
Le 12/06/2024

Figure 12 : Guide d'exploration, atelier Jeunes Chercheurs

Liste de thèmes à explorer

	Thème	Difficulté	Document à chercher
Sport	Les Jeux Olympiques de Paris en 1924		5/7/1924, inauguration des Jeux olympiques au stade de Colombes [cérémonie d'ouverture] : [photographie de presse] / [Agence Rol]
	Les clubs de sport féminin au XXe siècle		La Femme sportive : organe officiel de la Fédération des sociétés féminines sportives de France / [gérant Maurice Boutet]
Littérature	Les poèmes d'Arthur Rimbaud		Une saison en enfer, 1873/ A. Rimbaud [Livre imprimé, 1ère édition]
	Mme de Sévigné et Mme de Grignan : la relation épistolaire mère/fille.		Contrat de mariage de François-Adhémar de Montell de Grignan avec Françoise-Marguerite de Sévigné. (28 janvier 1669). ET Lettres de Marie Rabutin-Chantal, marquise de Sévigné, à Mme la comtesse de Grignan, sa fille... Vol. I
Jeux	Nintendo		Collection Charles Cros. Consoles de jeux vidéo (média : multimédias). Consoles de jeux vidéo Nintendo - NES. Console de jeux vidéo Nintendo - NES, n° de série PE5192654
	Les cartes à jouer (ou jeux de cartes)		[Jeu de tarot parisien anonyme à enseignes italiennes] : [jeu de cartes, estampe] ([Original])
Presse	"J'accuse... !" (L'Aurore 1898)		Émile Zola. J'accuse !.., 11-13 janvier 1898. [Manuscrit]
	La Sorbonne dans la tourmente Mai 68		Le Mouvement social : bulletin trimestriel de l'Institut français d'histoire sociale
Musique	Boléro, Ravel		Bolero : [pour orchestre] / Maurice Ravel [Partition imprimée]
	Les archives de Bérurier Noir et du mouvement punk		Fonds Tomas Heuer (a.k.a mastO). Activité de photographe. Photographies pour Bérurier Noir. Séance photo du groupe pour la carte de vœux de l'année 1988, 1988
Cuisine	Les tartes sucrées et salées		Tartes sucrées et tartes salées / Sylvie Thiébault, 1977
	Les recettes traditionnelles Dijonnaises		5 menus humoristiques [Maurice Piat s.d.]
Nature	Le jardin du Luxembourg (Paris)		[Jardin du Luxembourg] : [dessin]
	Entretenir les arbres et arbustes fruitiers		Catalogue méthodique des arbres fruitiers, cultivés dans les jardins et pépinières de L. Noissette,...

Figure 13 : Guide d'exploration p.1, atelier Jeunes Chercheurs

Partie 1 : Votre profil

Connaissez-vous Gallica ? (entourez la réponse)

Oui Non

Avez-vous déjà utilisé Gallica ?

Oui Non

En trois mots, qui êtes-vous ?

Comment cherchez-vous de l'information habituellement ?

Quel sujet avez-vous choisi dans la liste proposée ?

Partie 2 : Vos attentes

Qu'attendez-vous d'une bibliothèque numérique ?

Que pensez-vous pouvoir y trouver (type de collections, outils de recherche et de visualisation, ...) ?

Qu'est-ce qui vous intéresse le plus dans une bibliothèque numérique ?

- Rechercher des documents
- Explorer les collections
- Découvrir / me laisser porter
- Lire
- Constituer un corpus
- Consulter les contenus de médiation
- Autre, précisez :

Arrêtez-vous ici, vous pourrez poursuivre le guide d'exploration une fois que vous aurez commencé à naviguer dans Gallica.

Figure 14 : Guide d'exploration p.2, atelier Jeunes Chercheurs

Partie 3 : Vos ressentis

Complétez le tableau suivant (1 étoile étant mauvais et 4 étant bon)

	Convivial/ esthétique	Agréable/ ergonomique	Compréhensible	Pertinent
La page d'accueil	☆ ☆ ☆ ☆	☆ ☆ ☆ ☆	☆ ☆ ☆ ☆	☆ ☆ ☆ ☆
La barre de recherche	☆ ☆ ☆ ☆	☆ ☆ ☆ ☆	☆ ☆ ☆ ☆	☆ ☆ ☆ ☆
La recherche avancée	☆ ☆ ☆ ☆	☆ ☆ ☆ ☆	☆ ☆ ☆ ☆	☆ ☆ ☆ ☆
La liste de résultats	☆ ☆ ☆ ☆	☆ ☆ ☆ ☆	☆ ☆ ☆ ☆	☆ ☆ ☆ ☆
Le visualiseur de documents	☆ ☆ ☆ ☆	☆ ☆ ☆ ☆	☆ ☆ ☆ ☆	☆ ☆ ☆ ☆
Les pages sélection (blog...)	☆ ☆ ☆ ☆	☆ ☆ ☆ ☆	☆ ☆ ☆ ☆	☆ ☆ ☆ ☆
Gallica en général	☆ ☆ ☆ ☆	☆ ☆ ☆ ☆	☆ ☆ ☆ ☆	☆ ☆ ☆ ☆

Quels mots décrivent le mieux Gallica ? (Entourez 3 mots maximum)

Vous pouvez ajouter vos propres mots

Un laboratoire Une ruche Un phare Un bazar Une forêt
 Un outil Un tunnel Un écosystème
 Une galaxie Un océan Un couteau-suisse Un labyrinthe
 Une mosaïque Un trésor Un réservoir

Gallica c'est plutôt ... (Entourez 3 mots maximum)

Pour les curieux.ses Pour les chercheur.e.s Pour tou.te.s
 Pour rechercher Pour découvrir Pour lire Pour se cultiver

Figure 15 : Guide d'exploration p.3, atelier Jeunes Chercheurs

Partie 4 : Votre bilan

Qu'est-ce qui vous plaît dans Gallica ?

Y a-t-il des choses que vous ne comprenez pas ou qui vous questionnent ?

Qu'est-ce qui vous manque dans Gallica (fonctionnalités de recherche, d'affichage, de paramétrage, ...) ?

Avez vous des idées pour améliorer Gallica et ses modalités de recherche ?

Figure 16 : Guide d'exploration p.4, atelier Jeunes Chercheurs

Partie 2 : Témoignages

Eléonore : Le but dans ce temps deux, c'est donc de savoir est-ce que c'était une expérience agréable ? Est-ce que vous avez réussi à vous dépatouiller avec Gallica ? Est-ce que vous avez réussi à trouver le document qui vous était demandé ? Est-ce que vous avez quinze documents pertinents ? Si vous avez réussi, quelles étaient vos stratégies de recherche ? Par quelle fonctionnalité êtes-vous passé ? Qu'est-ce qui était utile ? Qu'est ce qui ne l'était pas ? Et si vous n'avez pas réussi, ça nous intéresse aussi pour savoir quels sont les problèmes rencontrés. Est ce qu'il y a des stratégies qui fonctionnent mieux que d'autres ? D'autres qui au contraire ne marchent pas ? Et si vous n'avez pas trouvé à quoi c'est dû ? Voilà, des retours vraiment très généreux pour l'instant, on pourra affiner après.

N. K. : Mon thème c'était les tartes salées et sucrées et donc j'ai commencé par écrire "*tartes salées et sucrées*" sur le moteur de recherche. Et du coup j'ai commencé à naviguer dans les résultats pour voir si je ne trouvais pas par exemple le document qui était demandé et je ne le trouvais pas. Du coup, j'ai essayé de sélectionner un peu les documents.

Eléonore : Et tu as cherché comment ?

N. K. : Juste en tapant dans le moteur de recherche d'abord et ensuite j'ai été dans les thèmes et j'ai sélectionné vie domestique par rapport à la cuisine. Donc du coup, je regardais, il y avait un petit peu tout, mais je me suis dit qu'il fallait que je trouve le document. J'ai cherché le document donc j'ai cherché par le titre, ça ne fonctionnait pas du coup je ne le trouvais pas. Donc j'ai noté l'auteur...

Eléonore : Toujours dans la recherche simple ?

N. K. : C'est ça. Du coup, j'ai regardé parmi les documents qu'il avait, puis j'ai fini par le trouver. Euh. En scrollant quand même, je pense. Et puis voilà. Après, j'ai essayé d'utiliser un peu la recherche avancée, mais pas plus que ça et voilà. C'était mon expérience.

Eléonore : Ok. Et donc qu'est ce qui a été plus facile ? Qu'est ce qui a le mieux fonctionné ?

N. K. : Ce qui a le mieux fonctionné dans le fait d'avoir sélectionné le thème déjà. Et puis la précision aussi, parce que quand j'écrivais "*tarte sucrée et salée*" c'était un peu trop. Donc quand je modifiais, je mettais juste "*tartes*", je trouvais beaucoup plus de documents. Ensuite, j'écrivais tartes sucrés puis tartes salées.

Eléonore : Et est-ce qu'en tapant juste “tarte” tu avais que des documents pertinents dans ta liste de résultats ?

N. K. : Heu non...

Eléonore : Et comment tu fais pour choisir si c'est un document pertinent ou pas ?

N. K. : Bah je regardais. Moi je me suis basée sur les recettes, je cherchais des recettes. Donc des fois il n'y avait pas de recette, c'était des livres et puis il y avait le mot tarte qui apparaissait mais ce n'était pas ce que je recherchais.

Eléonore : Et globalement, c'était plutôt facile ou difficile ?

N. K. : C'était plutôt facile parce que c'était un thème précis je pense.

Eléonore : Ok. Est-ce que cela fait écho à l'expérience de quelqu'un ? J. B., tu veux réagir ?

J. B. : Ben moi, c'est l'inverse.

Eléonore : Ok, c'est bien aussi !

J. B. : Je pense que j'ai eu de la chance, j'ai utilisé juste la recherche simple, j'ai noté juste “Boléro, Ravel” et j'ai trouvé le document demandé en premier. C'est le premier qui est apparu et j'ai trouvé, par contre, j'ai trouvé quinze documents mais ils n'étaient pas tous pour le coup pertinents. Du coup j'ai juste sélectionné moi et “Boléro” et “Ravel” dans la description sinon je ne prenais pas. Et après ? Du coup c'est vraiment en lisant la description, parce que des fois ça parle d'autres Ravel, d'autres...

Eléonore : Tu as utilisé la recherche simple d'une façon, on va dire “unique” parmi les 4 là parce que tu as tapé “Ravel, Boléro” ou “Boléro, Ravel”, pourquoi tu as tapé ça ?

J. B. : Parce que c'était noté sur la feuille comme ça.

Eléonore : Ok, donc tu as juste recopié le titre. [Rire de J. B.]

J. B. : Oui !

Eléonore : Ça c'était une fonctionnalité intéressante de mettre en rapport l'auteur et l'œuvre. Et donc tu as dit que c'était plutôt facile. Est-ce qu'il y a des fonctionnalités qui étaient plus faciles à utiliser que d'autres ou qui ont plus aidé ?

J. B. : Du coup, pour le coup je n'ai pas

Eléonore : Tu as vraiment utilisé que la recherche simple ?

J. B. : Ouais.

Eléonore : Et pareil pour la pertinence, comment est-ce que tu juges si un document est pertinent ou pas ?

J. B. : Du coup, comme j'utilisais seulement deux mots-clés, j'attendais que ces deux mots-clés là apparaissent dans les documents recherchés et en fonction, en fait, ça a fait la sélection tout seul.

Eléonore : Et est-ce que tu as eu beaucoup de bruit ?

J. B. : Pas mal oui.

Eléonore : Et comment tu fais pour réduire le bruit ?

J. B. : J'aurais dû mettre des guillemets ?

Eléonore : Par exemple ? Mais tu n'as pas eu besoin de mettre de filtre.

J. B. : Non parce que j'ai trouvé assez facilement, en fait un document sur deux était à peu près pertinent. Du coup, comme c'était quinze documents à trouver, c'était plutôt rapide.

Eléonore : Ok, très bien. Est-ce que quelqu'un va réagir ? J. V. tu acquiesces ?

J. V. : J'ai fait un peu pareil que les filles, j'ai tapé les mots de mon sujet. J'ai trouvé surtout beaucoup d'images du coup...

Eléonore : Oui, mais qu'est-ce que tu cherchais ?

J. V. : Je ne savais pas trop ce qu'on entendait par "document important". Du coup, du coup, j'ai essayé d'explorer les différents types de documents. J'en ai gardé un peu de chaque en fait. Il y avait pas mal de choses du coup, je suis tombée sur les bons documents en ligne direct au début.

Eléonore : Et qu'est-ce que tu as utilisé comme fonctionnalité ? Juste la recherche avancée ? La recherche simple ? Tu es passée par les sélections ?

J. V. : J'ai utilisé des filtres à côté sur les types de documents, j'ai essayé les dates mais je ne me suis pas trop arrêtée dessus finalement. Et du coup, enfin, j'ai lu les descriptions pour bien m'assurer que ça correspondait bien à mon sujet. Ce n'est pas parce que des fois c'était écrit

“Luxembourg” mais ça ne correspondait pas au jardin. C'est tout. Mais après je ne suis pas sûre que j'ai vraiment trouvé les documents les plus pertinents ?

Eléonore : Mais tu as trouvé 15 documents qui étaient plus ou moins en rapport avec le jardin du Luxembourg, c'est ça ?

J. V. : C'est ça.

Eléonore : Et donc c'est une expérience qui était plutôt simple ?

J. V. : Oui.

Eléonore : Est-ce que ça ressemble à ce que tu utilises habituellement comme moteur de recherche ?

J. V. : Bah à part pour les cours, franchement ?

Eléonore : Ok.

Eric : C'est les facettes que tu as utilisées.

J. V. : Oui, c'est ça.

Eléonore : Oui, une recherche simple avec l'utilisation de facettes pour affiner.

J. V. : Oui, c'est ça.

Eléonore : Ok, une expérience réussie du coup ?

J. V. : Je ne sais pas ?

Eléonore : On n'est pas tous obligés d'avoir des expériences réussies, on peut aussi avoir des déceptions. [Réactions] Voilà. Il n'y a pas de problème, de toute façon pour être honnête et sincère, (et c'est aussi ce qu'on disait en atelier avec les collègues), le but n'est pas de dire que le moteur fonctionne, le but c'est aussi de savoir ce qui ne fonctionne pas. Donc allez-y sur ce qui fonctionne mais n'hésitez pas à nous dire aussi ce qui ne fonctionne pas parce que ça nous intéresse aussi et entre guillemets, surtout... Oui, K. G., tu voulais dire ?

K. G. : Moi je me suis un peu ... Ce n'était pas facile d'avoir la main sur le moteur de recherche. Déjà, j'ai vu le thème (“J'accuse... !”), je croyais que ça allait être plus facile. Je mets “J'accuse... !”, j'ai aussitôt l'article ou bien le truc que je cherchais. Dès que j'ai vu en bas, c'était écrit “15 résultats par page”, j'ai dit “Bingo! J'ai trouvé mes quinze articles”. Mais non, ce n'était pas vraiment ça parce que dans les 15 articles, c'était juste le mot que j'ai inscrit dans

le moteur de recherche qui paraissait dans tous les documents qui étaient joints dans la page et c'était pas vraiment le thème que je cherchais. Déjà, le titre même du thème laissait, ça ne peut pas reporter à plusieurs choses déjà, donc on ne peut pas tomber sur des articles intéressants sur le coup, il faut les ouvrir et en les ouvrant, il y a un problème de visualisation. Apparemment, ce sont des documents scannés et les caractères sont trop petits, c'est peut-être parce que je suis myope ? Je ne sais pas mais...

Eléonore : Ce n'était pas océrisé ?

K. G. : Oui enfin je ne sais pas, ce n'était pas facile de le lire pour savoir si c'est pertinent ou pas. Il y avait ça comme difficulté que j'ai trouvée du coup je n'ai pas trouvé mes quinze articles qui pouvaient être intéressants sur le thème "J'accuse... !" de la presse et j'ai compris que c'était juste... je ne sais pas si c'est un journal ou un magazine ? Et il est parti sur les écrits d'Emile Zola pour utiliser peut-être le style qui a été utilisé dans "J'accuse... !" pour peut-être aborder une thématique quelconque et ce n'est pas donné à tout le monde donc c'est normal que je ne trouve pas tout ce...

Eléonore : Est-ce que tu connaissais bien le thème "J'accuse... !" ? Emile Zola ? L'affaire Dreyfus ? Etc...

K. G. : Oui !

Eléonore : Donc tu étais suffisamment armée pour savoir un petit peu comment chercher, mais c'est l'interface Gallica qui fait que tu n'as pas réussi à trouver ?

K. G. : Tout à fait.

Eléonore : Alors la presse c'est vrai est particulièrement difficile. Les soucis de visualisation dont tu parles sont assez communs. On a déjà eu ce genre de retours. Par exemple J. B. ou est-ce que tu connaissais bien le Boléro de Ravel ou est-ce que tu es partie... ?

J. B. : Non, pas spécialement.

Eléonore : D'accord, tu as eu un coup de chance en tapant le titre ?

J. B. : Oui.

Eléonore : N. H., tu avais une expérience similaire un peu complexe ?

N. H. : Oui, je n'ai pas vraiment trouvé, mais j'ai pas trouvé quelque chose par rapport à mon sujet (les recettes traditionnelles dijonnaises). Oui, après j'ai sélectionné par thématiques, j'ai

choisi “arts, loisirs et sport”, après j’ai pris “patrimoine gourmand”. Et après j’ai choisi l’“écrit gourmand”, j’ai rien trouvé et j’ai opté par “métiers et savoir-faire”, mais j’ai pas trouvé aussi quelque chose par rapport à mon...

Eléonore : Oui, donc tu as utilisé les sélections. Est-ce que tu as utilisé la recherche simple ?

N. H. : Hum oui mais je n’ai pas trouvé aussi. Je pense que c’est moi le problème.

Eléonore : [en rigolant] Normalement, ce n’est pas l’usager le problème. [Rires]

N. H. : Oui.

Eléonore : Et est-ce que tu as utilisé la recherche avancée ?

N. H. : Non.

Eléonore : D’accord. La recherche avancée, j’ai vu que tu as regardé pour le guide d’exploration...

N. H. : Oui.

Eléonore : Pourquoi tu ne l’as pas utilisée par exemple ?

N. H. : Parce que je ne connais pas comment elle fonctionne ? Comment faire exactement ?

Eléonore : Trop complexe ?

N. H. : Voilà. Elle était claire et tout mais je ne comprenais pas comment faire et tout.

Eléonore : Est-ce que tu as l’habitude d’utiliser les fonctionnalités de recherche avancée sur d’autres moteurs ? Par exemple sur Google ?

N. H. : Oui.

Eléonore : Qu’est-ce que tu utilises comme filtres habituellement ?

N. H. : Bah je ne sais pas, juste en bas. Bon, après il y en a d’autres mais je ne sais pas comment dire. Sinon sur Lillocat ...

Eléonore : Sur Lillocat, tu utilises la recherche avancée ?

N. H. : Oui.

Eléonore : Et ça te donne des résultats pertinents ?

N. H. : Pour moi.

Eléonore : Et là, c'était plutôt que tu n'avais pas de résultats ou qu'ils n'étaient pas pertinents ?

N. H. : Non, je n'avais pas de résultats et je ne connais exactement comment j'utilise, comment je filtre et tout.

Eléonore : Donc tu n'as pas réussi à constituer deux corpus, ni à trouver le document demandé ?

N. H. : Oui.

Eléonore : Mais tu as trouvé quelques documents en rapport ?

N. H. : Oui.

Eléonore : Est-ce que vous voulez réagir ?

A. T.P. : Il faut reconnaître que j'ai eu aussi un peu de chance. Après le sujet de ma recherche c'était qu'un seul mot du coup c'était plus facile pour moi de tomber sur -a priori- sur des documents que je cherchais.

Eléonore : Peut-être que tu peux rappeler ton thème pour tout le monde.

A. T.P. : Nintendo, tout simplement.

Eléonore : Et comment tu as fait pour trouver des documents en rapport avec ce thème-là ?

A. T.P. : Alors donc dans la barre de recherche, j'ai tapé "Nintendo" et le premier résultat que j'ai eu, c'était le document que je devais retrouver d'ailleurs. Par rapport aux 15 autres documents pertinents sur le sujet, j'en ai trouvé une dizaine. Donc dans la dizaine, il y avait six qui étaient des objets et donc ils m'avaient beaucoup plus intéressé parce que ça parlait de ces consoles, ça montrait différents objets autour de Nintendo et il y avait des livres qui mentionnaient Nintendo quelque part dans leur corpus. Je pense que c'était plus cela. Donc après, j'ai essayé de les chercher par thème, donc c'était par ci, par-là, je suis passé par les thématiques, je pensais que je rentrais dans Arts et Loisirs mais non. Donc j'ai dû reprendre par le type de documents et vu que j'avais trouvé les jeux dans l'objet, donc je les ai retrouvés dans Objets, collection Charles Cros. Et donc c'est par là que je suis retombé après, plus tard sur le document, quand on le cherche. En gros, c'est comme ça que j'ai fait.

Eléonore : Et pour trouver des documents qui n'étaient pas des objets justement, comment as-tu fait ?

A. T.P. : Par des filtres au niveau des facettes qui étaient présentées j'ai juste (parce que dans les types de documents), il avait montré et les livres quelques presses et revues et les objets. Donc j'avais sélectionné d'abord les objets pour être sûr que la plupart parlaient effectivement de Nintendo et ensuite j'ai exploré les livres.

Eléonore : Donc c'était plutôt facile ?

A. T.P. : C'était plutôt facile.

Eléonore : Et quelles étaient les difficultés, alors ?

A. T.P. : Les difficultés par rapport au bruit par exemple ? J'ai remarqué qu'il ne prenait pas strictement Nintendo. C'est-à-dire qu'à des endroits soit. Je ne veux pas dire qu'il écrivait mal, mais il proposait "Natendo" ... ?

Eléonore : Des expansion orthographiques ?

A. T.P. : Exactement ! Et personnellement j'aurais peut-être préféré... J'ai trouvé ça bien mais pas bien en même temps. Dans le cas où j'aurais pu moi-même me tromper dans l'orthographe, ça aurait pu m'aider. Mais c'est du coup, j'aurais apprécié que justement, dans les facettes qu'il propose d'autres mots qui se rapprochent un petit peu de celui que j'ai et que ce soit à moi de choisir...

Eléonore : La possibilité d'activer ou non les expansions ?

A. T.P. : Exactement ! Plutôt que de remplir tout Parce qu'il y avait énormément mais beaucoup de résultats avec des expansions.

Eléonore : Donc non pertinents ?

A. T.P. : Exactement, beaucoup de bruit sur des termes qui n'étaient pas du tout le mot-clé exact que j'avais en tête.

Eléonore : Est-ce que tu as réussi à filtrer un peu ? Justement, pour retirer le bruit ?

A. T.P. : Alors justement, pour retirer ces bruits, non. Les filtres que j'ai là, c'est plus des filtres de livres, de types de documents. Donc justement, je ne pouvais pas... Après, je n'ai pas exploré la recherche avancée.

Eric : Tu as essayé un peu non ?

A. T.P. : J'ai regardé mais je n'ai pas lancé, Je n'ai pas lancé quand je n'ai pas essayé. J'ai juste vu la page de la recherche avancée, mais je n'ai pas fait la demande dessus pour voir si dans la recherche avancée, il prend strictement le mot ou s'il ... J'aurais voulu essayer tout à l'heure mais je n'ai pas eu le temps. Je n'ai pas été jusqu'au bout.

Eléonore : Ce que j'ai vu de ce côté-là. Si vous avez exploré la recherche avancée, mais plus pour le guide d'observation, vous n'êtes pas allés naturellement vers la recherche avancée, surtout J. B. qui n'a presque pas utilisé je crois ? [J. B. acquiesce] N. K., je sais que t'as fait des tests. Ça s'est passé comment ?

N. K. : Il y avait beaucoup d'informations, c'était un peu complexe. Je pense que j'ai très bien fait, c'est que pour le coup c'est vraiment précis. Ça permet d'ajouter de la précision, après, ... il y a beaucoup de choses que je ne comprenais pas forcément, comme par proximité. Après, il y a une explication. Voilà beaucoup d'informations, c'est peut-être un peu impressionnant quand on clique dessus et qu'on voit tout ça, mais ça permet de chercher assez précisément.

Eléonore : Et est-ce que quelqu'un a eu une bonne expérience avec la recherche avancée ? Une recherche avancée qui a fonctionné ? Je vois non chez J. V., non chez J. B.. Personne n'a réussi à faire une recherche avancée qui fonctionne, qui permette de filtrer ?

N. K. : Je n'ai pas testé plus que ça...

Eléonore : Tu as testé un petit peu quand même. Mais de ce que j'ai vu, tu a eu soit pas de résultat du tout, soit des résultats mais qui n'étaient pas pertinents par rapport à ce que tu voulais. [N. K. acquiesce] Donc pas de bonnes expériences, non mais ce n'est pas grave. E. T., peut-être que tu veux nous expliquer un petit peu ton parcours ?

E. T. : Bon, moi j'y suis allée directement par la recherche par thématique et j'ai trouvé pas mal de documents, beaucoup plus des images, des images, j'en ai trouvé, mais par rapport au document à chercher, là je suis directement allée dans la barre de recherche...

Eléonore : La recherche simple ?

E. T. : Oui. Mais lorsque je l'avais pris, j'étais passé par là recherche par thématique, je n'avais pas trouvé le document à chercher. C'est dans la recherche simple que j'ai retrouvé. Bon, je n'ai pas aussi vraiment utilisé la recherche avancée mais j'ai regardé un peu par rapport aux livres. J'ai vu quatre livres. Sur le thème mais c'était qu'un seul... Le reste ne parlait pas...

Eléonore : C'était un seul livre, cad un seul document mais qui apparaissait plusieurs fois dans la liste de résultats ?

E. T. : Non. Il y avait quatre documents différents, mais y avait un seul qui parlait des Jeux Olympiques ?

Eléonore : D'accord, donc il y avait trois documents non pertinents.

E. T. : Voilà.

Eléonore : Et dans ce cas-là, comment est-ce que tu fais pour filtrer ? Est-ce que t'as réussi à filtrer ?

E. T. : Non.

Eléonore : D'accord, donc t'as eu des résultats qui étaient non pertinents avec beaucoup de bruit.

E. T. : Mais par rapport aux livres. Mais par rapport à la recherche que j'ai fait par thème, j'ai eu des documents. Lorsque j'ouvre, je vérifie un peu si c'était des "Jeux Olympiques Paris 1924".

Eléonore : Et qu'est-ce qui n'a pas fonctionné du coup ?

E. T. : Moi je voulais avoir des documents textes. Et c'était que des images. Pour avoir des informations, etc, mais il n'y avait que des images.

Eléonore : Et dans ce cas-là, qu'est-ce que tu as mis en place pour essayer de trouver autre chose que du texte ? Qu'est-ce que tu as utilisé ?

E. T. : Moi je n'ai rien utilisé...

Eléonore : Mais en revanche, tu as constaté que ta liste n'était que des images.

E. T. : Oui, bien sûr.

Eléonore : mais tu aurais aimé avoir d'autres types de documents ?

E. T. : Voilà !

Eléonore : Effectivement, peut-être qu'il n'y avait pas de texte sur les Jeux Olympiques, c'est possible. Bon toutes ces questions, elles nous amènent quand même à questionner cette notion de "bruit" et de "pertinence". Et c'est vrai que vous en parlez quand même beaucoup de bruit

dans la liste de résultats. Finalement, comment est-ce qu'on fait pour choisir si un document est pertinent ou non ? Qu'est-ce que vous avez utilisé comme critère pour savoir celui-là j'ai envie de le mettre dans mon corpus, celui-là me plaît ?

J. V. : Moi, c'est vraiment quand les termes de mon sujet ressortaient dans la première page, dans le titre ou les sous-titres.

Eléonore : Et pour toi, jardin tu l'as mis au pluriel et au singulier ?

J. V. : Je n'ai pas fait attention mais je pense que j'ai mis que du singulier.

Eléonore : Et donc si on t'avait proposé "parc du Luxembourg", ça n'aurait pas été pertinent ?

J. V. : Je vérifiais aussi que c'était bien en France aussi, à Paris.

Eléonore : Et comment tu vérifiais ?

J. V. : Je lisais la description, il y avait Paris.

Eléonore : Et donc tu ouvrais le document ?

J. V. : Ouais, je regardais un peu la première page au moins.

Eléonore : Le tri ! [Rires] Ça nécessite quand même beaucoup de temps. Tu as fait ça sur combien de documents ?

J. V. : En fait, pas tant parce qu'il y avait beaucoup de photos et d'images avant, mais informations c'était très facile de les voir.

Eléonore : Dans la liste de résultats tu veux dire ? C'est facile de voir les infos dans la liste ?

J. V. : Quand ça s'affiche oui, enfin dans la liste de résultats je veux dire, c'est facile de les voir dans la liste. Mais même quand ça s'affiche, oui enfin dans les résultats, même quand j'ouvre les fichiers en fait. Souvent quand c'est des images, c'est facile aussi. Pour les textes, les textes après pareil, c'était plutôt facile, je regardais quand même la première page.

Eléonore : Et donc pour trier, ta liste de résultats, tu es obligée d'ouvrir le document ?

J. V. : Je les ai ouverts à chaque fois que je ne me suis pas contenté de me dire... je les ai tous ouverts en me disant, tiens c'est intéressant. Je regardais à chaque fois que ça correspondait.

Eléonore : Donc tu as dû avoir beaucoup d'onglets.

J. V. : Je ne sais pas en fait.

Eric : Tu utilises les favoris toi beaucoup ?

J. V. : Oui, je les ai mis en favoris en fait.

Eléonore : d'accord, donc tu as constitué ton corpus avec beaucoup d'onglets et en plus tu as repris les références ?

J. V. : C'est ça.

Eric : Oui mais après.

Eléonore : Ok pour faire du tri, c'est intéressant. De ce côté là aussi, vous avez ouvert beaucoup d'onglets, surtout vous deux [N. K. et J. B.]. Beaucoup d'onglets, genre 8 ou 10. Comment vous faites pour déterminer ce qui est pertinent ou pas ? Qu'est-ce qui motive votre choix ? Qu'est-ce qui fait que vous ouvrez un nouvel onglet (et le gardez) ?

J. B. : Moi pour le coup, c'était juste... Très vite, j'ai fait une première sélection avec les termes que j'ai mis dans la barre de recherche, si ces termes ne sont pas là, je ne vais pas plus loin.

Eléonore : Non pertinent direct ?

J. B. : Non pertinent.

Eléonore : Donc Boléro tout seul ou Ravel tout seul c'est non pertinent ?

J. B. : C'est ça ! Parce qu'il y a beaucoup, enfin Boléro c'est beaucoup de résultats.

Eléonore : Tu avais quoi d'autre par exemple ?

J. B. : Il y a beaucoup de Ravel, je ne savais pas. Boléro, ça me parlait de vêtements... ? Donc je me suis concentrée essentiellement sur Ravel. Les deux. Je regardais le type de document.

Eléonore : Tu cherchais quoi comme type de document ?

J. B. : Je ne cherchais pas, un type en particulier, c'est juste que forcément, on avait de la musique, c'était telle ou telle symphonie. Je savais que c'était pertinent pour ce que je recherchais. Pour le coup, il y a eu des documents où il y avait Boléro et Ravel. En fait, c'était une gouache et en fait c'était juste la mention de la gouache, mais on ne la voyait pas. Donc j'ai considéré que ce n'était pas pertinent. Mais j'ai dû ouvrir le document pour voir qu'en fait il y avait juste le nom.

Eléonore : Et du coup les onglets que tu as gardés, c'était ton corpus on va dire ?

J. B. : Oui Je n'avais même pas vu qu'il y avait un truc "favoris".

Eléonore : Oui, bien "favoris" en fait, tu l'as sur tous les navigateurs. Et N. H., pour la pertinence, du coup, comment t'as fait ? Puisque tu t'es retrouvée avec beaucoup de bruit.

N. H. : Oui. J'essaye de filtrer mais...

Eléonore : Et donc toi tu cherchais "Recettes", "cuisine" ? Quels étaient les termes en fait que tu avais en tête ?

N. H. : Voilà, cuisine, recettes. J'ai essayé avec le "menu humoristique", mais je n'ai pas vraiment...

Eléonore : Du coup, tu cherchais plusieurs termes. Et est-ce que tu aurais aimé par exemple que le moteur te propose des termes en rapport avec ton sujet ?

N. H. : Ah oui !

Eléonore : Tu cherches "menu", il te propose "recettes", tu cherches "recettes" il te propose "repas" ?

N. H. : Oui ! Par exemple. Pourquoi pas, ça aide à trouver quelque chose quand même.

Eléonore : Oui, parce que là, du coup, tu t'es retrouvée un peu démunie face à ta liste de résultats.

N. H. : Oui...

Eléonore : Et du coup la pertinence, ça aurait été donc tous ces mots-clés là et Dijon ?

N. H. : Oui.

Eléonore : Seulement ?

N. H. : Oui.

Eléonore : Et il n'y a rien d'autre qui aurait pu être pertinent ? Par exemple le nom d'une spécialité dijonnaise ?

N. H. : oui, c'est vrai, je n'ai pas pensé à ça... [rire]

Eléonore : Par exemple.

N. H. : Oui, c'est vrai, parce que je ne connais pas très bien.

Eléonore : On n'est pas tous spécialistes de la cuisine dijonnaise [rires].

N. H. : Je ne connais pas, mais c'est vrai que c'est une bonne idée.

Eléonore : Souvent, si jamais tu es coincé, tu peux chercher avec d'autres mots-clés et ça c'est vrai que c'est quelque chose que vous n'avez pas du tout fait. C'était marrant d'ailleurs, en tout cas pas de ce côté-là, personne n'a cherché par Google. Vous êtes restés dans Gallica. Alors vous avez été très sages [rires] et vous avez fait vos recherches dans Gallica, mais beaucoup d'utilisateurs, voyant que la recherche ne fonctionne pas dans Gallica, vont sur Google tapent "recette dijonnaise Gallica" ou "recette dijonnaise BnF". Et ensuite, dans la liste de résultats de Google et ils arrivent directement dans le visualiseur de documents sans passer par le moteur de recherche. Peut-être que tu peux essayer de voir ce que ça donne. Tu nous dis dans cinq minutes ce que tu as réussi à trouver ? [Rires]

N. H. : Oui, d'accord.

Eléonore : Qui d'autre ? Donc peut-être que pour "J'accuse... !", ça marcherait aussi ? Peut-être avec "J'accuse... !" entre guillemets avec BnF ou Gallica, ça donnerait des résultats qui sont plus pertinents en tout cas d'autres résultats.

A. T.P. : Oui mais dans ce cas-là, pourquoi passer par la BnF ?

Eléonore : C'est-à-dire ?

A. T.P. : Pourquoi sur le coup, pourquoi est-ce que ce serait pertinent de passer par la BnF ? J'y pensais pendant que je travaillais. Après, comme je disais, je ne connaissais pas déjà très bien Gallica. Donc pour moi, j'ai eu l'impression, en explorant qu'il fallait plus ou moins savoir ce qu'on voulait avant de l'explorer parce que justement, si on cherche un objet, on tombe sur l'objet, on a toutes les facettes de l'objet, tous les détails sur l'objet. Si on cherche un document particulier dont on veut connaître le contenu, c'est plus facile. Mais par contre, si on ne sait pas exactement ce qu'on veut mais qu'on a tout simplement une idée générale, qu'on part sur : "Je veux faire un sujet sur la Nintendo en particulier", sans forcément parler de l'objet. Mais je voudrais bien savoir ce que Internet me propose sur le sujet, à mon sens, c'est un peu plus difficile avec Gallica. Comme je le dis, je ne connais pas forcément Gallica. D'ailleurs je l'avais noté dans mes trucs, ça aurait été bien d'au départ qu'il y ait un petit paragraphe qui nous dise en quelques lignes : "Ben Bienvenue Gallica, ici vous aurez un ensemble de documents sous différents formats sur ..." Je ne sais pas ou "L'objectif de Gallica c'est de faire tatata". Et comme ça ce n'est pas nous dire quelques mots, en quelques lignes l'objectif du site. Et ça nous

rassure un peu plus sur ce qu'on devrait chercher et comment on devrait chercher. Et dans la pertinence des résultats, on saurait ce qu'il faudrait ou pas considérer parce que l'outil est conçu de sorte que... je pense que c'est là que je me suis rendu compte qu'il y avait plus de documents historiques plutôt que de documents actualisés. Et ça, je ne savais pas vraiment, c'est une critique, je trouvais qu'il y avait trop de... Voilà, j'aime bien que ce soit très visuel, mais qu'il y a trop d'images, de beaucoup d'images du Moyen-Âge et de cette époque-là plutôt que des choses récentes.

Eléonore : E. T. avait l'air d'être d'accord avec ça ? Peut-être, qu'est-ce que tu aurais à ajouter ? Est-ce que t'as vécu la même chose de ne pas savoir dans quoi tu cherches d'être un peu perdue ?

E. T. : Oui pour moi, il y avait beaucoup d'informations puisque sur les Jeux Olympiques il y avait beaucoup d'informations, mais il fallait que je clique sur chaque résultat pour pouvoir voir le contenu. Donc je clique et je vois les athlètes sur le terrain et tout ça et l'image qu'on me propose et là je choisis ça comme étant pertinent. Il y a des résultats qui ne rentrent pas dans le cadre des JO.

Eléonore : Et donc pareil, tu as été obligée d'ouvrir le visualiseur de documents pour savoir si c'était pertinent pour ta recherche mais tu ne vois pas depuis la liste de résultats ?

E. T. : Exactement. Il y a des images dans la liste de résultats avec des titres mais il faut, puisque j'avais sur l'une d'entre elles, je suis restée plus prudente pour vérifier au fur et à mesure.

Eléonore : Ok et dans la présentation initiale de Gallica, qu'est-ce que vous pourriez attendre ?

A. T.P. : Je n'ai pas compris ?

Eléonore : Dès qu'on arrive sur le site, tu as dit que tu voulais "Bienvenue sur Gallica". Qu'est-ce que tu voudrais ensuite ? La typologie de documents ? Le nombre de documents ? La période ?

A. T.P. : Dans ma tête, je me dis que -je sais plus où j'ai déjà vu ça-, mais que quand on sait plus ou moins ce qu'on veut... parce qu'en explorant, moi j'ai l'impression que c'est vraiment idéal quand on sait plus ou moins ce que ça veut. Savoir à quoi ressemblait une Nintendo en 1968, en 1970 ? On arrive, je cherche un objet : "que recherchez-vous ?" -> par type de document, ... Je vois ça quand même un petit peu là, mais peut-être plus (comment dire) plus... que ce soit le visuel, plus impactant, plus affiché.

Eléonore : Et donc là il y a écrit sur la page d'accueil le nombre de documents qui sont les documents numérisés.

A. T.P. : Oui, 10 millions.

Eléonore : Oui, et 10 millions c'est quoi, c'est beaucoup ou pas beaucoup ? C'est énorme ? C'est peu ? On ne sait pas ?

J. V. : Moi je pense que c'est bien ?

Eléonore : C'est bien, d'accord.

J. V. : Ça dépend à quoi on compare après mais c'est déjà pas mal.

Eléonore : Ok, c'est bien mais est-ce que c'est une idée du nombre de documents à la BnF ?

J. V. : Non, je ne sais pas.

Eléonore : Gallica c'est 10 millions de documents numérisés mais ils ne disent pas que c'est environ 10 % des collections de la BnF. Alors tout le monde n'est pas d'accord sur le comptage etc. Mais, il y aurait un peu plus de 100 millions de documents conservés à la BnF et "seulement 10 millions" sur Gallica. En fait, quand on voit le bruit qu'on a dans 10 millions de documents et quand on sait que tous les ans on numérise de nouveaux documents et que donc on étend la collection Gallica. En 2010 (Arnaud est passé rapidement dessus) mais en 2010, c'était 1 million de documents. Le moteur de recherche a été créé vers 2010, mis en production vers 2013. En dix ans, on est passés de 1 million de documents, maintenant à 10 millions de documents. Et donc finalement, est ce que le moteur de recherche est toujours adapté ? C'est la grande question qu'on pose au travers de cette enquête, c'est est-ce qu'on a besoin de changer la recherche dans Gallica ? J. B., je vois que tu dis oui ?

J. B. : Bah, au niveau du bruit, qu'on a tous eu je pense dans nos recherches, il faut affiner. Il y a choses qu'il faut enlever, il y a des choses qui sont pertinentes. Moi, encore une fois, j'ai eu de la chance, j'ai trouvé ce que je voulais sans trop de souci même le bruit ne m'a pas posé tant de problème que ça.

Eléonore : Mais, tu avais déjà un thème très spécifique.

J. B. : Bah voilà mais effectivement, si en plus c'est pour appâter entre guillemets de nouveaux usagers, comme le disait A. T.P.. C'est compliqué si on n'a pas de thème précis, enfin jamais de la vie je vais aller sur Gallica pour chercher quelque chose.

Eléonore : Et tu as vu le blog, les sélections etc, là tu peux y aller sans avoir de thème précis, et te laisse découvrir, enfin normalement les sélections sont là pour te donner à découvrir la collection. Est-ce que ça fonctionne ?

J. B. : Non. Enfin si je suis honnête non.

Eléonore : C'est bien, il faut être honnête, soyons honnêtes.

J. B. : Il faudrait connaître Gallica et le blog. Honnêtement, j'ai déjà utilisé Gallica, j'ai jamais consulté le blog. Jamais. Parce que je ne sais même pas ce que c'est en vrai. J'y vais spécifiquement pour faire mes recherches et comme je sais que c'est quelque chose où il faut déjà avoir son idée et c'est juste un outil de recherche et pas pour explorer, je ne vais pas sur Gallica, surtout pas pour découvrir de nouvelles choses et me laisser porter par...

Eléonore : Donc Gallica, c'est fait pour une recherche précise ?

J. B. : De mon point de vue, ouais.

J. V. : Je pense que la recherche spécialisée, c'est surtout pour des documents particuliers, anciens, des collections d'images, ... Enfin, on sait qu'on peut trouver des documents de bonne qualité en fait. C'est plus ça, je pense. Mais sinon, à part ça, je pense qu'en général...

Eléonore : Tu ne te verrais pas le dimanche matin, te connecter sur Gallica pour aller lire... ?
[Rires]

J. V. : Non [rire].

A. T.P. : Et j'ai remarqué aussi. Même justement pour éliminer un peu le bruit dans les propositions similaires quand (justement j'étais dans les objets) sur la Nintendo dans les propositions similaires, il proposait des auteurs ? Sauf que déjà en vérifiant l'auteur sur l'objet que j'ai trouvé qui était pertinent pour moi. Je ne sais pas vraiment si c'est le même auteur quand on propose un document dans les auteurs, alors que c'est écrit auteur similaire. Mais justement, moi j'aurais voulu que dans les propositions qui se rapprocheraient justement, il y ait des propositions dans les auteurs, mais les propositions utiles. C'est vrai que c'est fort de dire des propositions dans le type de documents parce que c'était ici des objets et si on m'avait proposé des documents similaires sans proposer des objets, on va dans le meilleur des documents, mais je pense qu'on aurait pu... parce que l'auteur similaire, je n'ai aucune idée... ?

Eléonore : C'est des auteurs similaires mais qui n'ont pas de lien avec la recherche ?

A. T.P. : Aucun lien. D'ailleurs, de similaire... je ne sais pas ce qu'ils ont de similaire avec... [rires]. Parce que là, autant on a mis Nintendo &co parce que c'est compagnie etc ? Du coup, et même dans "auteurs similaires", je ne vois pas le nom de l'auteur tout de suite, mais dans la visualisation (déjà que je trouve très bien, c'est très bien franchement, cette...)

Eléonore : La visualisation de quoi ?

A. T.P. : Des documents. C'est-à-dire ce qui sort quand on a le résultat. Je trouve ça très bien, ça nous permet avant d'avoir le résultat d'avoir une idée de l'avant-veille, mais je ne sais pas ce qu'il a l'auteur de similaire de l'autre en fait, du coup aucune correspondance donc fatalement ça m'aide pas vraiment à aller vers et à approfondir.

Eléonore : En fait, tes retours tournent beaucoup autour de l'explicabilité. Tu aimerais qu'on t'explique un peu plus le résultat du fonctionnement du moteur, dans quelle interface tu te situes, la collection. Ce que tu disais tout à l'heure, c'est que tu aimerais avoir plus d'explications...

A. T.P. : Plus d'explications oui ! C'est déjà ce que je dis, plus d'explications en général oui, mais même là, dans la proposition de documents similaires, quel que soit le format qu'il y ait, qu'on trouve une certaine cohérence avec. Parce que comme je le dis là, je n'arrive pas à comprendre, il n'y a pas de lien, j'ai fait le test sur une autre recherche et dans les documents similaires, il n'y avait strictement aucun rapport.

Eléonore : Oui, on a eu des exemples de personnes qui cherchaient "citrons" à qui le moteur a proposé "Citröen", qui cherchaient les "typologies" et on eu "typographies" ...[rires]. On en a eu plein. Parce qu'en fait, il y a des expansions orthographiques. Il y en a quatre, dont il y a par lemme, par homonyme, etc. Et en fait, elles sont activées d'office. Donc quand vous allez chercher « Citron », vous allez aussi avoir « Citroën ». Parce qu'il y a une lettre d'écart, et le moteur se dit "s'il a fait une faute d'orthographe, je lui donne aussi Citroën". Et donc plus les mots sont longs, plus le nombre de caractères qui peut varier est grand. On peut aller jusqu'à quatre caractères qui varient. Donc typologie/typographie, ça veut dire que pour tout ce qui est typo, etc, il va proposer un nombre de documents énorme. Et en fait, ça, ça a été activé en 2013 parce qu'il y avait des erreurs d'OCR dans les documents, parce que c'était surtout des textes et là, vous avez fait beaucoup de recherches sur des images, etc. L'OCR ne se pose pas. Et en fait, une des questions qu'Arnaud avait posé c'est : est-ce que ces expansions qui sont activées automatiquement, il ne faudrait pas les désactiver et proposer à l'utilisateur d'activer l'extension dont il a besoin ?

J. B. : Bah si ! C'est même évident parce que sur un million de documents c'est encore faisable vu le nombre de thématiques qu'il y a, c'est faisable d'activer. Mais là, c'est à coup de 1 million de documents ajoutés en plus chaque année, bah forcément, il y a de plus en plus de documents qui reviennent en faisant nos recherches. Et donc c'est de moins en moins pertinent. C'est à ce moment-là que c'est à l'utilisateur d'avoir la possibilité de mettre ça en place.

Eléonore : D'activer ou de désactiver les extensions, mais qu'elles soient visibles, c'est vrai que là, elles sont visibles nulle part. Vous pouvez les chercher, vous ne les trouverez pas. C'est activé dans le back office. Sinon, le désactiver et laisser libre cours à la personne qui recherche, K. G., tu voulais dire ?

K. G. : Il faudrait les désactiver et laisser libre court à la personne qui recherche ce qu'il veut pour retrouver facilement. Et surtout que même avec tout ce qui est déjà intégré. Quand tu cliques dessus, en fait, on te dit que l'écriture ou l'orthographe n'est pas ... Ce qui est important c'est de l'afficher dans le moteur de recherche. Comme quoi c'est un truc qui est déjà dans les bases de données.

Eléonore : Orthographe exacte, recherche exacte, etc.

K. G. : Oui, si on clique sur le document que le moteur de recherche te propose, une page s'ouvre pour te dire que là, pour cette orthographe, on a sélectionné ce qui correspond à la recherche proposée. [J. V. acquiesce]

Eléonore : J. V. ?

J. V. : Ah non non.

Eléonore : Tu es d'accord ?

J. V. : Oui !

Eléonore : Ok, tu es d'accord avec et tu n'as rien à ajouter. La question un peu sous-jacente, c'est est-ce qu'il faudrait donner à l'utilisateur la main pour changer plus de paramètres pour paramétrer Gallica en fonction de ses besoins ? Parce que là on vous a donné des thèmes assez différents. Il y avait la presse, il y avait du son, je n'ai pas mis de vidéo, mais j'aurais pu puisqu'il y en a, je n'ai pas mis manuscrit mais on aurait pu. En fait, la typologie documentaire (il y a des objets) est vraiment assez vaste et diversifiée. Et donc finalement, est-ce que quelqu'un qui ne recherche que des images, on va lui proposer du plein texte qui ne correspond pas à la recherche sur les images ? Et donc est-ce que vous aimeriez par exemple que dans le compte

utilisateur, le compte personnel, vous puissiez dire “Moi, je fais plutôt ce genre de recherches” que le système vous propose d'activer certains paramètres par défaut qui sont possibles à changer après en paramétrages, et que donc, à chaque fois que vous vous connectez sur Gallica, vous avez ce Gallica personnalisé proposé d'office ?

K. G. : Après la difficulté qu'on a à chercher sur le moteur de recherche c'est qu'on n'a pas facilement une prise en main sur les filtres, mais là c'est comme un casse-tête chinois. Tu cherches, tu avances, tu ne trouves pas. Pourtant, sur d'autres moteurs de recherche, tu insères une typologie d'informations (images, textes, etc) et hop, tu vas là tu trouves, tu ne trouves pas, tu fermes et en 5 minutes tu trouves ce que tu veux ! Mais là, tu as l'impression que tu fais une course, un jeu de piste où tu ne trouves jamais. Par exemple, surfer sur ce moteur de recherche. Ce n'est pas forcément. Et ça fait que ça ne donne pas envie de revenir dessus, je ne me vois pas le temps de faire une recherche très sérieuse dessus [rire]. Pendant une journée pour trouver qu'un seul document, franchement moi je ne m'y mettrai pas. [Rires]

Eléonore : Oui, ça se comprend. C'est vrai qu'en plus le tien était particulièrement difficile à trouver. Donc ton expérience est partagée par d'autres, y compris des gens qui sont bibliothécaire à la BnF et qui n'utilisent pas Gallica. Non mais ça pose un cadre, c'est certain. Mais donc quand tu dis “éparpillés”, peut être que tu aimerais non pas qu'il y ait une recherche simple et une recherche avancée, des filtres, avec la possibilité de rajouter des guillemets, des and, des not, des « or » dans la recherche ?

K. G. : Bah des filtres, comme sur tous les autres moteurs de recherche surtout qu'on voit que là, il y a une certaine crédibilité. Les documents sont précieux et c'est crédible en fait, quand on doit le mettre dans un travail de recherche, dans nos recherches, on sait que quand le document vient de Gallica, c'est un truc authentique, pas comme tous les autres moteurs de recherche qui peut avoir des trucs piochés par ci par là. Ça serait intéressant que la personne qui cherche puisse y accéder facilement et pouvoir l'utiliser tout le temps.

Eléonore : C'est la première fois que j'ai ce retour là en atelier, cette authenticité ou en tout cas ce côté patrimonial des documents. Mais pour l'instant, les gens n'en ont pas parlé et finalement, en quoi on devrait les chercher différemment ? Qu'est-ce que tu voudrais pour signaler ce côté patrimonial des documents ?

K. G. : Ouf, ça je ne sais pas, je ne me suis pas préparée à ça. [Rires]

Eléonore : Ce n'est pas grave. Ce n'est pas grave même des retours blancs, ça va aussi ? Bon, on va prendre une petite pause si ça vous va pour revenir en forme pour le temps trois.

Partie 3 : Améliorations

Eléonore : Nous t'écoutons.

N. K. : Par rapport à la charte graphique de Gallica, c'est très gris, très maussade. Et du coup, on s'est dit qu'il faudrait peut-être améliorer le jeu de couleurs.

Eléonore : Oui, pour mettre plusieurs couleurs ? Une couleur plus vive ?

N. K. : Voilà !

J. B. : Pas forcément plus vive mais plus de contraste.

Eléonore : Pour mieux différencier ce qui est cliquable et ce qui ne l'est pas ?

J. B. : Voilà. En fait, on a une ligne en haut et on ne la voit pas.

K. G. : La barre d'outils, on ne comprend pas, ce n'est pas différencié, juste une phrase pour savoir sur la barre d'outils. Il faut faire attention pour dire que ce sont les différentes catégories ou sections de recherche quoi.

N. K. : Oui et des couleurs plus attractives.

[Approuvé par les autres.]

Eléonore : Ok, ensuite ?

N. K. : Ensuite, au niveau du thème. Moi quand je cherchais et que je cherchais par thème du coup bah par exemple j'ai cherché la cuisine, mais du coup il y a eu le thème de "vie domestique" donc j'allais sur ça mais je n'avais pas de sous-thème "cuisine" ou des choses plus précises. Donc peut-être préciser avec des sous-thèmes spécifiques.

Eléonore : Tu voudrais plus des thèmes ? Des thèmes plus généraux ? Mettre des entrées qui partent un peu plus (parce que peut-être que "vie domestique" ça ne parle pas réellement aujourd'hui ?). Qu'est-ce tu voudrais exactement ?

K. G. : C'est parce que c'est très général, on voudrait que ce soit plus spécifique. On garde "vie domestique" et dedans on met "cuisine".

Eléonore : Mais du coup on aurait plus de clics avant d'arriver au document, là il y en a déjà beaucoup ? Beaucoup de clics quand on passe par les sélections, N. H., tu l'as beaucoup expérimenté pour la cuisine dijonnaise. [Rires]

N. H. : Ah oui !

Eléonore : Donc tu as énormément de clics d'affilés pour arriver à trouver un document (pertinent ou pas ce n'est même pas la question).

N. H. : Oui !

Eléonore : Donc vous voulez encore rajouter des échelles ?

N. K. : Dans mon expérience, juste quand je regardais sur le côté, quand je cliquais sur "vie domestique", j'aurais trouvé ça simple d'avoir en-dessous "cuisine", ...

Eléonore : Ménage, vaisselle, ... ?

N. K. : Oui !

Eléonore : Ok.

J. B. : Après on a mis "amélioration du référencement" ? En fait, on ne sait pas trop dire comment appeler ça, mais moi du coup j'ai mis "Boléro" et "Ravel", je mets les deux mots. En fait, des fois j'avais des résultats où il y avait qu'un seul des deux mots-clés dans les plus pertinents avant d'autres documents où j'avais les deux mots et qui étaient du coup pour le coup plus pertinents. Mais qui arrivaient en dessous...

Eléonore : Donc changer l'ordonnancement de la liste de résultats pour un autre ordre ?

J. B. : Oui !

Eléonore : Et donc comment tu classes l'ordre ?

J. B. : En fonction des recherches, moi j'ai deux mots, je veux que les deux ressortent.

Eléonore : Donc les deux mots apparaissent au début, ensuite tu as les résultats avec un des deux mots ?

J. B. : C'est ça en fait, le plus intéressant ce serait par exemple ce qu'on a dans les titres. Pour mettre effectivement les documents avec un titre qui contient un des deux mots en premier c'est pertinent mais que par la suite, quand on est dans la description, on ait les deux mots recherchés qui viennent en premier quoi.

[...]

Eléonore : Les résultats sont déjà classés par pertinence sauf qu'on ne sait pas ce que c'est la pertinence pour le moteur de recherche. [...] Il y a beaucoup de critères qui ne sont pas visibles pour l'utilisateur et qui expliquent pourquoi on a telle ou telle liste de résultats.

N. K. : Oui mais ce n'est pas bien pour l'utilisateur.

Eléonore : Vous avez fini vos 3 améliorations ?

J. B. : Oui, après on en avait deux petites. C'est juste que (alors pour le coup je n'ai pas fait de recherche avancée donc je ne sais pas comment ça fonctionne), mais juste sur le volet de gauche, on peut choisir le type du document...

Eléonore : Les filtres ? Les facettes ?

J. B. : Sauf qu'on peut en choisir qu'un, c'est chiant.

Eléonore : Normalement, on peut les cumuler.

J. B. : Ah bah je n'ai pas réussi. Sur la facette juste à gauche je n'ai pas réussi.

N. K. : Et du coup, moi ce que j'avais trouvé aussi, c'est que par exemple quand on scrolle au niveau des résultats, on finit par trouver un document, je clique dessus. Et puis après, quand je retourne en arrière, je retourne tout en haut des résultats donc j'ai perdu ma recherche et du temps.

Eléonore : Ok, merci pour vos améliorations. [Au groupe 2] Est-ce que vous voulez partager votre liste ?

E. T. : Oui, nous on trouve que la page d'accueil est trop chargée : beaucoup d'images, de textes et tout.

Eléonore : Donc, qu'est-ce qu'il faut changer ? Moins d'éléments ?

E. T. : Oui.

Eléonore : Moins d'images plus de textes ?

E. T. : Non.

Eléonore : Moins de choses ?

E. T. : Oui, alléger.

A. T.P. : Je pense qu'il y a beaucoup, il y a énormément d'informations.

Eléonore : Oui et elle est très très très longue.

A. T.P. : Oui très longue et je pense qu'on pourrait regrouper les informations.

Eléonore : J'ai compté, il y a 9 scroll si on compte le footer. [Rires] C'est beaucoup !

A. T.P. : Oui et énormément de trucs qu'on pourrait regrouper dans un menu. Mais je pense que ça Et on rejoint aussi le fait que la barre de recherche aurait pu être plus visible. Je me disais même qu'il y a des bases de recherche sur certains sites et qui défilent quand on scroll.

Eléonore : Qui restent ... ?

A. T.P. : Oui, qui restent affichées même quand on scroll.

Eléonore : Ok donc une barre de recherche plus visible, si on doit alléger vous enlevez quoi ?

A. T.P. : Sur la page d'accueil ?

Eléonore : Oui. Parce qu'il y a des tutoriels et des outils, des entrées par type de document, des entrées par la médiation (blog etc), il y a des entrées par publics (les classes, les professeurs, ...), pour ...

A. T.P. : Franchement, moi j'aurais enlevé tout ça. Et comme je dis, j'aurais proposé peut-être un menu qui permette de rentrer sur une page où il y a ça. Mais sur la première page, je trouve que... c'est limite violent.

Eléonore : La page d'accueil, elle date de 2017. [...] Mais elle est longue pour arriver jusqu'au footer. Je sais que N. H. tu as vu le footer, je t'ai vu l'utiliser [rires] mais je ne sais pas si les autres, vous l'avez vu, vous avez regardé dedans ?

A. T.P. : Non. Mais c'est un mélange et c'est là que j'ai vu qu'il y avait un "à propos" parce que je ne l'avais pas vu dès le début et c'est lui qui explique, qui parle un peu de Gallica et qui parle de comment est-ce qu'il marche.

Eléonore : Donc la page d'accueil de Gallica doit présenter Gallica, c'est ça ?

A. T.P. : Oui, en fait, j'en avais parlé plus tôt, quel qu'il y ait ne serait-ce que trois lignes, mais qui explique un peu tout dans un petit résumé de Gallica. Je pense que ça nous permet de savoir un peu...

E. T. : Que ce soit bien visible !

A. T.P. : Ça nous permet de savoir un peu dans quoi on met les pieds.

Eléonore : Ok.

E. T. : Et puis nous avons trouvé aussi que la recherche avancée est trop complexe.

Eléonore : Oui. Et donc là, qu'est-ce qu'on fait ? On retire des champs ? On ... ?

A. T.P. : Oui, je pense qu'il faut retirer des champs. Je ne sais pas si d'autres ont trouvé des champs pertinents ? La proximité par exemple ? Je ne sais pas ? L'année d'édition encore que peut-être quand on sait ce qu'on cherche ? Mais a priori quand on ne sait pas vraiment ce qu'on cherche...

Eléonore : Effectivement, N. K., tu as utilisé non la date d'édition ?

N. K. : Oui.

Eléonore : Et ça a marché ?

N. K. : Oui. Enfin, oui j'ai trouvé, mais après... pour mon ... Non en fait ça n'a pas marché...
[rires]

Eléonore : Tu sais, tu n'as pas besoin de me dire que ça a marché pour me faire plaisir. Ce ne me fait pas plus plaisir que ça ait marché ou que ça n'ait pas marché ! [Rires]

N. K. : Enfin ça a marché mais pas pour des résultats que je considérais comme étant non pertinents.

Eléonore : Oui, effectivement. Donc des filtres qui sont plus ou moins pertinents en fonction des usagers.

A. T.P. : Oui, c'est ça.

Eléonore : Ok, ensuite ?

A. T.P. : Et bon, une dernière chose peut-être, enfin deux petites choses. Moi j'aurais pensé à une sorte de carte mentale. J'en parlais un petit peu plus tôt, c'est-à-dire quand je sais plus ou moins ce que je veux, je peux peut-être partir de là. Je recherche, qu'est-ce que vous recherchez ? Un objet : un objet dans quel thème ? C'est-à-dire qu'à chaque fois que je pars d'un format ou d'une autre typologie de documents, je pars d'un format, quand je clique sur l'objet, par exemple, il va me demander dans quel thème se trouve cet objet et quand je clique dessus quel sous-

thème ? Et je pense que ça peut affiner un petit peu plus. Si ça se présente un peu comme une carte mentale.

Eléonore : Et ce serait où ? Dans la liste de résultats ?

A. T.P. : Je pense non, je pense que ça peut même être sur la page d'accueil.

Eléonore : Ok et comment on fait pour l'activer ? Enfin, pour rechercher quelque chose dedans ?

A. T.P. : Ce que je vois, ce que j'imagine, c'est quelque chose qui dit tout simplement que recherchez-vous? Et quand on clique dessus, ça peut le montrer comme je dis d'abord les différents types de documents et les différents formats de documents (si on cherche un objet, si on recherche un livre, si on recherche une vidéo). Et ensuite quand on choisit le type de document, on peut avoir une liste de thèmes, si on recherche un objet, est-ce qu'il est dans la littérature ? Enfin proposer tous les thèmes ou les sous-thèmes qu'on a à disposition jusqu'à ce qu'on tombe un petit peu plus sur un ensemble de documents répondant un petit peu à tout ça.

Eléonore : Donc il y a deux choses, que recherchez-vous ? On propose des types de documents, etc. Et à partir de ce qui est tapé, proposez des thèmes en lien ?

A. T.P. : C'est ça !

Eléonore : Et donc là, je tape J'accuse, on me propose Emile Zola, on me propose l'Aurore, on me propose d'aller explorer ...

A. T.P. : Exactement oui !

Eric : il faudrait commencer par thème et ensuite arriver sur les types de documents.

A. T.P. : Oui, c'est pour ça que j'ai dit que ça dépend. C'est vrai que c'est pour ça que je disais que ça dépend beaucoup de si on sait ce qu'on recherche. Pour moi, je trouve que c'est plus facile, enfin je préfère commencer par le format si j'ai en tête le format, si je cherche un livre.

Eléonore : Une enluminure...

A. T.P. : Ah oui, exactement que je passe par là et je pense que ça réduit beaucoup plus quand je passe par là avant de chercher le thème. Pour moi, je pense que ça m'amène vers des résultats plus précis que de passer par le thème avant de descendre vers ce type de documents.

K. G. [s'en va] : Merci beaucoup, je dois partir mais c'était une belle expérience, découvrir Gallica et en même temps contribuer à son amélioration, si jamais ça arrive ? Et après, je crois

qu'on est peut-être très exigeants dans le sens où c'est une bibliothèque et dans toutes les bibliothèques, on ne trouve pas tout ce qu'on veut, mais c'est peut-être aussi ça. Merci !

Eléonore : Merci à toi pour ta participation ! Donc ok, très bons retours. On ne nous a pas proposé de carte mentale jusqu'à maintenant mais on nous a proposé des datavisualisations avec des cercles en disant "Je cherche Emile Zola", pour l'instant il dit juste (imaginons qu'on a 10 000 résultats), il dirait "dans Gallica, il y a 8 000 résultats, dans la Bibliothèque de Paris, il y en a 150, ..." et il ferait des cercles comme ça pour dire ce qu'il y a. Ensuite il pourrait dire ce qui est océrisé, dans quels documents on peut faire de la recherche plein texte etc. Donc la carte mentale, ça complète très bien la datavisualisation qu'on nous a proposée, ce serait une autre approche mais effectivement très complémentaire.

A. T.P. : Ah, après la deuxième chose, je ne sais même pas si c'est possible d'avoir une messagerie un peu comme sur Lillocat.

Eric : Sur l'Ent, pour échanger avec des documentalistes ?

A. T.P. : C'est ça. Au départ je parlais sur un chatbot, mais je me dis que c'est la même chose que la carte mentale du coup, parce que en partant d'un mot-clé, il va faire pareil comme sur le moteur de recherche. Mais messagerie pourquoi pas et a la possibilité de discuter avec un documentaliste...

Eléonore : Instantanément ?

A. T.P. : Instantanément si c'est possible. Parce que je pense oui et certainement si c'est possible ou si on laisse un mail peut-être pour le lire plus tard. S'il a du mal à retrouver la référence.

Eléonore : Alors là, il y a déjà deux choses. Il y a l'assistance Gallica avec un collègue qui répond aux mails et aux questions des Gallicanautes, c'est dans le footer [rires] Oui, c'est tout en bas de la page d'accueil, il faut déjà être motivés pour y arriver. Il y a déjà quelqu'un qui fait ça pour Gallica. Et pour la BnF plus généralement, il y a un service de chat pour discuter avec les documentalistes, c'est Sindbad. On peut y poser une question, demander : je cherche des manuscrits illustrés en couleur. Impossible qu'un usager lambda arrive à faire cette recherche-là parce qu'arriver à trouver des manuscrits en couleur, c'est d'une grande complexité. Donc ils vont sur Sindbad, ils posent la question et là, il y a un conservateur du département des manuscrits qui va aller sur Gallica [...] et essayer de répondre à l'utilisateur avec la meilleure requête et donc ils peuvent donner à l'utilisateur les requêtes à taper directement dans Gallica.

Donc en fait, il y a déjà les deux, juste, peut-être que ce n'est pas suffisamment mis en avant ? Mais oui, vous pouvez avoir des conservateurs qui vous répondent et qui vous aident.

J. B. : Oui mais ça c'est dommage que ce soit dans le footer et on ne le voit pas, enfin l'utilisateur ne le voit pas parce qu'il descend jusque-là.

Eléonore : Sindbad n'est même pas indiqué sur Gallica, il faut aller sur les sites de la BnF.

J. B. : Non mais même sur Gallica, du coup j'imagine qu'il répond toujours aux mails des mêmes personnes parce qu'il y a 3 personnes qui sont au courant ? [Rires]

Eric : Et en même temps, c'est peut-être un choix hein. Comme ils n'ont pas les moyens, ils ne veulent pas le mettre en avant ?

Eléonore : C'est plus que la page d'accueil n'est pas rationnelle, faire 9 scroll avant d'arriver en bas etc, comme le reste, ça a été finetuné pendant des années, aujourd'hui, on hérite de ça et ça va être complètement remis à plat. Il y a une collègue qui s'occupe de la page d'accueil, j'ai vu les premières ébauches, la barre de recherche va être énorme, il va y avoir un burger menu au lieu d'avoir les entrées (un menu qui se déroule pour avoir les entrées), 3 scroll pour arriver jusqu'en bas, ils ont retiré toutes les entrées blog etc. Il reste une entrée blog, une entrée moteur de recherche, une entrée actualités de la BnF et la dernière m'échappe. Mais il y a quatre entrées sur la page d'accueil. Là maintenant, la question c'est dans quel ordre mettre ces blocs-là ? Ce sera l'objet d'un atelier en juillet, où on va demander aux gens de les classer dans l'ordre dans lequel ils souhaiteraient que les blocs apparaissent. Nous on pense que la barre de recherche devrait être en haut.

J. B. : Ah bah oui, ce serait mieux !

Eléonore : Mais peut-être que les usagers vont trouver que, d'abord, il faudrait le blog ou d'abord il faudrait autre chose. On laisse la possibilité et on est à l'écoute de ce que les gens ont à nous dire. Et oui, la refonte de la page d'accueil, c'est un gros sujet.

J. V. : Du coup, je ne sais pas si ça a déjà été dit, mais quand on tape un mot dans la barre de recherche, on trouve que ça ne donne pas des choses cohérentes en fait. Par exemple, on avait mis "recettes" et ça nous donnait "budget des recettes" enfin ça continuait, il ne donnait pas juste "recettes" et quelque chose en fait. Ce n'est pas optimal. Après on se dit parfois peut être qu'on ne tape pas le bon mot dans la barre de recherche et du coup ça ne trouve pas les documents pertinents. Par exemple, si vous donnez des synonymes un peu à côté ou quoi, peut-être qu'on aurait plus de documents du coup. Et on a aussi trouvé la barre de recherche avancée

trop compliquée, je pense que c'est lié à toutes les fonctionnalités... ou ce n'est pas assez bien décrit peut-être ? Je pense. On ne sait pas forcément bien ce que ça veut dire en fait ?

Eléonore : Ok et est-ce que vous avez ouvert des bulles d'aide contextuelle ? Je sais que N. K. l'a fait... beaucoup ! [Rires]

J. V. : Je crois ?

Eléonore : En fait les bulles d'aide contextuelle c'est pour chaque fonctionnalité, il y a un petit texte qui décrit comment c'est censé être utilisé.

J. V. : ... ? Heu... ouais ?

N. H. : Après, on peut ajouter là les types de documents sur la recherche avancée, genre...

J. V. : Ça c'est assez simple à utiliser, c'est bien par rapport au reste qui est plus compliqué je pense.

Eléonore : Mais si on cherche sur "le ciel". Je voudrais faire une recherche iconographique sur "le ciel" et donc je veux des estampes, des photos, ... Et donc, comment je fais une recherche iconographique sur le ciel ? Ou le printemps si on n'entre que par type de document ? Images c'est vaste ?

J. B. : Faut proposer d'abord l'entrée par thème.

N. H. : Il faut d'autres thématiques.

Eléonore : Ok donc plusieurs entrées. Mais c'est une vraie question, laquelle on met en premier ? Parce qu'il y a des collègues qui cherchent "le ciel" [rires] pour les réseaux sociaux etc. Alors ils cherchent le ciel, le printemps, enfin ce genre de choses, enfin des choses assez vastes. Et donc dans la barre de recherche avancée, vous proposez plusieurs entrées ?

N. H. : Voilà.

A. T.P. : Il y a je pense les types de documents.

Eléonore : Oui. [...] Ok, et du coup vous avez donné vos trois améliorations ? Des synonymes, une recherche avancée trop complexe, c'est quoi le 3e ?

Eric : Ce n'est pas les synonymes sur la recherche, c'est que les mots-clés proposés sont pas bons : "recettes" et "bilans des recettes" ?

J. V. : Oui, ce n'est pas logique en fait.

Eric : Oui, il y a des propositions dans le moteur de recherche, il y a un nom pour ça et qui sont pour le coup trop filtrés et qui sont présélectionnés. Parce qu'à mon avis, dans le corpus, il y a beaucoup plus de bilans des recettes que de recettes de cuisine. Et je pense que c'est pour ça que tu as ce mot qui est remonté. Ce n'est pas ce que toi [Eléonore] tu disais. A force de vouloir être intelligent, et ben on fait l'inverse dans le moteur de recherche.

J. V. : C'est compliqué.

Eric : Parfois c'est contre-productif. Ce n'est pas tant la synonymie.

J. V. : Après la synonymie c'était bien parce que parfois on n'a pas d'idées de ce qu'on cherche.

Eléonore : Oui, c'est deux propositions ?

A. T.P. : Oui.

J. V. : Oui.

Eléonore : Ok, on a plein de retours, on va passer à la conclusion avant de se laisser.

Le questionnaire biblionautes

Le questionnaire est disponible en ligne en suivant ce lien :

<https://app.questionnaireweb.com/f/154067/4pebxva2GJXrvqw9hq8sRx>.

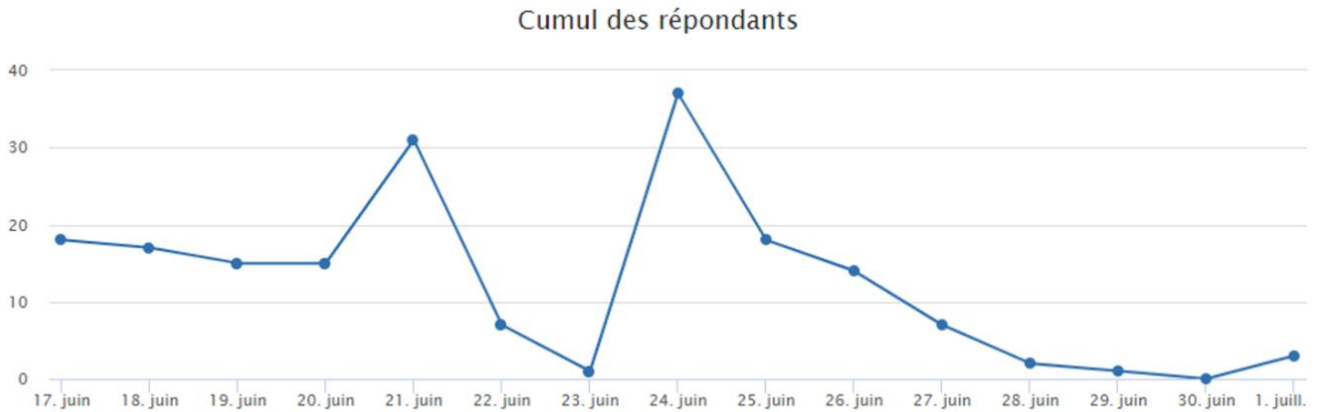


Figure 17 : Cumul des répondants au cours de la période de diffusion de l'enquête

Extraction des réponses au questionnaire biblionautes

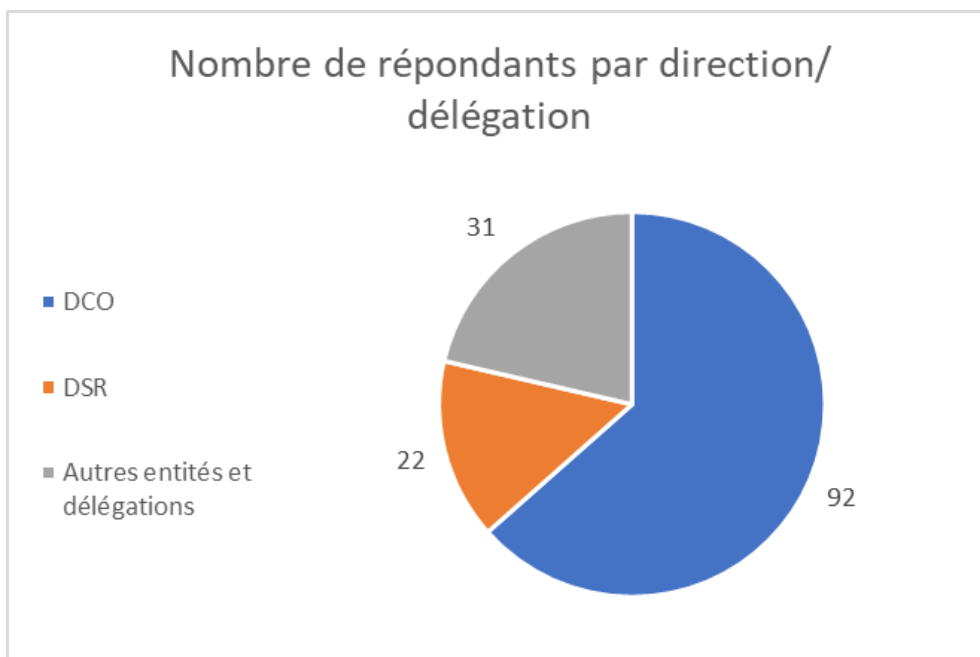


Figure 18 : Nombre de répondants par direction/ délégation

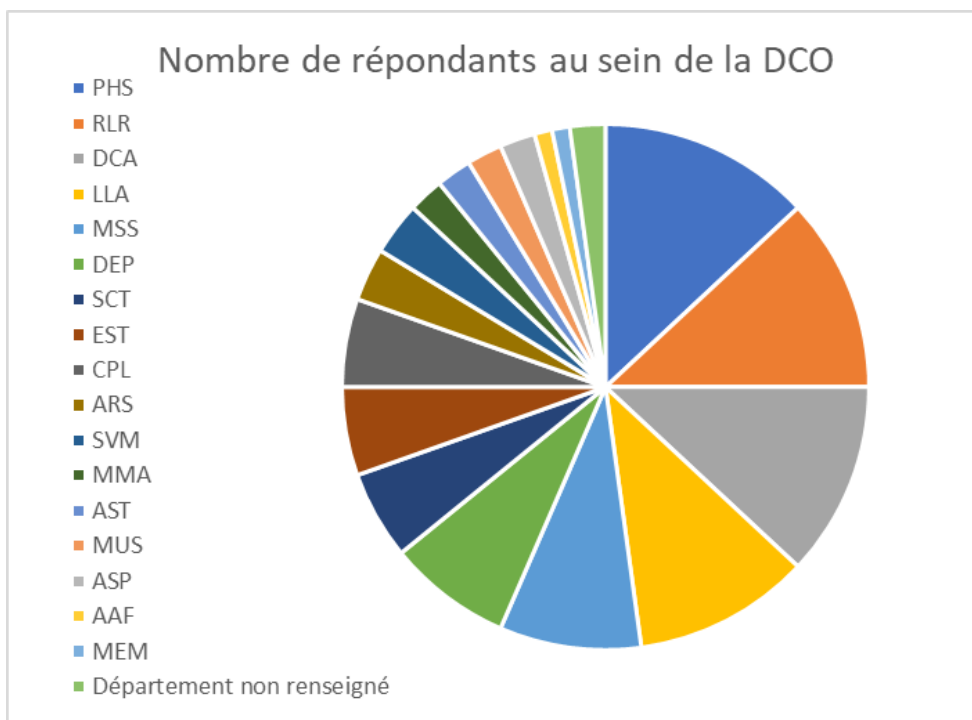


Figure 19 : Nombre de répondants au sein de la DCO

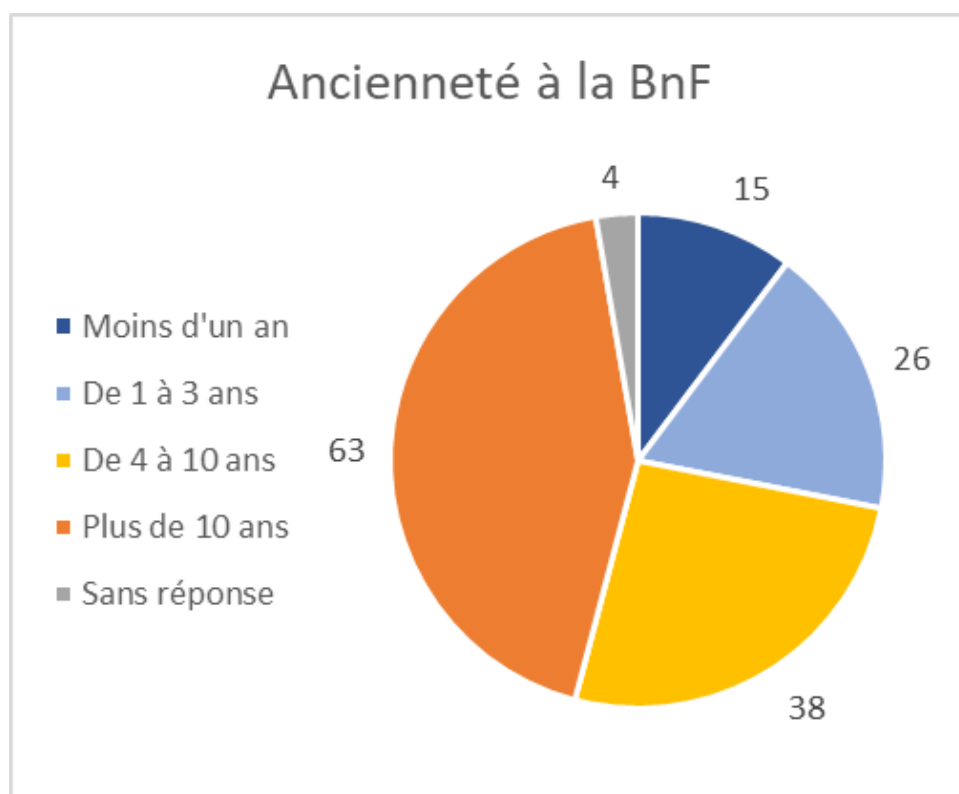


Figure 20 : Ancienneté à la BnF

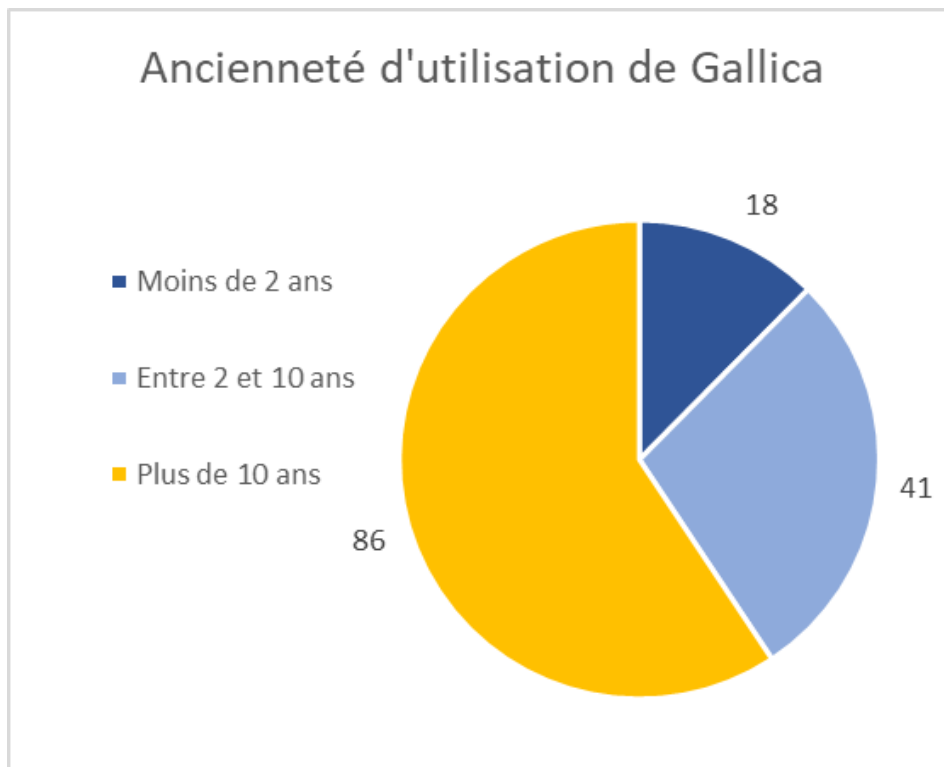


Figure 21 : Ancienneté d'utilisation de Gallica

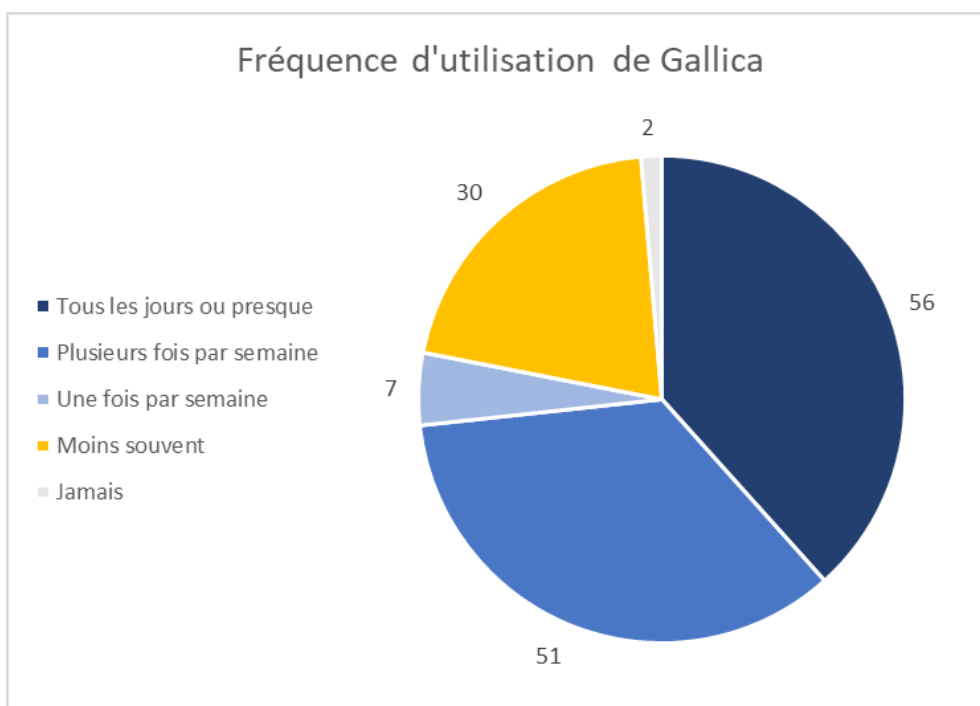


Figure 22 : Fréquence d'utilisation de Gallica

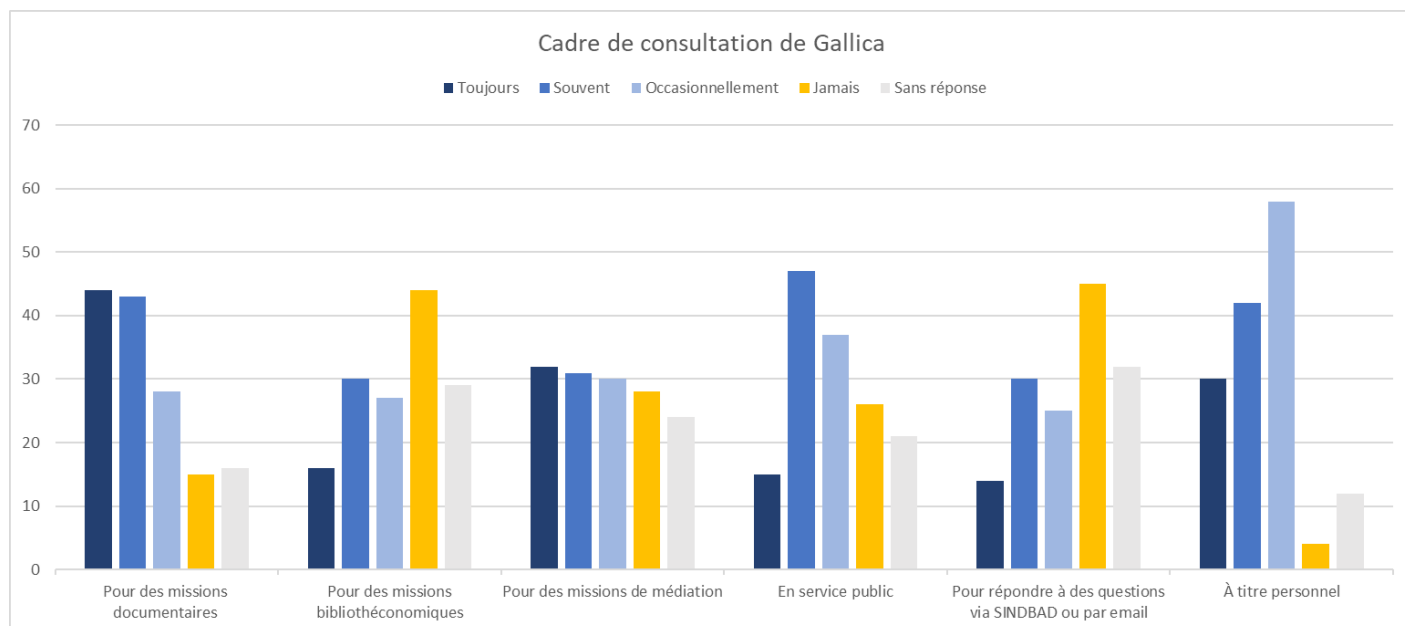


Figure 23 : Cadre de consultation de Gallica

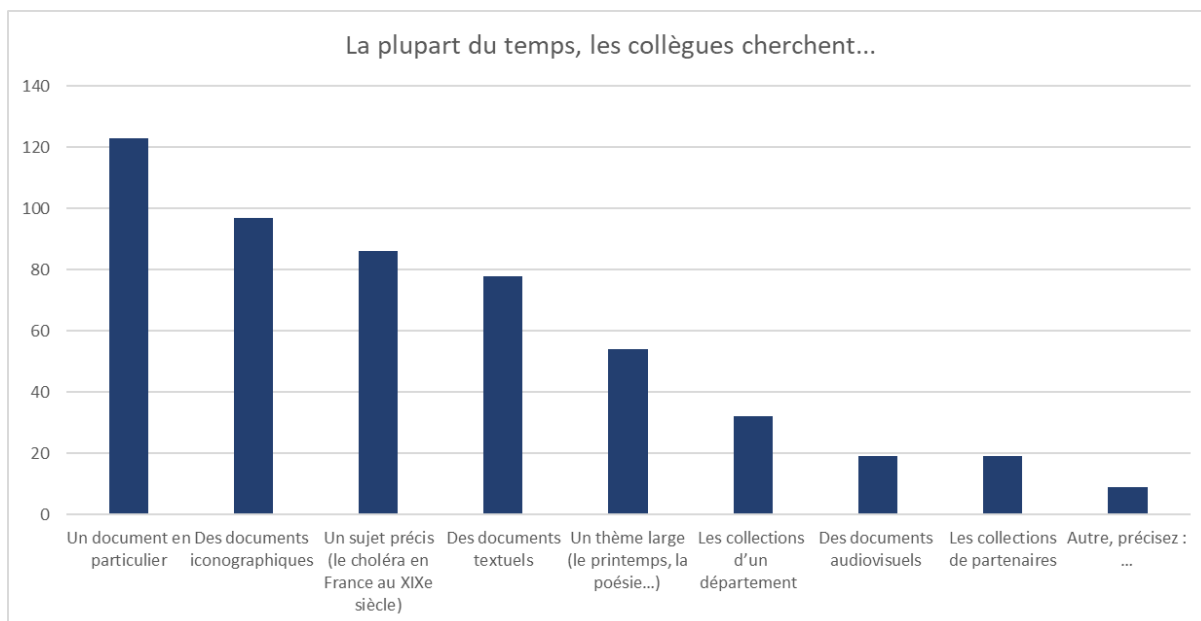


Figure 24 : La plupart du temps, les collègues cherchent...

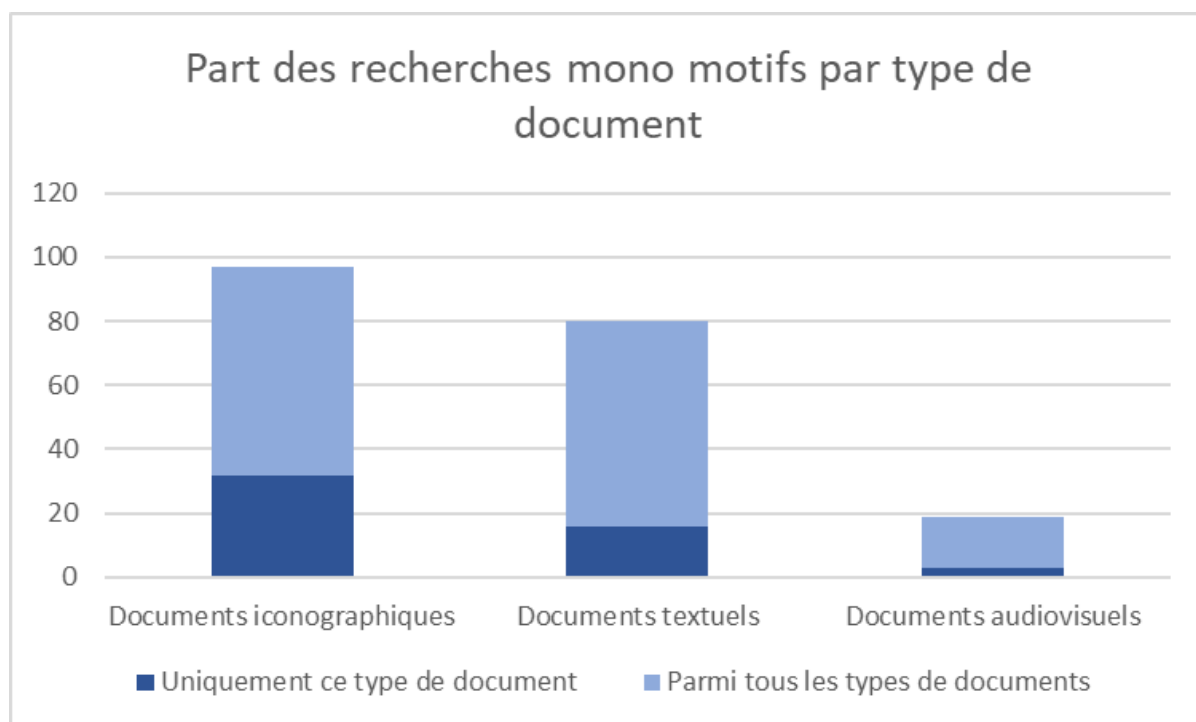


Figure 25 : La part des recherches mono-motifs par type de document

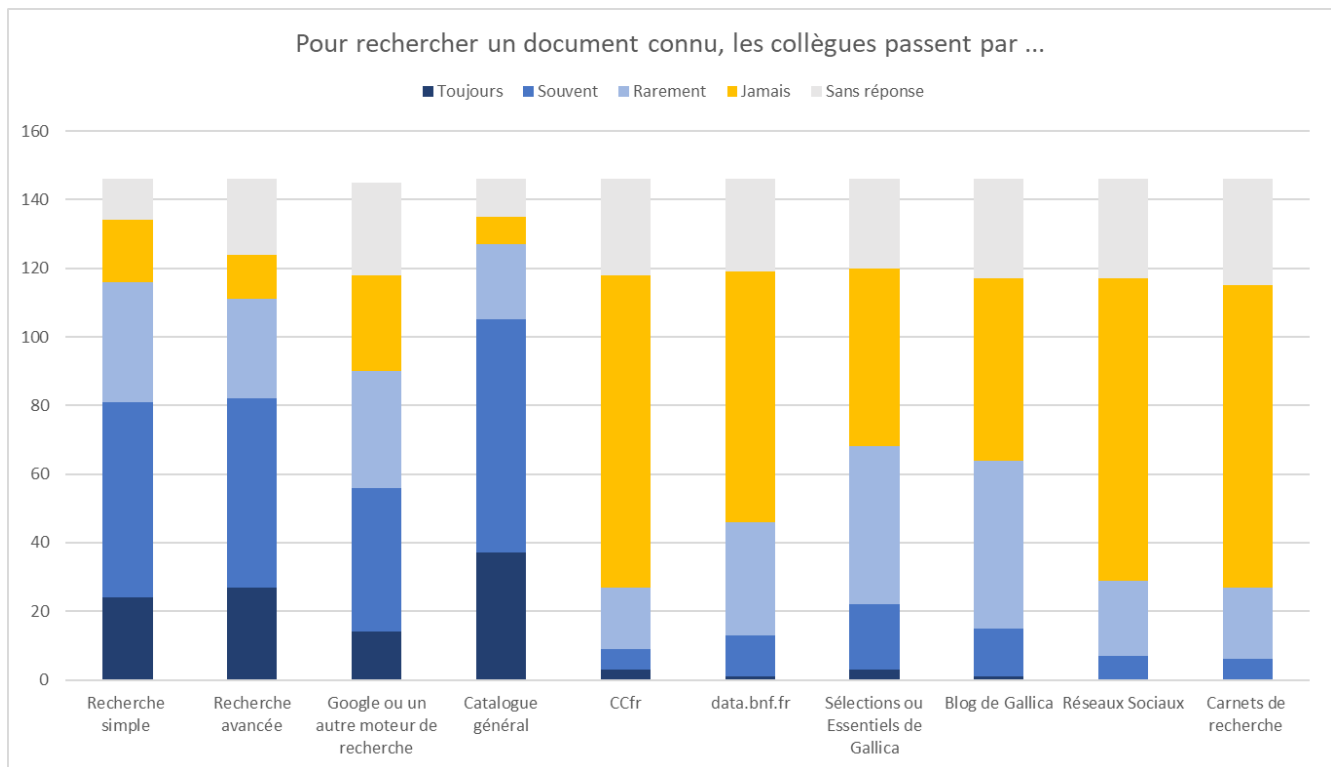


Figure 26 : Pour rechercher un document ciblé, les collègues passent par...

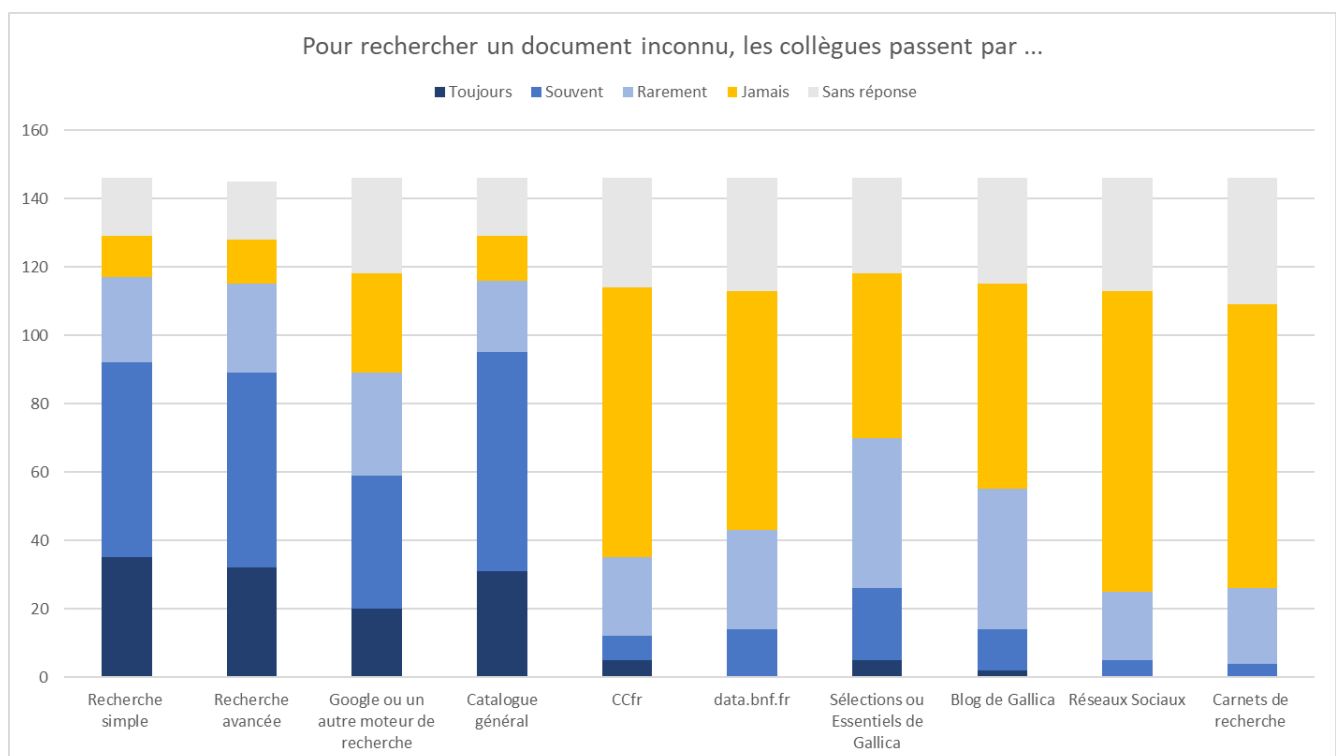


Figure 27 : Pour une recherche exploratoire, les collègues passent par...

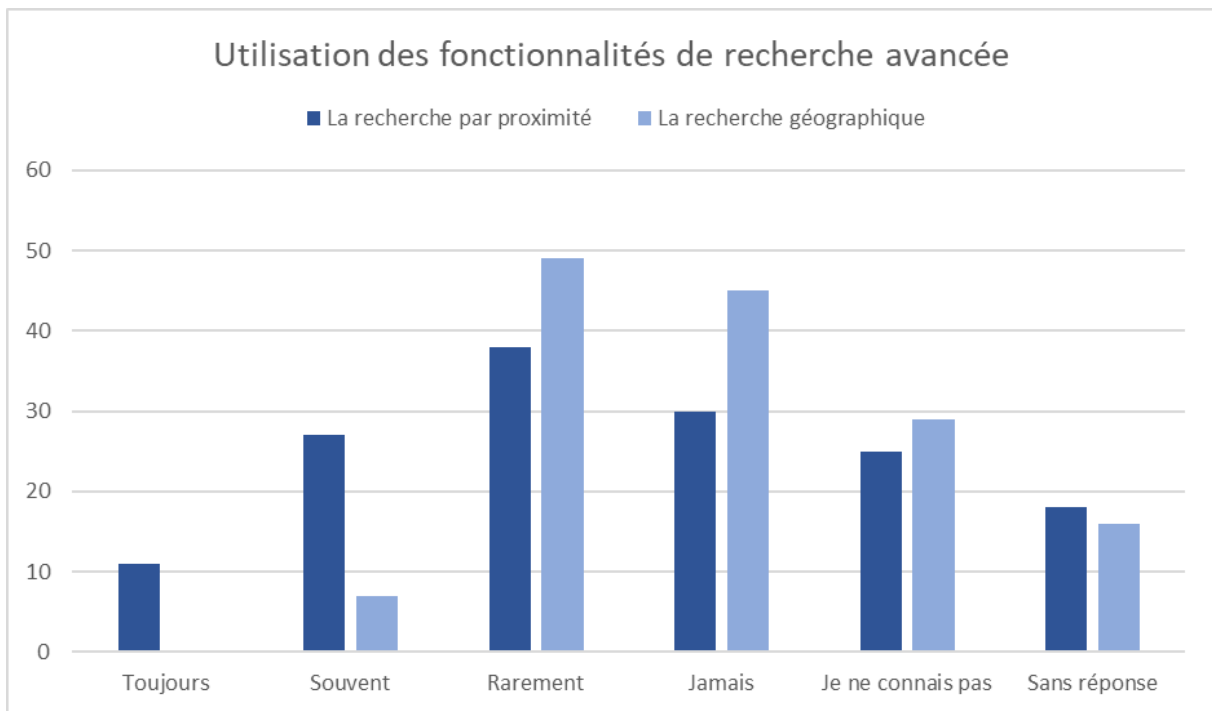


Figure 28 : L'utilisation des fonctionnalités de recherche avancée

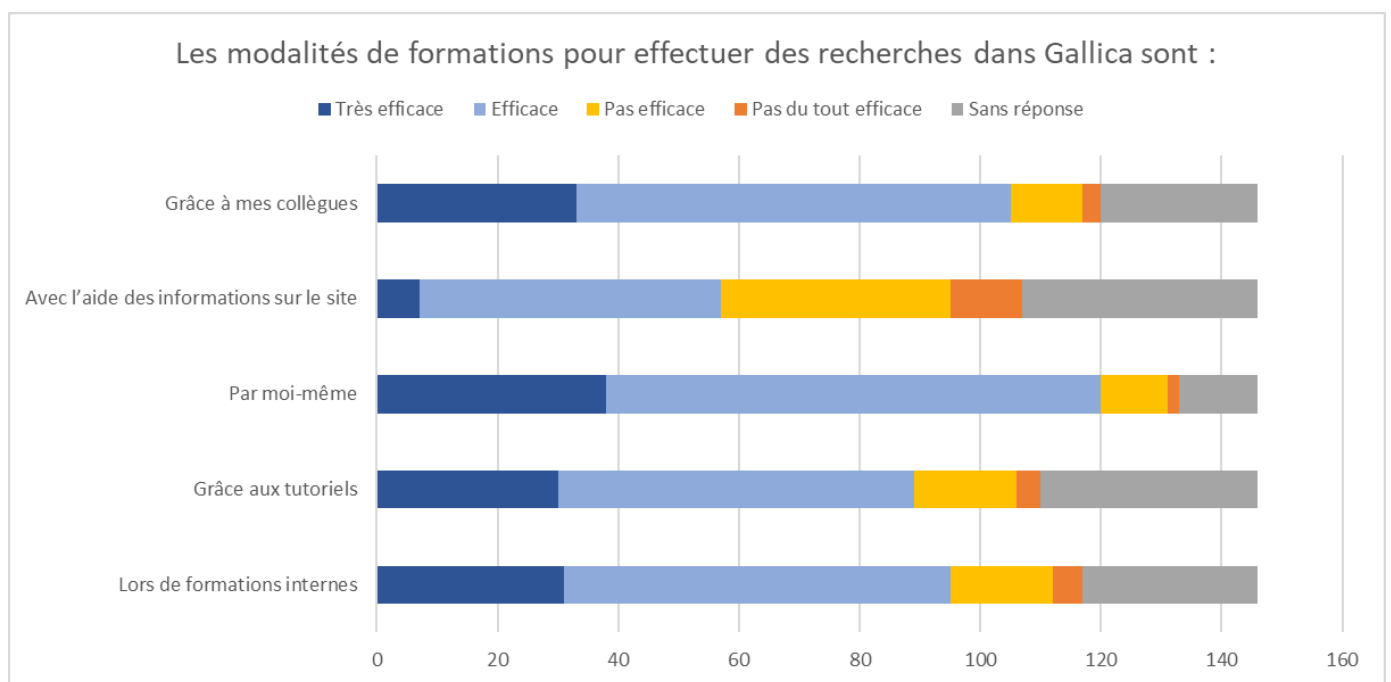


Figure 29 : L'efficacité des modalités d'apprentissage de la recherche dans Gallica

De quels types de formation complémentaires auriez-vous besoin ?

trouver un document avec la cote sans numéro d'ark

schéma arborescence

Pourquoi pas un service téléphonique dédié ?

Fonctionnement du moteur de recherche, astuces de recherche

Sur les API

présentation des onglets, page d'accueil et volets

Explorer tous les outils disponibles sur Gallica et toutes les bases de données liés à la BnF

Exportation des résultats, rapport de recherche

les contenus du catalogue

Exporter fichiers en haute définition

Recherches dans le Journal Officiel car pas toujours évident à manipuler, pour le reste, c'est nickel

Changer le système de recherche de Gallica, seule solution...

Icono

Recherche avancée

entrée dans la recherche de mots et de corpus via l'OCR

webinaires par zoom ?

Formation
Apprendre à être plus efficace car souvent je me perds.
des formations flash quand une nouvelle fonctionnalité est proposée ou pour une piqûre régulière de rappel
une formation sur Gallica
Une formation assurée par ceux qui travaillent sur Gallica, et qui connaissent ses récentes évolutions, c'est-à-dire une formation assurée par le DCP
Tutoriels
Formation sur complémentarité retronews Gallica, formation subtilités recherche avancée
mieux connaître la recherche avancée dans Gallica, les liens entre les différentes bases de données BnF (savoir passer de l'une à l'autre). Savoir également pourquoi certains documents n'ont pas été numérisés et ne sont pas visibles dans Gallica
Je ne savais pas qu'il existait des formations, ça m'intéresse carrément
Formation pointue sur la recherche avancée et la récupération des données

Figure 30 : Les formations à la recherche dans Gallica demandées par les répondants

Tableau obtenu après suppression des cellules vides et regroupement des réponses similaires.

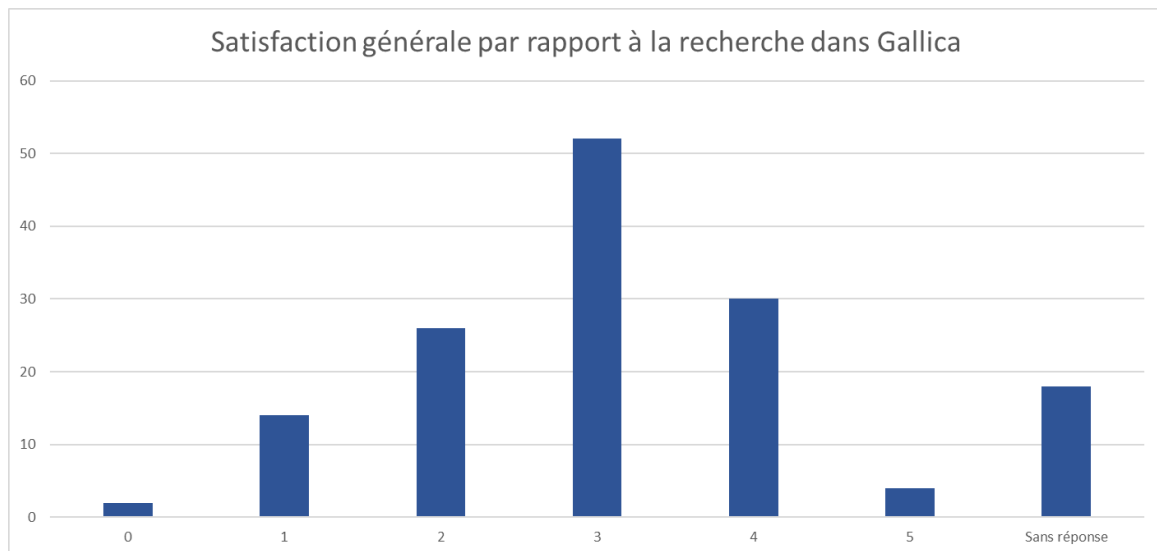


Figure 31 : Satisfaction générale des collègues par rapport à la recherche dans Gallica

$$\text{Moyenne théorique} = \frac{0 + 1 + 2 + 3 + 4 + 5}{6} = \frac{15}{6} = 2.5$$

$$\text{Moyenne} = \frac{(0 \times 2) + (1 \times 14) + (2 \times 26) + (3 \times 52) + (4 \times 30) + (5 \times 4)}{2 + 14 + 26 + 52 + 30 + 4}$$

Figure 32 : Calculs relatifs à la moyenne obtenue

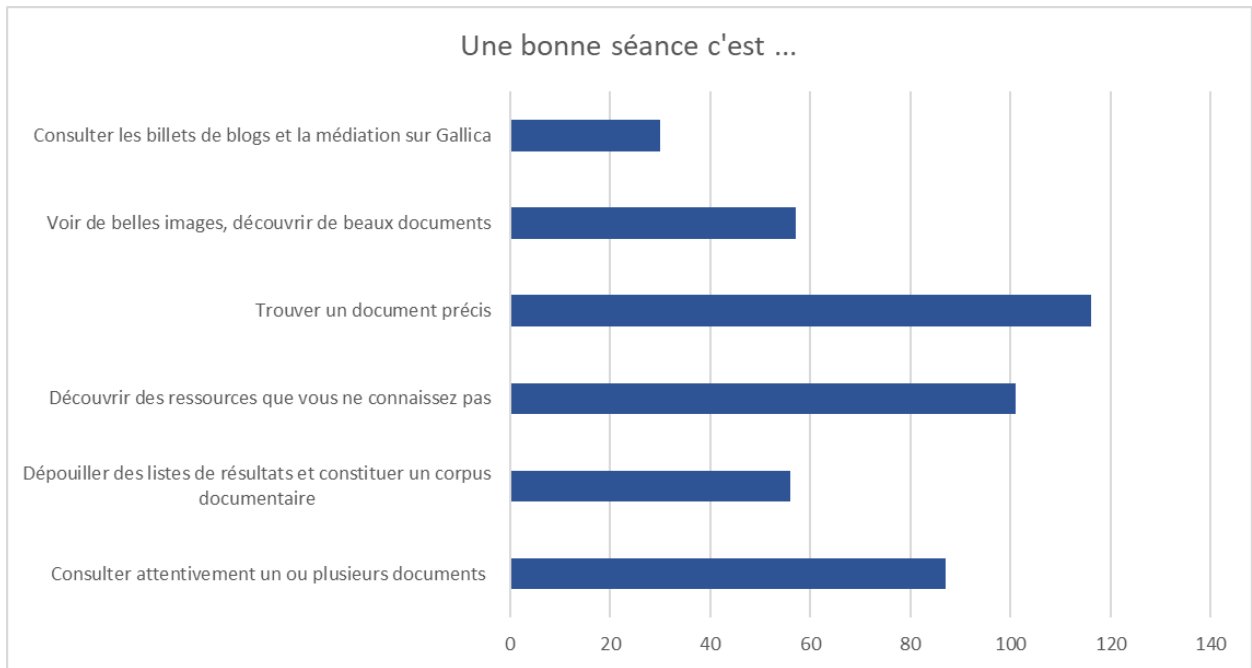


Figure 33 : Une bonne séance sur Gallica c'est ...

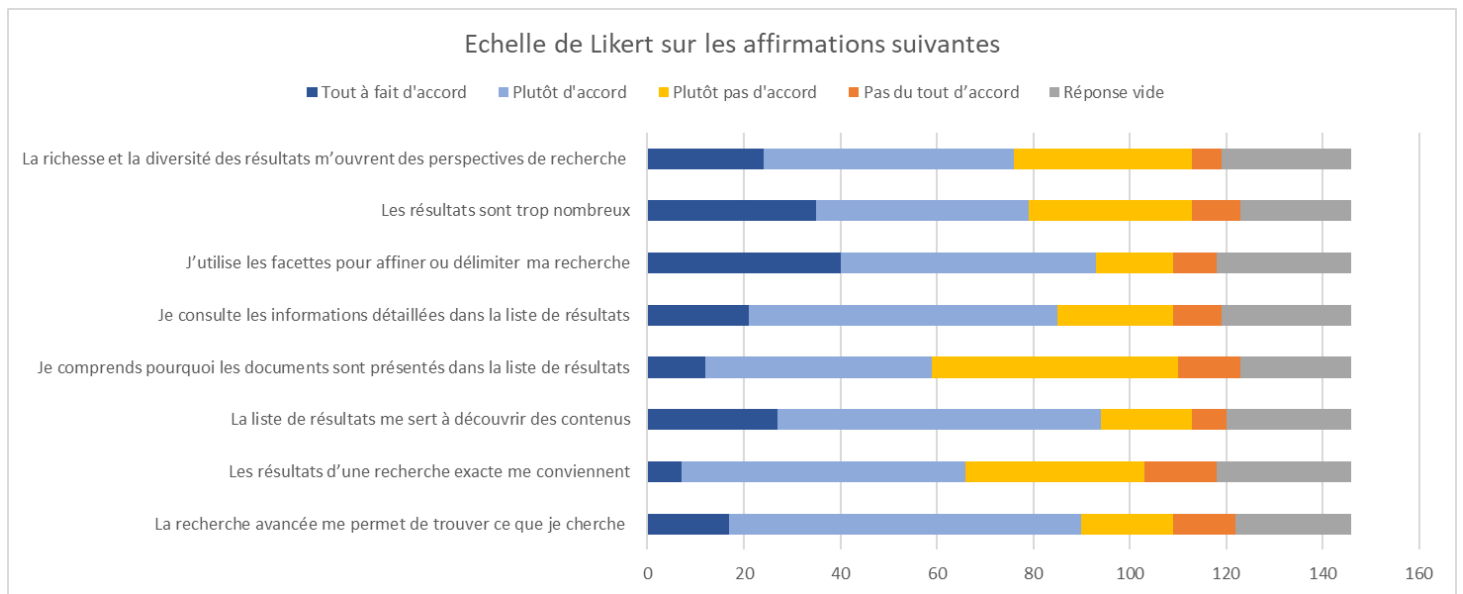


Figure 34 : Échelle de Likert sur plusieurs affirmations

Quelle fonctionnalité préférez-vous dans la recherche avancée de Gallica ?
Aucune
Recherche par proximité (20 réponses)
J'aime pouvoir chercher dans un document, une occurrence. Ce qui évite de tout lire si je n'ai pas le temps.
Le fait de pouvoir affiner ma recherche
la possibilité de chercher par auteur ou sujet mais sans plus, la recherche avancée m'aide rarement à faire une recherche performante. isfaite des résultats qui ressortent
recherche plein texte (mais à améliorer) ; recherche par proximité
Recherche OCR
OCR
facettage pour filtrer les résultats
La date de mise en ligne
la recherche par année d'édition
Recherche par cote (4 réponses)
la rubrique "Sujet"
Recherche par type de documents

Sélection par type de document et dates
liste déroulante des accès auteurs, sujets etc
protocole iiif
je ne sais pas, je ne l'utilise que pour mes recherches personnelles
la recherche par type de documents
typologie et années d'édition
aucune
Aucune, je passe systématiquement par le catalogue général
Pertinence des résultats ; présentation du numéro de périodique concerné (au lieu de renvoyer vers le titre du périodique) ; ergonomie de l'interface
Type de document
la recherche par type de document et par date
dates
proximité
La recherche par date de fascicule. Hyper pratique pour éviter de trop se perdre ou aller droit au but
la périodicité
Par type de document

Recherche plein texte
recherche par type de documents couplée avec recherche sujet
Notice et texte intégrale, année d'édition
Date de mise en ligne
La recherche par proximité ! génial ! Mais j'aimerais réussir par proximité de plusieurs mots et pas seulement 2
Le tri par période
Type de document
Source/cote et Date de mise en ligne
recherche avancée plein texte dans la presse numérisée par période
Recherche par date
texte intégral
type de documents
OCR
Cela dépend des besoins : pouvoir associer ou exclure plusieurs critères, pouvoir lancer la recherche sur le titre, toute la notice, Tout, Source/Cote, la recherche par proximité
Je ne sais pas
Surtout le tri une fois les résultats affichés

recherche par auteurs ou titres et types de document ou tout mot dans la notice
Le classement par types de documents et aires géographiques, la presse par thèmes
notice + format
La recherche par types de documents
la richesse des critères

Figure 35 : Fonctionnalités de la recherche avancée préférées par les répondants

Y a-t-il des fonctionnalités que vous ne comprenez pas ou n'utilisez pas ? Si oui, lesquelles ?
La recherche avancée est une usine à gaz. Je ne l'utilise jamais.
recherche par proximité, par format, géographique, par date de mise en ligne
Je ne suis pas seul à dire que le moteur de recherche, même si il a évolué, n'est pas encore satisfaisant
par proximité : je souhaite et je cherche souvent à l'utiliser mais je ne comprends pas ce que signifie "à une distance maximale de": que compte-t-on ? le nombre de mots ? articles et déterminants compris ou non? quid de la ponctuation?
par date de mise en ligne
la recherche exacte
la recherche par collection

La recherche géographique
La recherche géographique
recherche par collection et par thème
Je n'utilise pas la recherche par format, la recherche par collection ou thème, la recherche géographique. Au sein du bloc "Par bibliothèque ou site", j'utilise le menu déroulant "Saisir ou sélectionner une bibliothèque" mais pas "Saisir ou sélectionner un site" (car j'utilise ce bloc uniquement lorsque j'ai besoin d'exclure les collections des bibliothèques partenaires (indépendamment du site de consultation) pour me concentrer sur les collections BnF).
Recherche par proximité,
La recherche par thématiques, par format, par date de mise en ligne
Exportation des résultats
je ne l'ai jamais utilisé, mais je l'indique aux visiteurs
toutes, pas encore suivis de formation
Je ne comprends pas les suggestions de recherche dans la barre, et je ne comprends pas que dans les filtres "auteurs" une fois la recherche lancée on ne voit que le début de la liste alphabétique.
date de mise en ligne
Cf ci-dessus : les facettes pour affiner ou délimiter ma recherche
Proximité

je n'utilise pas la recherche par collection ou thème, par bibliothèque et par date de mise en ligne
recherche par année d'édition, je ne comprends pas comment elle fonctionne pour les périodiques, car elle me semble ne pas prendre en compte la date de publication des fascicules
type de document (les données sont fausses), par format ("né numérique" est peu compréhensible), recherche géographique
Je n'utilise jamais la fonction par format car dans mon cas, ça ne sert pas à grand chose
Globalement le moteur de recherche gallica est très décevant. Le catalogue général est beaucoup plus performant.
N'utilise pas : recherche par bibliothèque, par collection, par format, géographique, par date de mise en ligne.
Je n'utilise pas "Par date de mise en ligne"
recherche par proximité et recherche géographique
Non utilisées : Par bibliothèque ou site, par format, recherche géographique, date de mise en ligne
Autocomplétion dans la barre de recherche simple (renvoie souvent vers une absence de documents !)
année d'édition et format car ils restreignent trop la recherche
Je ne comprends pas toujours bien la fonction "Affiner", après l'avoir utilisée peut s'afficher une liste de ressources sans surlignage avec le mot recherché
Proximité, Recherche géographique

Format, Collection ou Thème, Recherche géographique (je les comprends mais ne les utilise pas.)
Je ne sais pas utiliser la recherche avancée dans Gallica.
anomalies fréquentes de résultats en recherche simple, "bruit" dans les résultats,
recherche par site et par collection : peu pertinentes pour mes recherches. Le rapport de recherche n'est pas très utile non plus
Aucune
Je n'utilise pas "par collection ou thème"
J'utilise rarement la recherche par thème ou collection, ou la recherche géographique, car j'ai peur de passer à côté de résultats et de trop restreindre ainsi la recherche
J'utilise très rarement : Par bibliothèque ou site, Par format, Par collection ou thème, Par date de mise en ligne et la recherche géographique mais je comprends leur intérêt pour certains besoins ou recherches.
Recherche avancée, la recherche par proximité et la recherche géographique
Je n'utilise pas par bibliothèque ou site ni par format (je n'en ai pas l'utilité)
Trouver le numéro exact d'une revue (à partir de son numéro ou de sa date)
la recherche avancée par formats
Difficile de chercher par date d'édition
les fonctionnalités qui n'opèrent pas sur tous les documents : sujet Rameau par exemple.

Figure 36 : Fonctionnalités mal comprises ou peu appréciées des répondants

Qu'est-ce qui vous manque pour chercher dans Gallica ?
UNE recherche avancée plus simple.
Plus de stabilité des résultats. Efficacité moteur de recherche.
Une recherche par proximité plus claire. / faire ressortir plus aisément des résultats par sujet
l'ocrisation des textes anciens (XVIe siècle en particulier)
Plus d'OCR
images en couleur
accès à un dictionnaire avec les variantes des termes indexés
La recherche par image
accès clair à la notice du document
différencier les photographies des estampes et du dessin
La recherche par localisation dans les départements BNF, la recherche par corpus ou collections
la recherche d'images semblables à une image que l'on fournit
Recherche d'images
Indexation, manque de documents présents dans la Banque d'Image

de bonnes métadonnées, comme dans les catalogues ou même data.bnf.fr
les données complètes des catalogues et un moteur de recherche efficace
L'indexation des images dans les documents textuels (livres + presse et revues)
Une recherche à partir d'une carte interactive.
tout y est
du temps
Mise à jour des cotes / des suggestions de recherche / attractivité de l'interface
les données des DEUX catalogue (et non pas seulement du CG comme le questionnaire l'a présenté) sont mal mappés, heureusement data.bnf.fr facilite un peu la tâche pour les termes indexés
du temps
Une meilleure pertinence ; présenter le numéro de périodique où se trouvent les résultats (au lieu de présenter le titre dans sa globalité)
Des résultats classés par pertinence, une meilleure description des documents, de la recherche fine (images dans des documents textuels par exemple), une meilleure OCRisation
Un équivalent de l'ocr pour les cartes
la troncature avec * qui a été supprimée ! ce serait plus qu'utile de remettre cette fonctionnalité de base !
Des métadonnées plus fiables sont ce qui me manque le plus pour effectuer des recherches précises dans Gallica

passer par les données du catalogue, avoir les mêmes résultats, en recherche par notice, que dans le catalogue. La troncature. Que les documents présents dans Gallica soient tous indexés avec les champs de recherche de Gallica (géographie, thème...). Améliorer les tris. Je pense aux lecteurs : simplifier les noms des champs, qui sont parfois des termes propres à la BnF

La possibilité de mettre des troncatures

Un bon moteur de recherche : on peut indiquer un titre précis avec des "..." et ne pas tomber sur le résultat...

Pans les facettes, permettre de filtrer les résultats par décennie (et pas seulement par siècle).

Synthèse du contenu du document en plus des métadonnées

un moteur de recherche de qualité

je ne sais pas où trouver les tutoriels mentionnés précédemment dans l'enquête

la troncature (comme dans le catalogue général)

Des résultats via la recherche simple qui remontent en premier les mots exacts recherchés et prioritairement placés dans le titre et/ou les noms d'auteurs. PS : je n'ai pas touché à la question suivante.

La fonction recherche d'images, mais en même temps il y aura moins de "mérite" lors d'une trouvaille de "pépité"

la reconnaissance d'images

C'est très difficile de trouver un document dont on connait la cote, les documents proposés ne sont souvent pas pertinents

Un vrai moteur de recherche avec une vraie indexation !

Une meilleure indexation des contenus notamment les images
la pertinence des résultats
Des résultats de recherche qui répondent de façon plus précise et pertinente à la chaîne de caractères saisis.
Une recherche simple plus performante. Et des facettes plus précises (par décennies pour les dates d'édition). Pouvoir séparer les numérisations en noir et blanc de celles en couleurs
Rien
On a l'impression qu'on trouve plus facilement sur ressource en passant par Google en indiquant Gallica + la recherche.
la possibilité de lancer une recherche sur toutes les images, y compris celles issues d'ouvrages ou de revues ; la troncature * ; pouvoir associer ou exclure davantage de termes
Je ne sais pas exactement
L'idéal serait de continuer à faire de courtes vidéos de présentation (premiers pas, puis possibilités diverses), si possible également sous-titrées en français pour l'accessibilité, et disponibles en anglais.
Une meilleure sélection des documents par pertinence, par exemple commencer par les titres qui comprennent les mots clés, plutôt que de chercher dans le plein texte.
Un accès direct à Gallica intramuros depuis le catalogue général, qui serait bien utile dans le cadre d'une recherche sur site. Sans quoi, il est nécessaire de passer par biblionautes, ou d'y accéder à partir d'un document numérisé en particulier.

L'habitude, sans doute : je suis si habituée au CG que c'est mon premier réflexe. Je passe par les Essentiels lorsque je travaille pour des publics extérieurs, lorsque fais de la valorisation ou prépare un cours.
Distinction ouvrages/périodiques notamment lorsqu'on recherche un numéro en particulier
un développement de recherche dans le texte plus fin (création de corpus de thèse plus fins) / développement de la recherche dans les collections-valorisation des collections
une formation / un tuto, j'ai l'impression de rester en surface, par méconnaissance, et de passer à côté de résultats pertinents
un meilleur moteur de recherche pour le blog et les sélections
Pouvoir interroger en langage naturel
efficacité et discrimination des résultats
trouver rapidement ce que je recherche

Figure 37 : Ce qui manque aux répondants pour chercher dans Gallica

Classement	Proposition	Moyenne
1	Améliorer l'ergonomie, simplifier et alléger l'interface	4,734513274
2	Clarifier la liste de résultats	4,787610619
3	Ajouter des fonctionnalités et paramètres	5,592920354
4	Ajouter de la granularité documentaire	5,920353982
5	Expliquer la recherche	6,469026549
6	Faire converger les fonctionnalités	6,539823009
7	Proposer de la visualisation de données	6,601769912
8	Ajouter des fonctionnalités innovantes	7,008849558
9	Proposer une recherche spécifique pour certains types de documents	7,221238938
10	Présenter la collection numérique	7,283185841
11	Simplifier la recherche	7,318584071
12	Pouvoir personnaliser l'interface	8,522123894

Figure 38 : Classement par ordre de préférence/ priorité des propositions d'amélioration de la recherche dans Gallica

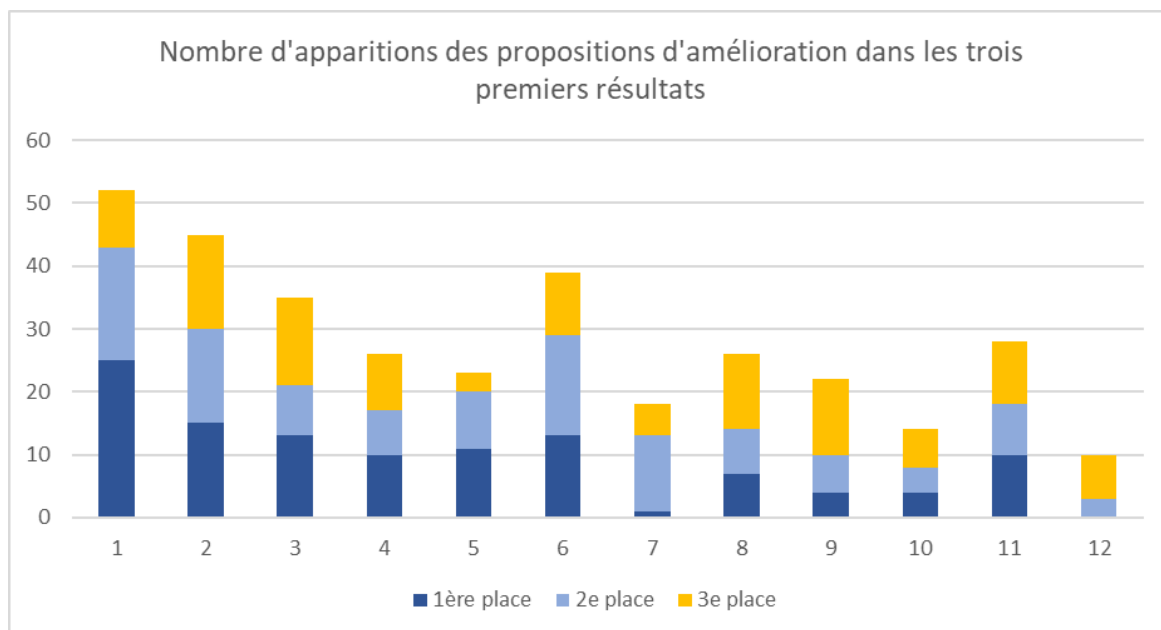


Figure 39 : Nombre de fois où chaque proposition s'est retrouvée en tête du classement (1^{ère}, 2^e ou 3^e place)

Souhaitez-vous ajouter quelque chose ?

Améliorer le zoom qui parfois ne fonctionne pas.

Pour moi, c'est surtout l'affichage/consultation des documents qui pose problème

Consulter les documents avec le moyen des flèches droites gauches est fastidieux un menu déroulant serait appréciable

Pas toujours simple de chercher une image si elle n'est pas individuelle.

J'utilise fréquemment le blog gallica, dans lequel j'ai besoin de faire des recherches (pour vérifier si un sujet d'article a déjà été traité ou non par exemple) et la barre de recherche fonctionne très mal. En fin de compte je passe par Google en tapant "blog gallica" + le mot clé qui m'intéresse. Je procède régulièrement de même pour Gallica car j'ai remarqué que je trouvais par Google des résultats que je ne trouve pas avec la recherche avancée. Je suis familière de cette recherche avancée, mais en fin de compte, Google me sert tout autant.

Un des principaux problèmes actuels est l'affichage des résultats : quand on effectue une recherche avec plusieurs expressions en plein texte, seulement une de ces expressions (la dernière) s'affiche dans les résultats. Par ailleurs, il faudrait vraiment pouvoir distinguer la presse des revues savantes car ce sont deux types de périodiques qui n'ont pas du tout la même utilité pour les chercheurs.

améliorer les rapports de recherche

Difficile de hiérarchiser dans la question 4 ! La colonne de gauche sur la carte d'identité des documents n'est pas facile avec tous les volets déroulants, il y a sûrement l'ergonomie à revoir mais je me rends compte que c'est complexe. Attention entre Gallica et LES ESSENTIELS tout seul (pas les Essentiels de Gallica) : il y a confusion injuste . Gallica est évidemment bcp plus riche et LES ESSENTIELS ne renvoient même pas sur Gallica... Gallica est un outil formidable !

Un problème de contenu : J'avais écouté et vérifié techniquement, en 1999 pour le département SVM, le corpus audio des cours de G. Deleuze, j'avais comptabilisé 177 cours pour une durée de 400 heures. Par la suite, les CD des cours ont été versés dans Gallica. Comme les dates des cours apparaissent sans ordre chronologique dans Gallica, il est difficile de mener une écoute assidue sur un semestre, sur une année, on risque fort de louper des séances ... - Sur le site de Gallica, à la page de présentation du corpus des cours, est indiqué 180 cours, 413 heures. De plus, dans les intitulés des cours, il manque les cours sur Leibniz qui datent de 1980 (5 cours en tout). Merci à vous !

(j'ose) Que la recherche sur Gallica soit aussi simple que lorsque l'on réalise sa recherche sur google

L'utilisation du moteur de recherche Gallica est toujours une expérience pénible ; tous les collègues, à ma connaissance, préconisent aux lecteurs d'effectuer leurs recherches avant la consultation dans Gallica

Je constate un décalage entre la recherche en mode texte (OCR) via la recherche avancée et via la petite loupe à droite d'un document : lorsque l'on place des guillemets dans la recherche avancée pour chercher une expression exacte, cela est bien pris en compte dans la liste des résultats, en revanche lorsqu'on place des guillemets dans la petite loupe d'un document, les guillemets ne sont pas pris en compte et les résultats affichent donc des mots isolés. La recherche avancée me sert donc à contourner ce problème : une fois le document qui m'intéresse identifié, je retourne dans la recherche avancée et, dans le bloc "Notice et texte intégral", je copie-colle le titre, l'auteur ou la cote du document (ou tout autre élément me permettant de m'assurer que le document qui m'intéresse sera le seul à ressortir dans la liste des résultats), et j'ajoute une autre ligne "Texte intégral" avec entre guillemets l'expression exacte que je recherche. Autrement dit, j'utilise la recherche avancée pour contourner les défauts de la petite loupe de recherche dans un document précis. Il me semblerait donc très utile que les modalités de recherche via la petite loupe soient améliorées, avec prise en compte des guillemets pour une recherche exacte, et (soyons fous) avec l'ajout d'une option "recherche par proximité" une fois arrivé sur le document, avec le même fonctionnement que via la recherche avancée.

Il faudrait améliorer l'océrisation et la vitesse d'affiche, très irrégulière.

mettre en avant gallica sur le site
sur l'écran précédent, la question "Parmi la liste suivante, quelles seraient les principales améliorations pour faciliter la recherche dans Gallica ?" est vraiment infernale à utiliser et beaucoup trop de paramètres proposés.
J'adooore Gallica ! Merci à toutes les équipes impliquées.
Faire un lien vers Europeana
vous semblez ignorer que certains documents numérisés sont catalogués dans BnF archives et manuscrits, c'est préoccupant sur la connaissance de l'environnement numérique de Gallica et les enjeux autour de la fouille de données non éditorialisées
Gallica ne me semble pas adapté à des recherches de type bibiothéconomique, croisées et avec filtres
Prioritaire : améliorer la pertinence des résultats dans la recherche en plein texte.
Longue vie à Gallica !
J'utilise de manière très fréquente la fonction "type de document"
remettez la troncature !
Ce serait interessant de pouvoir exploiter davantage de métadonnées des documents numériques, date de umérisation, institution de conservation de l'original
non
Globalement Gallica est un super outil, incontournable aujourd'hui pour la recherche ou la médiation. Mais son moteur de recherche est catastrophique. Heureusement il y a le catalogue général (et Google...)
Je pense qu'il faut plutôt ajouter des possibilités d'affinage, d'autres outils de précision à combiner que simplifier la recherche. Et merci encore pour Gallica, c'est super !
Pour trouver la numérisation d'un manuscrit c'est beaucoup plus facile de passer par BAM que par Gallica

Je n'ai jamais vu un moteur de recherche aussi mal fichu et moins performant que celui de Gallica.

J'aurais souhaité des questions concernant la recherche dans Gallica et dans BAM (tous les documents de la BnF ne sont pas dans le Catalogue général ;))

Il faut vraiment améliorer la recherche simple. Même si l'on met des guillemets les résultats sont trop imprécis. Pareil pour la fonction "Affiner (Lancer la recherche dans ces résultats)" qui serait très utile si elle marchait bien, mais qui en général annule tous les résultats de la première recherche. Il serait utile aussi de pouvoir rechercher les documents qui contiennent des illustrations.

une traduction complète du site en anglais

Outil indispensable. Chapeau pour votre travail de maintenance, d'enrichissement et d'amélioration ! Les bruits dans les résultats peuvent avoir une dimension heuristique et ils obligent l'esprit de rester vigilant et en vie.

L'affichage des documents est souvent laborieuse et longue. QUand on utilise le zoom cela ne fonctionne pas de façon intuitive. J'ai cherché récemment à faire entendre des documents audios de personnalités célèbres impossible de les retrouver. Quand on cherche un groupe de mots on a des résultats de recherche non pertinents car les mots sont pris séparément. Les filtres ne sont pas toujours assez fin. Pour la presse, il est compliqué de voir le mot clé dans l'article. Les suggestions proposées à côté des documents consultés ne sont pas souvent pertinentes. Dommage que les informations bibliographiques détaillées n'apparaissent pas d'emblée. Beaucoup d'améliorations sur les fonctionnalités, l'ergonomie, la pertinence et la rapidité sont attendues sur ce merveilleux réservoir de contenus en ligne

le moteur renvoie trop de documents non pertinents à cause d'un paramétrage trop laxé sur la recherche plein texte (OCR)

Compliments et encouragements

Gallica est un outil exceptionnel, bravo aux équipes! Ma crainte est que des hackers s'en prennent un jour avec succès à ce bijou d'intelligence collective. Merci pour ce questionnaire, intéressant à remplir.

Non, c'est quand même drôlement bien, Gallica.

Bon courage chers collègues !

Mais Gallica, c'est merveilleux !

Gallica est un outil de recherche de très grande qualité. Les billets de blog permettent de découvrir certains thèmes et les ressources associées.

Gallica est un site de référence très complet et déjà assez ergonomique vu la complexité des sources et des documents accessibles.

Figure 40 : Champs libre pour que les collègues puissent s'exprimer