

Projet Savoirs

Cahier des charges



Réalisé par

Sylvain Besson

Julien Muzet

Online Poulat

Camille Roblin

Laurence Thévenieau

Master Humanités Numériques 2020 - 2021



Table des matières

1. Présentation du projet Savoirs et définition du besoin général.....	3
1.1. Une brève présentation du projet.....	3
1.2. Les parties prenantes de Savoirs.....	3
2. Le besoin du client.....	4
3. Le mode de gestion des données	5
3.1. Plan général de gestion des données	5
3.2. L'arborescence de la plateforme	8
3.3. Points sur le RGPD.....	10
4. Les livrables	12
4.1. Le questionnaire et l'autoformation	12
4.1.1. Questionnaire.....	12
4.1.2. L'autoformation.....	15
Une explication en plusieurs temps / étapes.....	15
4.2. Le profil	17
4.2.1. Le paramétrage du profil	17
4.2.2. Les collections	18
4.2.3. Suggestion de texte.....	20
4.2.4. Historique des parcours de lecture.....	20
4.2.4.1. Le fil d'Ariane	20
4.2.4.2. Accès à l'historique de recherche et sa visualisation.....	21
4.2.5. L'export de la recherche	22
4.3. L'espace texte.....	22
4.3.1. Les marque-pages.....	23
4.3.2. Le surlignage	23
4.3.3. Définitions de concepts	24
4.3.4. L'annotation	24
5. Le calendrier de développement	25
5.1. Ordre à suivre pour le développement des fonctionnalités.....	25
5.2. Le calendrier détaillé	26

1. Présentation du projet Savoirs et définition du besoin général

1.1. Une brève présentation du projet

Née d'une publication papier, *Lieux de Savoirs*, dirigée par Christian Jacob, Savoirs est une plateforme numérique qui, dans son esprit-même, s'approprie les concepts de sérendipité et de science ouverte. Ainsi, elle se veut être pour l'utilisateur un lieu où la connaissance est à la fois familière et source de découverte, le faisant emprunter des chemins qu'il connaît déjà et d'autres bien plus inconnus, créant des ponts qu'il peut décider de franchir... ou pas.

Le projet "Savoirs" s'inscrit dans le développement des approches interdisciplinaires dans le champ de la recherche, et plus largement dans le développement d'une logique *d'Open Science*. Mettre à disposition et valoriser les corpus pour les faire découvrir, c'est permettre à des utilisateurs de profil divers d'acquérir, par des fonctionnalités propres à la plateforme, une autonomie dans l'exploration et la manipulation des corpus ; surtout, leur permettre d'entrer dans une logique de sérendipité favorisant les liens, les dialogues inattendus et féconds : tels sont les objectifs généraux que se donne Savoirs.

Dans cette perspective, deux éléments sont essentiels : la nécessité d'avoir un algorithme de recommandation paramétrable par l'utilisateur et transparent pour éviter l'effet boîte noire, et celui d'avoir un profil utilisateur qui viendrait s'implémenter sur ce même algorithme. Ce cahier des charges décrira le développement du profil utilisateurs et les fonctionnalités imaginées pour celui-ci.

1.2. Les parties prenantes de Savoirs

Les parties prenantes de Savoirs sont triples. Tout d'abord, nous avons le client, qui est l'équipe de Savoirs dans laquelle nous retrouvons notamment :

- Christian Jacob, directeur d'études à l'EHESS en tant que chef de projet
- Jean-François Bert, sociologue et historien des sciences sociales à l'Université de Lausanne (IHAR)
- Axel Le Roy, doctorant à l'Université de la Sorbonne (labex OBVIL)
- Simon Dumas-Primbault, enseignant – chercheur à l'Ecole Polytechnique Fédérale de Lausanne (LHST), qui nous a suivis personnellement.

Ensuite, nous avons un ou plusieurs prestataires qui seront des développeurs lorsque l'équipe de Savoirs souhaitera réaliser l'espace profil. Enfin, nous avons les contacts, les étudiants du master Humanités numériques, avec chacun des spécialités différentes, encadrés par Jean – Philippe Magué, chercheur en linguistique à l'ENS de Lyon (ICAR).

Contacts :

- Sylvain Besson : Sylvain.Besson1@univ-lyon2.fr
- Julien Muzet : julien.muzet@univ-lyon2.fr
- Online Poulat : Online.Poulat@univ-lyon2.fr
- Camille Roblin : c.roblin@univ-lyon2.fr
- Laurence Thévenieau : laurence.thevenieau@free.fr

2. Le besoin du client

Ce document constitue une ébauche développée mais non – exhaustive d’un cahier des charges pour un développeur, dans la mesure où l’équipe *Savoirs* devra valider les propositions que nous formulons et ajouter les éléments que nous n’avons pas à notre disposition et à notre échelle. Nous nous sommes intéressés au développement de la notion de profil utilisateur. Il s’agira ainsi de permettre à un utilisateur de personnaliser son expérience en :

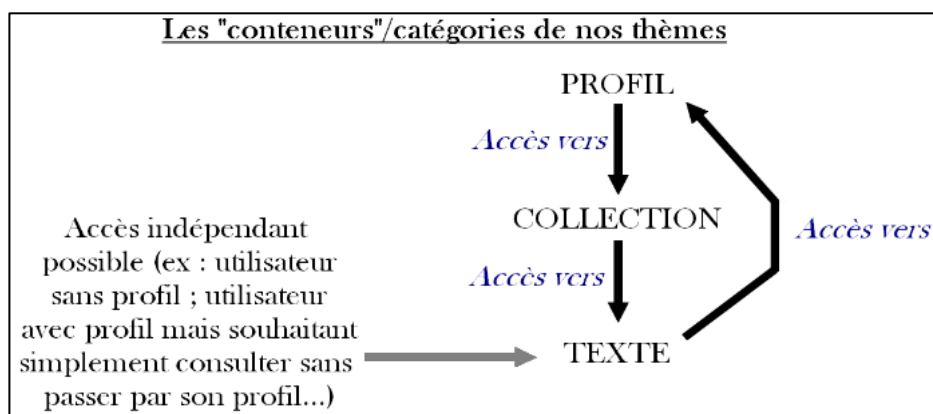
- modifiant et sauvegardant ses paramètres de recherche ;
- profitant d’un site suivant des principes éthiques “*open source*” et de transparence face au futur algorithme de suggestion en cours de développement ;
- profitant d’un site favorisant la sérendipité, l’interactivité et l’accessibilité: sur ce point nous souhaitons que le compte s’intègre dans les mêmes exigences non fonctionnelles de *Savoirs*, à savoir les règles d’accessibilité de niveau AA, une interface “*responsive*” en français ou anglais ou en italien au choix.

Furent distinguées trois catégories de réflexions que nous résumons ci-dessous avec les fonctionnalités que nous développerons en partie 4, sauf pour “communauté” (cf. infra) :

Compte utilisateur	Texte	Communauté
Espace méta-texte	Marque-page	Forum de discussion
Historique des parcours de lecture	Suggestion d'ajout de texte	Partage et mise en commun de notes
Question utilisateur et autoformation	Annotation	Proposition de correction du texte à l'auteur
Paramétrage du profil utilisateur	Surlignage	Communication avec l'auteur
Création de collections		
Export des recherches et favoris		

En termes imagés, *Savoirs* numérisant le concept complet de bibliothèque physique, la page texte serait le livre, le compte utilisateur serait le bureau de travail, et le thème de la communauté serait la possibilité de l’échange ou du projet de groupe. Ce dernier thème est écarté sur le court terme car il ne convient pas nécessairement aux besoins immédiats de la plateforme. Néanmoins, il devra être anticipé : par exemple, le système d’annotation pourrait à terme, et selon les volontés de *Savoirs*, inclure une fonctionnalité de partage.

D’un point de vue pratique, nous avons imaginé ce compte comme s’intégrant dans un site comportant une page d’accueil et l’ensemble des textes, ce que le schéma ci-dessous pourrait résumer par l’accès indépendant au texte. Dès la création des profils utilisateurs, nous pensons que nos apports se structureraient tels qu’illustrés par le du schéma ci – dessous :



3. Le mode de gestion des données

3.1. Plan général de gestion des données

Il faut, tout d'abord, rappeler que l'espace profil n'est pas une obligation pour l'utilisateur, mais qu'il s'agit plutôt de lui donner davantage de possibilités dans son usage de la plateforme. Par conséquent, il est possible d'imaginer trois types d'utilisateurs de cette dernière :

- Un utilisateur qui possède un compte
 - il aura accès à l'intégralité des fonctionnalités ;
 - il lui faudra, lors de sa première connexion, se créer un profil en y mettant les informations qu'il souhaite ;
 - il n'aura pas besoin de rentrer à nouveau ses informations.
- Un utilisateur peut souhaiter avoir un profil sur la plateforme, tout en refusant que certaines données soient collectées.
 - Certaines fonctionnalités seront impactées. À commencer par la recommandation qui ne prendra pas en compte les informations que l'utilisateur saisit lorsqu'il s'inscrit sur la plate-forme et remplit le questionnaire ;
 - Cela dépendra de la nature des données que l'utilisateur aura accepté de partager avec Savoirs. Il devrait être informé des fonctionnalités qui seront impactées par le refus de collecte de données ;
 - Si l'utilisateur s'oppose à toute collecte de données, aucune fonctionnalité de l'espace profil ne fonctionnera. Il vaudra mieux l'informer de ne pas se créer de profil, par un message d'avertissement, afin qu'il ne crée pas un compte inutilement.
- Un utilisateur ne souhaite pas avoir un compte (que ce soit à cause de ses données ou qu'il n'y trouve pas d'intérêt dans son utilisation) :
 - il pourra quand même accéder à la plateforme, mais n'aura accès qu'aux fonctionnalités de base qui sont les suivantes :
 - consulter les articles ;
 - choisir des critères géographiques, temporels et conceptuels (ils seront pas enregistrés après la fermeture de la session) ;
 - avoir des recommandations ;
 - naviguer avec la courbe temporelle pour voir les article consultés au cours de la session.

Au sein du projet Savoirs, nous retrouvons différentes données liées à l'utilisateur. Ce dernier va pouvoir les enregistrer lorsqu'il se crée un profil. Il pourra par la suite les modifier quand il le veut et autant qu'il le souhaite. Nous retrouvons, dans ces données, le pseudonyme (*string*), l'adresse mail (*string*), les champs disciplinaires de recherche (*string*) et le mot de passe (*string*). Ces derniers seront au format HTML puis stockés dans un premier temps dans un fichier JSON, puis, à terme, dans une base de données relationnelle PostgreSQL.

Le mot de passe devra en plus avoir un encryptage¹. De plus, à chaque utilisateur sera associée une adresse IP (*integer*) afin de relier une session à un profil. De même, il faudra des informations sur le moment où le compte est créé (*datetime*) et lorsqu'il est mis à jour (*datetime*). Toutes ces données seront conservées dans la base de données, du jour où l'utilisateur se crée un compte jusqu'au moment où il désire le supprimer.

¹Voir la partie 7.2 pour réaliser des mots de passe sécurisé sur Ruby :
http://french.railstutorial.org/chapters/modeling-and-viewing-users-two#sec:secure_passwords

Par ailleurs, il est possible pour l'utilisateur, à tout moment, d'avoir accès à ces données sur son **espace profil**, à l'exception des mots de passe pour des raisons de sécurité.

Il aura uniquement la capacité de voir le nombre de caractères dudit mot de passe. Il pourra également consulter ces données auprès du gestionnaire de ces dernières. Les données relatives à chaque utilisateur pourront être encodées dans le fichier JSON de la façon suivante dès qu'un profil est créé :

Figure 1 : Encodage d'un profil en JSON

```
{
  "cookie_id":
  "IP_adress"
  "pseudonym":
  "mail_address":
  "disciplinary_fiefs":
  "encrypted_password " :
  "created_at" :
  "updated_at " :
}
```

De plus, il sera nécessaire de gérer les données agrégées liées à **l'expérience de l'utilisateur durant son temps de session**. Tout d'abord, les différents parcours de lecteurs que l'utilisateur est amené à faire : ces derniers sont générés à partir des différents articles – qui pourraient être distingués par un id (*integer*) - que l'utilisateur consulte mais aussi les concepts auxquels les articles sont associés. Par ailleurs, l'utilisateur va pouvoir par lui-même choisir les concepts, les périodes et les zones géographiques qui l'intéressent. Les concepts et les zones géographiques auraient chacun un id. Une période, en revanche, serait composée de deux *integer*. Ainsi, il serait envisageable d'avoir ces **données conservées dans un fichier JSON** sous le format suivant :

```
{
  "settings" : [
    {"concept" : ["id_concept_1", "id_concept_2", "id_concept_3"]}
    {"place" : ["id_place_1", "id_place_2", "id_place_3"]}
    {"date" : [date_1, date_2]}  ]
}
```

Ces données sont assemblées afin de créer un parcours de lecture et sont stockées dans des fichiers JSON. Ils ne seront donc qu'en partie consultables à l'utilisateur puisque les paramètres qu'il a sélectionnés lui seront accessibles et qu'il pourra les modifier s'il le souhaite.

Néanmoins, les données collectées sur son parcours de lecture ne lui seront pas connues.

De la même façon, les différents articles consultés seront aussi conservés afin de pouvoir **renvoyer à l'utilisateur son historique**. L'arborescence dans laquelle le fichier JSON pourrait s'intégrer serait la suivante :

```
/nav_hist
  id_art_1.json
  id_art_2.json
  id_art_3.json
```

Figure 2 : Arborescence des données de l'historique de navigation

Viennent ensuite des données issues des **annotations**, des **commentaires** et des **surlignages**. Dans un fichier JSON, une annotation et un commentaire auraient le numéro du

paragraphe (*integer*) et le contenu (*string*). Ils pourraient être encodés de la façon suivante par l'intermédiaire d'un fichier JSON :

```
{
  "annotation" : {[num_paragraphe,"contenu_ann"]}
}

{
  "comments" : {[num_paragraphe,"contenu_com"]}
}
```

Figure 3 : Intégration des annotations et commentaires

La partie surlignage serait encodée de la même manière, mais en ajoutant au numéro de paragraphe (*integer*) la couleur en hexadécimal (*integer*) et le nom du thème (*string*). Cela se présenterait de la façon suivante :

```
{
  "highlight" : {[num_paragraphe,
#couleur, "nom_thème" ]}
}
```

A cela, il est possible d'ajouter **les favoris et les titres de collections**. L'arborescence pour les favoris, intégrée à un fichier JSON, se présenterait ainsi :

```
/bookmark
    /folder
        /subfolder
            id_art_1.json
            id_art_2.json
        /folder
            id_art_3.json
```

Quant aux **collections**, un fichier JSON regrouperait les différents noms de ces dernières. Il est possible de les envisager sous le format suivant :

```
{
  "collection_headers" : ["collection_header","collection_header","collection_header"]
}
```

La durée de conservation des données agrégées ainsi que de celles relatives à l'expérience de l'utilisateur sur la plateforme aura pour limite la suppression du compte par l'utilisateur – ou par un membre de l'équipe Savoirs. La raison en est que ces données sont indispensables - en particulier celles liées au parcours de lecture - pour assurer un bon fonctionnement du profil, bien que l'utilisateur ne soit pas contraint de fournir ces informations.

Enfin, **les données relatives à la suggestion d'ajout de texte** seront conservées le temps que l'équipe Savoirs traite la demande, qu'elle se solde par une acceptation ou par un refus de cette dernière. Le format dans lequel la suggestion d'ajout d'un texte pourrait se structurer de la façon suivante :

```
{
  "title" : ""
  "author" :
  "date" :
  "publication" :
  "volume" :
  "Issue" :
  "DOI" :
}
```

Figure 4 : Données recueillies lors de la suggestion d'un texte par un utilisateur

En définitive, il y aurait **un dossier pour chaque utilisateur de la plateforme**, susceptible d'être structuré comme suit :

```
/id_user
    id_user.json
/settings
/navi_history
    id_art_1.json
/annotations
    id_art_1.json
/comments
    id_art_3.json
/highlight
    id_art_5.json
/bookmark
    /folder
        /subfolder
            id_art_5.json
            id_art_3.json
/collection_headers
    collection_headers.json
/suggestion
    sugestion_1.json
```

La totalité de ces données seront hébergées au sein de l'EHESS et, de fait, tomberaient sous la responsabilité de cette dernière. Une personne devrait aussi être assignée à la vérification de la qualité des données.

3.2. L'arborescence de la plateforme

Nous avons réfléchi à une proposition d'arborescence reprenant les pages centrales de la plateforme : la présentation, les corpus et l'interface de lecture et de paramétrage des algorithmes, ainsi que les pages correspondants aux fonctionnalités que nous avons envisagées. Elle consiste en au moins deux arborescences distinctes : l'une pour un utilisateur qui peut se connecter via son compte et l'autre pour un utilisateur sans compte.

Utilisateur avec un compte : Entrée directe sur la page Corpus, ou Mes collections.

Arborescence menu principal

- Accueil (le fil des idées)
- Le projet
 - o Historique du projet
 - o L'encodage en xml tei
 - o Le thésaurus
 - o Comité et partenaires (figurant aussi dans la page d'accueil...)
- Le corpus
 - o Rechercher un texte
 - Marques-page, méta-texte, surlignage
- Profil
 - o Mes collections
 - Mon historique de lecture
 - o Paramétrage profil
 - Auto-formation
 - Soumettre un texte
- Contact

Utilisateur sans compte arborescence : Entrée sur l'Accueil, avec pop-up *créer un compte*

Arborescence menu principal

- Accueil (le fil des idées)
- Le projet
 - o Historique du projet
 - o L'encodage en xml tei
 - o Le thésaurus
 - o Comité et partenaires (figurant aussi dans la page d'accueil...)
- Le corpus
 - o Rechercher un texte
- Profil
 - o Créer un compte / s'inscrire
- Contact

3.3. Points sur le RGPD

L'information doit être facilement accessible à l'utilisateur, claire, concise et transparente, pour faciliter la compréhension et afficher d'emblée les choix et les fonctionnalités consultables et paramétrables par ce dernier. Localiser les informations ou données accessibles par l'utilisateur est important dans le cas d'un traitement de celle-ci par le site, comme cela sera le cas sur la plateforme Savoirs. La concision et la lisibilité devront donc être envisagées de manière attentive ; dans le design ergonomique du site - au moins dans les pages qui relèvent des données utilisateurs et de paramétrage.

Dans le cas de création d'un compte sur la plateforme Savoirs, plusieurs champs d'informations utilisateur pourront être récoltés par le site, lors du questionnaire d'inscription ou de la modification du profil utilisateur. Il conviendra de ne recueillir que les champs nécessaires au fonctionnement du site et de ses fonctionnalités principales. Dans le cas où certaines données recueillies permettraient l'identification, il faudra l'afficher clairement et permettre de refuser ces cookies ou requêtes spécifiques.

Une **donnée personnelle** se rapportant à tout champ permettant l'identification d'une personne physique, la plateforme Savoirs doit justifier le cas échéant, le besoin qu'elle aurait d'accéder à un tel type de données. Dans le cadre de la création du profil et du questionnaire, nous avons proposé le champ de pseudonyme ou d'identifiant unique en chiffre, pour permettre de masquer l'IP de l'utilisateur si c'est le choix qui est fait. Le nom et le prénom ou l'ORCID (identifiant chercheur) ne seront pas recueillis dans le cas d'un utilisateur lambda, qui ne se connecte pas. Mais la question se pose dans le cas où un utilisateur crée son compte. Il faudra ainsi lui laisser le choix de ne pas faire apparaître ces données. L'adresse e-mail néanmoins, a priori requise pour la création du compte, pourra elle suffire à l'identification, si elle suit le schéma « prenom.nom@... » ? Cela pose question quant aux éléments d'identification abordés. Il faudra que les engagements de la plateforme à cet égard soient clairs et que ces données ne soient pas exploitées dans un autre cadre que le maintien du compte ou le paramétrage de notifications : pas de partage à des entreprises/cookies tiers, ni commerciaux si ce n'est pas défini et dûment accepté par l'utilisateur dans les conditions d'utilisation et de confidentialité.

Si « les données résultant d'une pseudonymisation sont considérées comme des données personnelles et restent donc soumises aux obligations du RGPD », la CNIL considère néanmoins cette approche comme réductrice des risques et contribuant à la conformité du RGPD. L'anonymisation des données personnelles est la voie que nous conseillons d'envisager, de manière générale et dans le cas où Savoirs n'est pas encore utilisé comme outil de collaboration et de citation directe entre chercheurs.

Concernant le **consentement**, celui-ci doit être recueilli pour chaque utilisateur et concerner 4 critères « cumulatifs ». Il doit être :

- **Libre** : le consentement ne doit pas être contraint ni influencé. La personne doit se voir offrir un choix réel, sans avoir à subir de conséquences négatives en cas de refus.
- **Spécifique** : un consentement doit correspondre à un seul traitement, pour une finalité déterminée.
- **Éclairé** : pour qu'il soit valide, le consentement doit être accompagné d'un certain nombre d'informations communiquées à la personne avant qu'elle ne consente.
- **Univoque** : le consentement doit être donné par une déclaration ou tout autre acte positif clair. Aucune ambiguïté quant à l'expression du consentement ne peut demeurer.

Un exemple de notice de RGPD qui peut être installée sur le site et les différentes étapes d'installation (documentation/read me file) :

Installation
Add this line to your application's Gemfile:

```
gem 'gdpr'
```


And then execute:

```
$ bundle
```


Or install it yourself as:

```
$ gem install gdpr
```


Add this to layout:

```
<%= render 'gdpr/cookie_consent' %>
```


Add this to javascripts:

```
//= require gdpr/cookie_consent
```


Add this to stylesheets:

```
@import 'gdpr/cookie_consent'
```


Set the privacy policy url in the locales.

Un exemple des points nécessaires pour la mise en, place d'un registre de traitement unique de données (source : CNIL) :

- le cas échéant, **le nom et les coordonnées** du [responsable conjoint du traitement](#) mis en œuvre
- les **finalités** du traitement, l'objectif en vue duquel ces données ont été collectées
- les catégories **de personnes concernées** (client, prospect, employé, etc.)
- les catégories de **données personnelles** (exemples : identité, situation familiale, économique ou financière, données bancaires, données de connexion, données de localisation, etc.)
- **les catégories de destinataires** auxquels les données à caractère personnel ont été ou seront communiquées, y compris les sous-traitants auxquels vous recourez
- les **transferts** de données à caractère personnel vers un pays tiers ou à une organisation internationale et, dans certains cas très particuliers, les garanties prévues pour ces transferts ;
- les **délais prévus pour l'effacement** des différentes catégories de données, c'est-à-dire la durée de conservation, ou à défaut les critères permettant de la déterminer
- dans la mesure du possible, une **description générale des mesures de sécurité** techniques et organisationnelles que vous mettez en œuvre.

Les cookies sont à envisager selon leur nature. Certains sont temporaires et non nécessaires : comme les cookies de partage des réseaux sociaux - ici Twitter - les cookies de session de navigation sur le site, pour se souvenir de ses préférences temporaires, ou des cookies de simple consentement pour les cookies du site web spécifique. Les cookies de préférences sont acceptables ou non, comme la possibilité de créer un compte, d'accepter tous les cookies pour avoir accès à toutes les fonctionnalités, comme indiqué plus haut. Les cookies nécessaires assurent les fonctionnalités basiques du site web et ne nécessitent pas de consentement, mais il peut cependant être utile et requis, comme le préconise la CNIL ; d'informer l'utilisateur.

4. Les livrables

Nous décrirons les différentes fonctionnalités ainsi que leurs possibilités d'implémentation avec la logique suivante : d'abord ce qui est susceptible de présider à la création du profil, à savoir le questionnaire et l'autoformation, ensuite le profil et tous les éléments qui dépendent de ce dernier et enfin, l'espace texte, qui existe indépendamment du profil mais dont certaines fonctionnalités dépendront de la création de ce dernier.

4.1. Le questionnaire et l'autoformation

4.1.1. Questionnaire

Dans la mesure où nous cherchons une certaine efficacité dans le paramétrage du profil et par l'utilisateur, il semble préférable d'éviter les questions « ouvertes » pour le questionnaire, qui nécessiteraient un traitement trop long ou trop complexe en aval, d'autant que les utilisateurs ne savent pas forcément eux-mêmes formuler clairement ce qu'ils cherchent en venant sur la plateforme. Ainsi, nous proposons, comme détaillé par la suite, un questionnaire qui permette à la fois une certaine finesse dans la constitution des profils utilisateurs, et une efficacité qui permette un paramétrage ergonomique et réalisable des algorithmes de recommandation. Pour cela, le questionnaire se présentera sous la forme d'une série de questions fermées : l'utilisateur devra cocher la ou les cases correspondant à ses réponses ; chaque réponse conditionnera la question qui lui sera proposée après, selon une arborescence détaillée plus loin. Le but est de constituer des profils de plus en plus détaillés et personnalisés pour les utilisateurs.

Voici les **principales fonctions que nous avons définies pour le questionnaire utilisateur** :

- faire un premier paramétrage du profil pour proposer des parcours susceptibles d'intéresser le lecteur en trouvant un équilibre entre ciblage et ouverture
- disciplinaire tout en gardant un objectif de sérendipité. Pour cela, il sera nécessaire de faire un aller – retour entre les attentes de l'utilisateur et les propositions de la plateforme. Par conséquent, il faudra cerner les attentes de l'utilisateur en lui demandant, par exemple, des informations sur ses centres d'intérêt.
- faire attention à ce que les questions portant sur l'identité ne soient ni limitantes en terme de paramétrage, ni trop intrusives, pour respecter les normes imposées par le RGPD : la dimension facultative de certaines questions dans le questionnaire nous paraît un début de réponse valable pour éviter ces deux écueils.

Nous précisons donc que certaines des propositions faites ci-dessous peuvent rencontrer des oppositions à ce niveau, et nous les mettons en italique pour les identifier clairement.

Précision sur la collecte des données d'identité : dans la mesure où on imagine pouvoir faire de Savoirs un outil de mise en commun universitaire, ou du moins de partage du savoir, il pourrait être pertinent de demander l'identité réelle de l'utilisateur (notamment pour les chercheurs souhaitant utiliser la plateforme à but professionnel, cela pourrait permettre de faciliter le dialogue entre chercheurs et entre laboratoires, de situer chacun ses interlocuteurs). De plus, dans l'hypothèse du développement de fonctionnalités spécifiquement orientées vers

la mise en commun des textes et des bibliothèques, avec un système éventuellement d'annotations partagées sur les textes (co-éditions de notes...), cela permettrait de pouvoir authentifier les "co-éditeurs", voire de pouvoir citer d'autres chercheurs.

Néanmoins, il faudrait se demander si ces données sur l'identité peuvent apporter une réelle plus-value ; si on trouve ces demandes trop intrusives, on peut choisir de demander par exemple, à la place ou en complément de l'identité réelle, l'ORCID du chercheur.

Il paraît important, **tout d'abord, d'identifier le statut professionnel** - le statut de « lecteur » ; à partir d'une première sélection :

Profil chercheur (celui que nous avons le plus développé, car c'est la catégorie dont nous connaissons le mieux, par nos profils et nos études, les problématiques de "base") : voir le schéma d'arborescence ébauché correspondant au "profil chercheur"

— *Est-ce un chercheur ? Si oui, est-il membre d'un laboratoire ? A-t-il déjà travaillé sur des projets interdisciplinaires ? Si oui, quelles disciplines étaient concernées ? Quelles périodes ? Quel est son champ disciplinaire de référence ? Souhaite-t-il constituer dès maintenant une « bibliothèque de base » à partir de corpus de références déjà utilisés pour ses travaux ? Ce serait faisable à n'importe quel moment à partir de compte utilisateur, mais pourrait constituer une base pour un premier paramétrage du profil utilisateur.*

-- Ensuite, pour cibler les intérêts du chercheur : sur ses motivations (on suppose qu'il vient non pour son « loisir » mais principalement pour ses travaux) ; vient-il pour contextualiser un thème à traiter (élargissement, ouverture, sérendipité) ? Ou plutôt pour approfondir un travail de recherche disciplinaire en ciblant son approche (précision) ? Ces réponses seront déterminantes pour la programmation de l'algorithme et le paramétrage du profil. Là encore, il sera possible pour l'utilisateur de modifier ces préférences tout au long de son utilisation de Savoirs.

-- Il paraît important de ne pas négliger la dimension « ludique » ou du moins le principe d'« ouverture » sur lequel repose Savoirs, même si le chercheur choisit la réponse « approfondissement disciplinaire ». Pour cela, nous pourrions proposer une série de questions à part : quels sont ses goûts, ses loisirs, ses centres d'intérêt (périodes, thèmes, autres disciplines ; peut-être même personnages historiques, légendaires, fictionnels...). Une autre question dans cette série pourrait porter sur ses compétences : quelles langues comprend-t-il (il n'est pas forcément nécessaire de savoir parler couramment une langue pour lire des articles dans cette langue) ? A-t-il des notions dans une autre discipline ou un autre domaine (informatique, double-formation, pratique artistique) ? Ces questions pourraient permettre un ciblage à la fois plus fin et plus « ouvert » pour le profil, car ces compétences individuelles, sans nécessairement être reconnues professionnellement, sont transversales et peuvent enrichir l'approche et le savoir produit à terme par le chercheur.

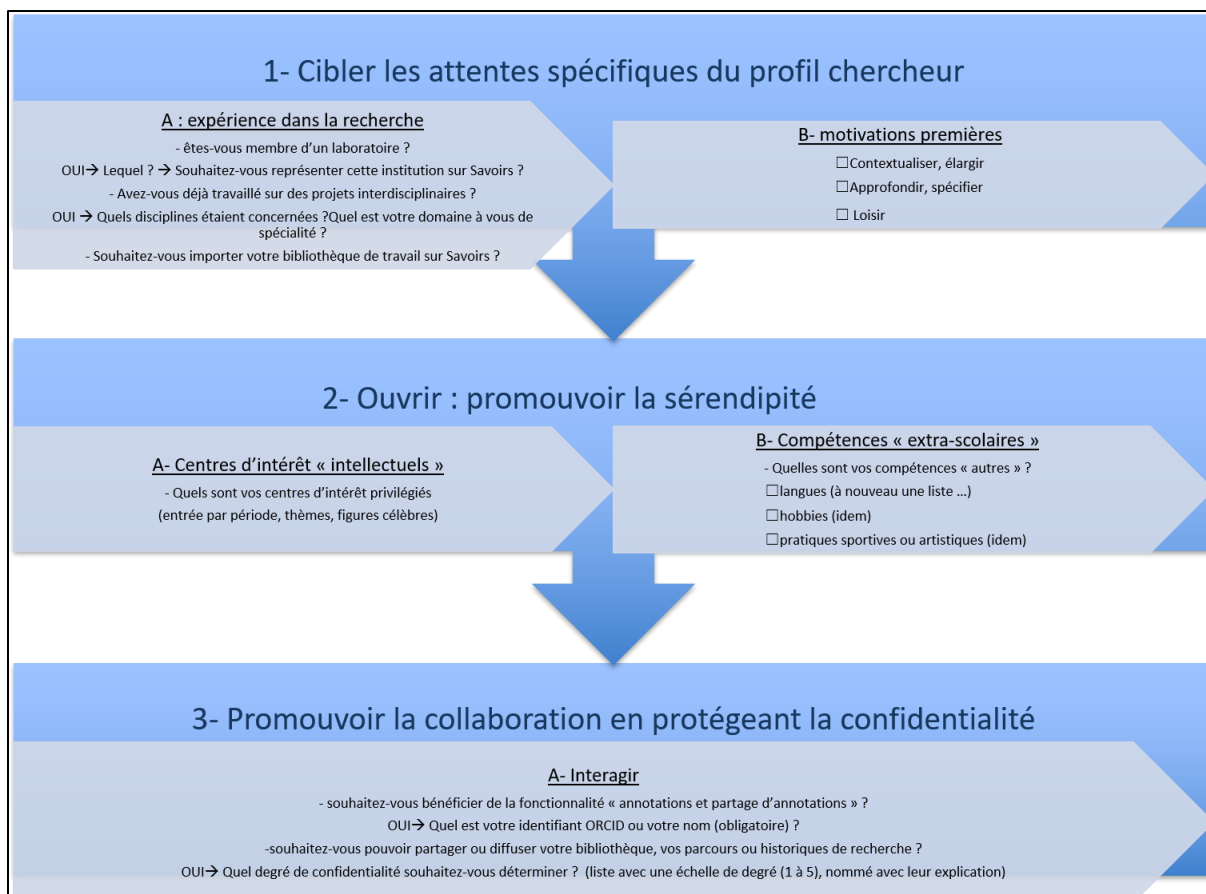


Figure 5 : Arborescence du questionnaire pour le profil "chercheur"

Profil étudiant

— Est-ce un étudiant ? si oui, vient-il initialement pour effectuer des recherches universitaires (mémoire, articles) ? Si oui, est-il familier avec les outils, le fonctionnement de la recherche, ou débute-t-il dans cette démarche ? Ces questions seraient pertinentes dans la mesure où une partie de l'auto-formation pourrait être adaptée en fonction : si le site établit des passerelles avec des outils bibliographiques extérieurs, par exemple. Sinon, vient-il pour compléter, approfondir un cours ? Par simple curiosité, ou pour s'informer sur un de ses loisirs ?

Profil autre

— Est-ce simplement un utilisateur qui souhaite utiliser Savoirs pour son loisir, ou à d'autres fins que la recherche (enrichissement personnel, se cultiver ...) ? Dans ce cas, il nous semble pertinent de pouvoir proposer des parcours différents, des textes peut-être plus ludiques, des œuvres ou extraits romanesques etc. Les parcours proposés reposeraient davantage sur des liens thématiques ou génériques. A nouveau, on pourrait demander à l'utilisateur ses goûts (époques...). On pourrait éventuellement demander aussi à ce type d'utilisateur combien de temps il pense vouloir consacrer à l'exploration de corpus sur Savoirs.

Voilà, pour résumer, ce que nous proposons comme **déroulé général du questionnaire** :

- 1- Une entrée par le profil « statut » : étudiant (il faudrait faire la distinction secondaire / université), chercheur ayant déjà une pratique des plateformes de recherche et une pratique professionnelle de la recherche, lecteur désirant explorer et se cultiver
- 2- Puis, entrée par les objectifs a priori recherchés par l'utilisateur : préciser, explorer, élargir, lire (tout simplement) ? Ces objectifs seront évidemment pris en compte mais relativisés dans l'algorithme, le principe de Savoirs restant l' « ouverture »
- 3- Ensuite, les goûts
- 4- Après, l'expérience (dans la manipulation des outils de recherche, dans la démarche de recherche)
- 5- Enfin, les compétences / Savoirs divers et pas forcément « disciplinaires » de l'utilisateur : compétences informatiques, langues maîtrisées, passions pratiquées etc.

4.1.2. L'autoformation

Une explication en plusieurs temps / étapes

L'autoformation présentera les principes de base de la plateforme avant de faire remplir le questionnaire, avec la possibilité d'avoir une formation plus approfondie une fois celui-ci complété. En effet, l'utilisateur ne sachant pas nécessairement ce qu'il vient chercher lors de son arrivée sur la plateforme, du fait notamment de la diversité des profils potentiellement attirés par Savoirs (chercheurs, mais aussi adolescents, ou amateurs), la présentation des fonctionnalités proposées par Savoirs peut déjà l'aider à cerner ses propres attentes, faire naître des idées qui serviront pour le questionnaire. Cette première autoformation, comme détaillée plus bas, ne serait qu'une brève entrée en matière visant à la fois à informer l'utilisateur et à le faire entrer dans la démarche originale de Savoirs.

L'autoformation pourrait prendre plusieurs formes et intervenir à différents moments pour l'utilisateur ; il serait possible de la faire en deux temps, pour plusieurs raisons : premièrement, cela permettrait de ne pas « surcharger » l'utilisateur en termes d'informations, ensuite, le fait d'en faire une partie avant, et une partie après le questionnaire utilisateur permettrait de « cibler » les autoformations proposées en fonction du profil de l'utilisateurs. (Cela sera développé dans la partie « Questionnaire »).

L'objectif général est de permettre et de faciliter l'appropriation de la plateforme de Savoirs par des utilisateurs aux profils divers. D'où, plus précisément :

1- La présentation brève, synthétique du projet Savoirs, pour tous les utilisateurs : cette partie doit être imagée, colorée, peut-être sous forme dynamique (gifs, courte vidéo), l'idée étant de susciter l'adhésion de l'utilisateur au projet, à la démarche d'exploration, d'ouverture aux savoirs. Le but est de faire entrer tous types d'utilisateurs dans cette démarche qui se veut singulière et interdisciplinaire. Il serait possible d'utiliser, au cours de cette présentation, des formules marquantes pour évoquer le projet, comme « construire une cartographie du savoir / de votre savoir etc. avec le système de parcours de lecture, de balisage, d'exploration... ».

2- Accompagnement pour la création du compte et le paramétrage profil :

Cette partie pourrait prendre la simple forme d'annotation accompagnant le remplissage du questionnaire ou le paramétrage des aspects « formels » du compte (police, couleur...), en écho à la partie de paramétrage profil.

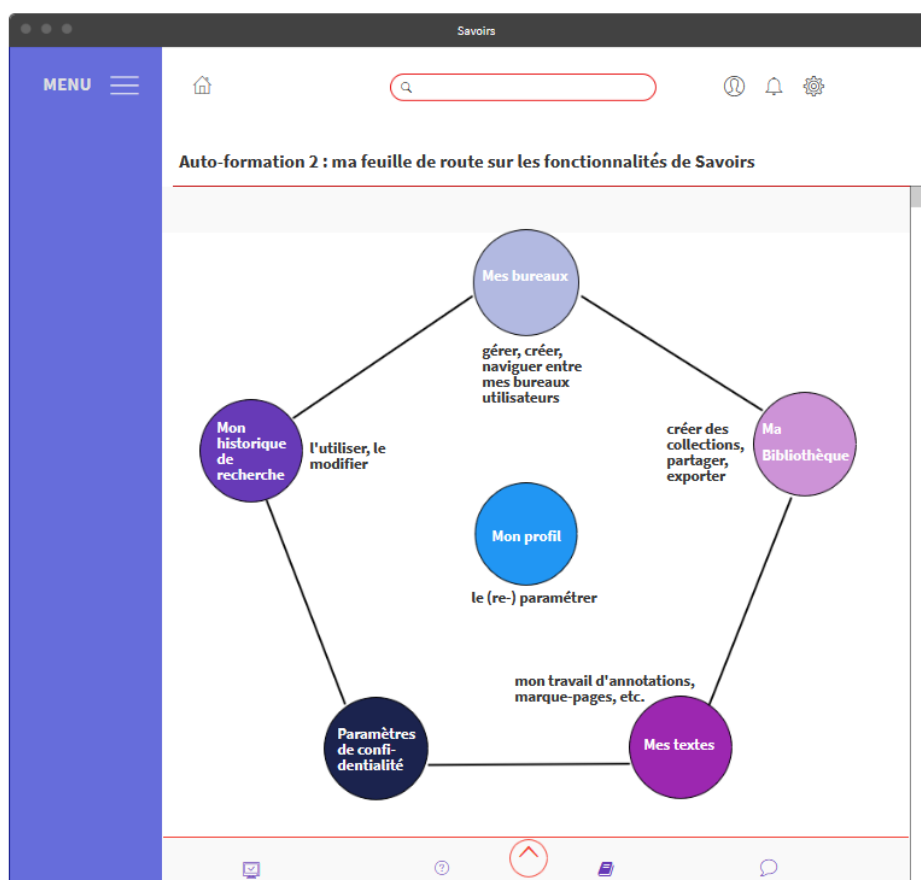
3- Une présentation des fonctionnalités spécifiques à la plateforme Savoirs :

Cette partie se présenterait sous forme de carte mentale (colorée), accessible à tout moment dans la barre de navigation de la plateforme, et présentée immédiatement après le questionnaire. L'utilisateur serait amené à explorer les différentes fonctionnalités avant de se lancer dans la recherche de corpus ou même la création d'un premier bureau. A partir de la carte mentale, qui permettrait de voir d'emblée les différentes fonctionnalités offertes par Savoirs, l'utilisateur clique sur telle ou telle fonctionnalité pour avoir un exposé détaillé de la démarche à suivre. L'utilisateur aurait ainsi accès aux entrées intitulées (à simplifier éventuellement en raisons de la place disponible sur la page : l'intérêt de la carte mentale étant la clarté) :

- comment modifier mes préférences de profil utilisateur ?
- avoir accès et utiliser mon historique de lecture
- créer un autre bureau / naviguer entre mes bureaux
- exporter mes recherches
- sur un texte, comment annoter / surligner / mettre un marque-page / suggérer un lien hypertexte ou un ajout de texte ?
- partager ma bibliothèque - mes textes travaillés
- gérer la confidentialité de mes recherches / de ma bibliothèque

Il serait souhaitable que l'utilisateur ait accès de nouveau aux explications guidées sur la partie « annoter/ surligner/ suggérer un ajout de texte » lors de sa lecture des textes eux-mêmes :

cela pourrait prendre la forme d'icônes sur le côté ou le bas de l'écran lorsque l'utilisateur lit un texte.



4.2. Le profil

4.2.1. Le paramétrage du profil

Le paramétrage du profil utilisateur sera accessible suite au remplissage du questionnaire, qui inscrira directement les différentes informations de l'utilisateur dans son profil. Ce dernier aura toutefois accès à son profil et pourra modifier, corriger ou supprimer ses informations. Certaines seront obligatoires à remplir pour assurer le bon fonctionnement de certaines fonctionnalités du site ou tout simplement la maintenance du profil utilisateur. Il sera évidemment possible de changer ses préférences et de retirer son consentement concernant toutes les données saisies, ainsi qu'avoir un contrôle direct sur celles-ci et le/les traitements qui leur sont appliqués. L'utilisateur pourra ainsi, depuis son profil, accéder à ses données et procéder à des demandes concernant ses dernières, ou se rendre compte de la manière dont elles sont utilisées. Les impératifs d'information, de lisibilité, d'accès et de modification facile, voulus par le RGPD seront pris en compte dans l'ergonomie et les possibilités de paramétrage.

Le screenshot présente l'interface utilisateur de la page 'Mon profil'. À gauche, une barre latérale bleue contient le logo 'Savoirs'. Le titre principal de la page est 'Mon profil', avec des liens de langue 'EN | FR | IT' et des boutons 'Mode sombre' et 'Mode clair'. La section 'Mes collections' permet d'ajouter de nouvelles collections à la bibliothèque, avec des boutons 'Asie', 'Afrique' et un bouton '+'. La section 'Mes informations personnelles' contient des champs pour 'Changer mon adresse mail', 'Changer mon mot de passe', 'Confirmation de mot de passe', 'Domaine de recherche*', 'Intérêts multidisciplinaires*', et un lien 'Mot de passe oublié ?'. La section 'Mes paramètres' inclut un slider pour la 'Taille de police', un 'Historique', et des boutons 'Utiliser l'auto-formation', 'Retirer le questionnaire', 'Effacer l'historique', et 'Supprimer mon compte'. Une section 'Envoyer une suggestion de texte' avec un bouton 'Envoyer' est également présente. Le pied de page contient des liens de navigation, des services numériques de l'EHESS, des coordonnées de contact, et des logos de l'École des Hautes Études en Sciences Sociales et PSL.

Figure 6 : 1ère maquette du profil

Paramétrage du profil utilisateur

Informations personnelles

Nom d'utilisateur/pseudo*

Nom Prénom

ORCID

E-mail*

Mot de passe* Confirmation du mot de passe Mot de passe oublié ?

Domaine de recherche

Intérêts multidisciplinaires

* champ obligatoire

Personnalisation

Préférences d'affichage : ☐ Clair ☐ Sombre

Taille de police : 10 Texte témoin

Auto-formation : Désactiver l'auto-formation Lancer le questionnaire

Historique : Effacer l'historique récent

Recherche préférentielle par : ☐ Titres ☐ Textes ☐ Lieux ☐ Dates

Vers les collections Autoformation Explorer les corpus de SAVOIRS Suggestion de texte

© 2020 Savoirs - EHESS - École des hautes études en sciences sociales Design : Sarah Meunier Construit avec : Laravel, Vue.js et Tailwind

Figure 7 : 2e maquette du profil : détaillée et avec indices RGPD latéraux et en rouge

4.2.2. Les collections

Dans le cadre d'un utilisateur enregistré, il sera amené à se créer des collections. Ces dernières sont pensées comme des petites bibliothèques permettant de faire le lien entre bibliothèque et Savoirs. Concrètement, cela lui permettrait de créer différents environnements de travail. L'intérêt est d'avoir des recommandations propres pour chaque collection, notamment dans le cas où un utilisateur travaille en parallèle sur plusieurs projets dont les thématiques sont différentes.

Au sein de cet espace, il pourra aussi, au centre de la page, modifier ses paramètres de recommandations en mettant ses préférences en termes de concepts, lieux et périodes. L'utilisateur aura également la possibilité de retrouver tous ses articles favoris qu'il pourra organiser par dossier dans un volet à gauche consacré à la collection. C'est aussi ici qu'il aura la capacité de consulter son historique dans le volet droit consacré à toutes les collections, soit par une recherche simple soit en utilisant le calendrier.

De plus, il pourra, dans une recherche avancée, représenter spatialement son historique ou bien exporter ce dernier. D'ailleurs, l'export s'étendra aux favoris dans le volet central ainsi qu'aux paramètres de recherche. Enfin, l'utilisateur disposera d'un bouton « home » pour retourner à la page d'accueil et des boutons servant à sélectionner diverses langues, à savoir le français, l'anglais et l'italien.

En matière de développement, il s'agit concrètement d'une page supplémentaire en HTML créée dès que l'utilisateur ajoute une collection. Les données générées seront stockées dans un fichier JSON. Le contenu des collections - à savoir les articles favoris - seront accessibles par des liens qui auront été créés lorsque l'utilisateur lisant l'article aura au préalable cliqué sur le bouton "Ajouter aux favoris" et choisi la collection souhaitée. Ce choix s'effectuera depuis le volet situé en bas de la page de l'article.

The screenshot displays the 'Savoirs' web application interface. The main header shows the 'Savoirs' logo, a home icon, the title 'Asie', and language options (EN | FR | IT). The interface is divided into three main vertical sections:

- Left Sidebar (Dark Blue):**
 - Dans cette collection:** A section for the current collection.
 - Mes favoris:** A list of favorite items, including 'Chine', 'Hérédité', and 'la transmission par l'hérédité'.
 - Nouveau dossier:** A button to create a new folder.
- Central Panel (Light Gray):**
 - Mes paramètres de recherche:** Search parameters section.
 - THÉSAURUS:** A dropdown menu labeled 'Sélectionner des concepts dans le thésaurus'.
 - LIEUX:** A dropdown menu labeled 'Sélectionner des lieux'.
 - DATES:** Date range selection with input fields for '1700' and '1877'.
 - Visualizations:** A bar chart showing data distribution over time, with a timeline slider below it.
 - Exporter mes favoris:** A section with a checkbox 'Exporter avec mes paramètres de recherche' and four buttons: CSV, XLSX, TXT, and RDF.
- Right Sidebar (Dark Blue):**
 - Dans toutes les collections:** A section for all collections.
 - Mon historique de recherche:** A section for the user's research history, featuring a calendar view for 'January 2021' and a 'Recherche dans mon historique' search bar.
 - Courbe spatiale de l'historique:** A spatial curve visualization showing the historical path of research.
 - Exporter mon historique:** A section with four buttons: CSV, XLSX, TXT, and RDF.

The bottom of the page contains a footer with the 'Savoirs' logo, contact information for 'Services numériques de l'EHESS', and logos for 'L'ÉCOLE DES HAUTES ÉTUDES DE LA SOCIÉTÉ' and 'PSL'.

Figure 8 : Interface utilisateur des collections, historique de recherche...

4.2.3. Suggestion de texte

La suggestion de texte permettra de fournir à l'équipe de Savoirs des suggestions pour des articles qui ne sont pas présents sur la plateforme et qu'il pourrait être intéressant d'ajouter à cette dernière. Cela pourra être fait dans la page profil pour un utilisateur enregistré – envisageable également pour un utilisateur non enregistré - avec un simple bouton permettant d'ouvrir une fenêtre pop-up. Dans cette fenêtre, on aurait la possibilité de mettre les différentes informations liées à un article sous la forme d'*inputs* que l'utilisateur peut remplir avec les éléments dont il dispose. Un bon exemple en termes de formats et d'inputs serait celui de Zotero. Ces informations pourraient ensuite être stockées dans un fichier JSON où l'équipe de Savoirs pourrait venir les récupérer et décider si elle donne suite ou non à la demande. Les critères d'acceptation ou de refus d'une demande seront à déterminer par l'équipe Savoirs.

4.2.4. Historique des parcours de lecture

4.2.4.1. *Le fil d'Ariane*

Nous avons réfléchi à plusieurs manières de représenter l'historique de recherche de l'utilisateur lorsqu'il disposera d'un compte.

Le fil d'ariane permettra à l'utilisateur d'avoir une vue en temps réel sur son parcours, et de pouvoir revenir d'un clic à l'étape précédente. La question est l'intégration de ce fil d'Ariane dans la navigation et sa nature exacte. En effet, s'agira-t-il d'un fil qui donnera l'historique et l'étape de chaque article lu précédemment - dans la lecture active, pas lors de nouvelles lectures, sauf dans un cas de reprise d'un parcours précédent. Ou alors, un fil qui laissera revenir à chaque article, au-delà des simples pages et catégories du site à la manière d'un fil d'Ariane classique.

Historique de recherche/navigation : en temps réel, fil d'Ariane

- Fil d'ariane classique

Maison > Présentation Savoirs

- Fil d'ariane d'articles (navigation articles précédents) en temps réel

L'accès au Savoir dans les grandes bibliothèques > Les Lieux de Savoir

- Fil d'Ariane thématique (par date/lieu...)

1452 > 1320 ou Concepts

- Un fil d'ariane sous format dynamique : de petits points ou ronds rouges, un peu comme le format de menu latéral droit utilisé dans le design actuel de la plateforme Savoirs. Affichés horizontalement au-dessus de l'article lu, le survol des petits points fera apparaître le titre de l'article lu actuellement et en allant vers la gauche, des articles précédents.

4.2.4.2. Accès à l'historique de recherche et sa visualisation

Proposer un affichage linéaire de l'historique n'est pas a priori la solution que nous avons retenue, étant donné qu'elle nous paraît moins adaptée pour la plateforme Savoirs. Nous pouvons néanmoins évoquer cette idée, qui reprendrait alors la forme classique des affichages d'historique dans le navigateur par exemple.

Nous proposons également une possibilité de recherche directement implémentée dans les parcours d'historiques. Celle-ci pourra permettre à l'utilisateur de voir s'il a, durant l'un de ses parcours, déjà exploré une thématique et d'y revenir pour poursuivre ses lectures sur cette dernière, à partir d'un point déjà visité s'il le souhaite.

La visualisation de l'historique par témoins de dates nous paraît assez intéressante, puisqu'elle permet clairement à l'utilisateur de retrouver ces recherches, sur un affichage calendaire et de retourner plus facilement « dans le temps », pour y avoir accès. C'est donc notamment sous la forme d'un calendrier assez classique, avec les dates d'activités sur la plateforme en surbrillance, que nous proposons d'animer cette proposition.

La visualisation en courbe temporelle ou *time curve* pourrait également convenir à Savoirs, puisqu'elle représente de manière assez claire, visuelle, les parcours de l'utilisateur. Savoirs étant une plateforme qui enjoint à l'exploration, à la sérendipité et au dépassement des « frontières » disciplinaires, nous pensons que cela se traduira de manière intéressante dans ces courbes. L'intérêt d'utiliser des visualisations de ce type est de pouvoir insérer plusieurs types de données dans la navigation ou l'historique de recherche. Les recherches de l'utilisateur sont initialement séparées par des intervalles temporels et ce, sur un axe droit à une dimension. La boucle temporelle permet de dépasser cette unique dimension en prenant en compte la proximité entre deux points/textes et en les rapprochant pour former une courbe, qui correspond à l'historique de recherche. Celle-ci illustre l'évolution du parcours de lecture ou de recherche et permet de relever visuellement et facilement les points d'attention, de rupture ou de changement au cours d'un parcours.

Nous nous posons également la question de la possibilité d'une autre forme de visualisation suivant ce principe, mais avec d'autres variables que les variables temporelles : à savoir les concepts du Thésaurus de Savoirs ou les lieux. Cela permettra plus facilement à l'utilisateur de rapprocher et comparer ses parcours de lecture précédents et les différences ou rapprochements entre eux, pour pouvoir s'appuyer sur les ressources déjà explorées ou des selon d'autres critères qui indiquent la similarité entre eux. Cela pourra constituer des éléments parties des algorithmes de Savoirs : les concepts ou les lieux sont particulièrement intéressants pour l'utilisateur lorsqu'il cherche à reprendre des articles consultés antérieurement et à retrouver des textes analogues à ceux qu'il aurait entrevus.

L'essentiel, du côté de l'implémentation est de bien respecter les normes des données alimentant l'algorithme générant la *time curve*. Et ce, afin qu'elle s'affiche correctement et sans erreur, sous un format de dates « YYYY – MM – dd » et d'horaire « HH ; mm ; ss ; ms », le tout en JSON. La question principale se posera sur la possibilité d'introduction de cet algorithme directement dans la plateforme Savoirs.

4.2.5. L'export de la recherche

Savoirs intégrerait deux fonctions d'exportation de données : la première est celle de l'exportation de l'historique général de l'utilisateur. Cette fonction située dans le volet droit dédié aux fonctions s'appliquant sur toutes les collections confondues permet à l'utilisateur de récupérer une chronologie des articles consultés ; seraient exportés la date de consultation, le titre de l'article et son auteur.

La seconde fonction d'exportation est celle des favoris, située sur le corps de la page de chaque collection. Cette fonction permet cette fois à l'utilisateur d'exporter des informations détaillées sur le contenu de ladite collection, à savoir :

- le titre de l'article
- l'auteur
- l'année de parution
- les indications de numéro de page
- le volume ou numéro de la publication
- le titre de la publication
- l'éditeur
- le lieu d'édition
- la langue de l'article
- la date de consultation de l'article
- le DOI.

En somme, l'intégralité des données traditionnellement nécessaires à l'identification des articles dans une bibliographie. En guise de paramètre facultatif, il sera permis à l'utilisateur d'exporter les paramètres de chaque collection pour donner davantage de contexte aux articles favoris.

Dans les deux cas d'exportation - historique comme favoris - il sera nécessaire de prévoir plusieurs formats d'export : nous recommandons le CSV, TXT, XLSX et RDF. Nous avons choisi le RDF car c'est un format qui peut être utilisé notamment pour Zotero et que nous disposons d'un script qui permet de l'exporter.

De façon plus concrète, l'exportation se fera de la façon suivante : pour l'exportation de l'historique, il faudra un script Ruby qui permette de récupérer dans un fichier JSON les informations à exporter, puis de choisir le format que l'on souhaite exporter².

4.3. L'espace texte

Un espace métatexte a été imaginé sous la forme d'un volet rétractable situé sous le corps de texte, à l'image des volets latéraux déjà réalisés. Ce volet contiendra les différentes fonctionnalités liées au texte qui seront détaillées ci-dessous. Par conséquent, le code utilisé sera le même que celui imaginé pour les volets latéraux.

²Pour les exports au format RDF, nous nous appuyons sur ceci : <https://gist.github.com/ktym/3985701>.
Pour les exports en CSV <https://gist.github.com/ktym/1672112>.

4.3.1. Les marque-pages

Savoirs intégrerait deux systèmes de marque – pages : le premier s'apparenterait à un système de favoris. L'utilisateur pourrait mettre les textes de son choix en favoris et les textes sélectionnés composeraient différentes collections. L'utilisateur accéderait à ces dernières via son profil.

Le second, lui, aurait l'utilité d'un marque-page de type papier : si l'utilisateur stoppe sa lecture en cours de route, il peut marquer le point où il s'est arrêté. Même s'il se déconnecte de son profil, le marque-page ne disparaît pas. En revanche, dès l'instant où il reviendra sur le passage marqué, le marque-page disparaîtra automatiquement. Pour que cette dernière fonction s'active, il faut que l'utilisateur ait quitté la page une fois et y revienne. S'il marque un passage, poursuit sa lecture et décide de revenir sur la partie marquée, le marque-page ne disparaîtra pas. L'utilisateur pourra également décider de supprimer le marque-page si finalement il se rend compte qu'il n'en a pas besoin.

Pour l'implémentation des marque – pages type favoris, cette bibliothèque ³en Ruby pourrait être utilisée. Puis, ces marque-pages pourraient être récupérés et placés dans un fichier JSON et à terme, une base de données relationnelle. En ce qui concerne les marque-pages de type « papier », ce code⁴, utilisant uniquement des bibliothèques python, pourrait convenir.

4.3.2. Le surlignage

Cette fonctionnalité permettrait à l'utilisateur de mettre en couleur des éléments qu'il considère comme étant importants, auxquels il peut associer des thématiques. Chaque thématique serait associée à une couleur précise qui aurait été définie au préalable par l'utilisateur. Evidemment, l'utilisateur peut modifier les paramètres du surlignage à tout moment. La fonction surlignage apparaîtrait dans le volet du bas, sous forme de bouton sur lequel l'utilisateur cliquerait.

Un volet comprenant une palette de couleurs ainsi qu'un espace texte – un espace par thématique – apparaîtrait alors. Le nombre de surlignages thématiques par défaut serait limité à trois, mais en cliquant sur un bouton dédié inséré au volet, l'utilisateur pourrait ajouter des surlignages supplémentaires – en théorie, autant qu'il le souhaite. Un stylage du texte serait possible : police, taille de la police. Cette dernière option peut être intéressante pour des personnes souffrant de problèmes de vue.

En étendant les fonctionnalités liées au tag, il serait possible pour l'utilisateur, en cliquant sur sa liste de tags personnels, de retrouver tous les documents comportant le tag sélectionné. Malgré la redondance que cela engendrerait – les documents pourraient apparaître plusieurs fois sous différents tags – cela permettrait à l'utilisateur de retracer sa réflexion. De

³ https://www.rubydoc.info/gems/bookmark_system/0.2.0

⁴ <https://github.com/dusk125/sublime-betterbookmarks>

plus, cette fonctionnalité est susceptible d'être intéressante pour le recueil de données dans le cadre d'études scientifiques visant à analyser les parcours réflexifs des utilisateurs de Savoirs.

En matière de code, il serait intéressant d'utiliser cet exemple⁵ qui permet de souligner un élément sur une page web et de l'adapter pour le projet Savoirs. Ce dernier utilise dans la grande majorité du Javascript mais aussi du HTML/CSS qui sont des langages avec lesquels Ruby est compatible et qu'il peut aussi gérer. Cet exemple-ci⁶, qui utilise presque intégralement du Javascript, pourrait également être avantageux à étudier. Il s'agirait ensuite de récupérer ces surlignages pour les mettre dans des fichiers JSON dans un premier temps, puis lorsque le nombre d'utilisateurs sera trop important, il faudra envisager de les mettre dans une base de données relationnelles.

4.3.3. Définitions de concepts

Cette fonctionnalité s'inscrit dans la logique de sérendipité voulue par l'équipe de la plateforme Savoirs. L'utilisateur, en passant le curseur sur un mot, pourra avoir accès à la définition du concept. En cliquant dessus, il aura la possibilité de lire la définition complète. Par ailleurs, nous conseillons éventuellement de se référer à titre d'exemple au site *Trismegistos* dont la forme est assez représentative du rendu imaginé pour cette fonctionnalité, à savoir une définition apparaissant au survol d'un mot grec⁷.

Au départ, un simple fichier JSON avec une définition et un lien vers une page Wikipédia – ou une autre plateforme Creative Commons pourrait suffire puisque le nombre d'utilisateurs sera limité. Cependant, à termes, il faudra envisager la création d'une base de données relationnelle qui sera remplie par une partie des membres de l'équipe – dans le souci d'avoir des définitions qui pourraient être jugées comme plus scientifiques et précises. Cela nécessitera, évidemment, une maintenance assez importante et une charge de travail supplémentaire pour l'équipe Savoirs.

4.3.4. L'annotation

Cette fonctionnalité a été imaginée de la manière suivante : l'utilisateur pourrait avoir le choix entre différents symboles pour ses annotations. Ce dernier apparaîtrait au-dessus du mot annoté ou du dernier mot du passage annoté. En cliquant dessus, il verrait alors le contenu de l'annotation. Il serait raisonnable de limiter le nombre de symboles disponibles à cinq : un cercle, un carré, un triangle, une étoile et un point d'interrogation encadré.

L'annotation se différencie du commentaire car elle est rattachée uniquement à certaines parties du texte (terme, passage, etc.). Le commentaire, quant à lui, concerne l'ensemble du texte. Pour l'aspect technique, en reprenant les remarques de compatibilité avec le Javascript et le HTML/CSS formulées pour la solution de surlignage, nous avons pu trouver un exemple⁸

⁵ <https://github.com/Jahn08/WEB-PAGE-HIGHLIGHTER>

⁶ <https://github.com/dossier/html-highlighter>

⁷ Voir ici : www.trismegistos.org/text/2

⁸ Voir ici : <https://javascript.developpez.com/faq/javascript/?page=fenetre>

détaillé et complet en javascript pour la création et la gestion des fenêtres de type “pop-up”. De la même façon, il s’agira de placer le contenu de ces annotations dans un fichier JSON dans le cadre d’une communauté d’utilisateurs réduite.

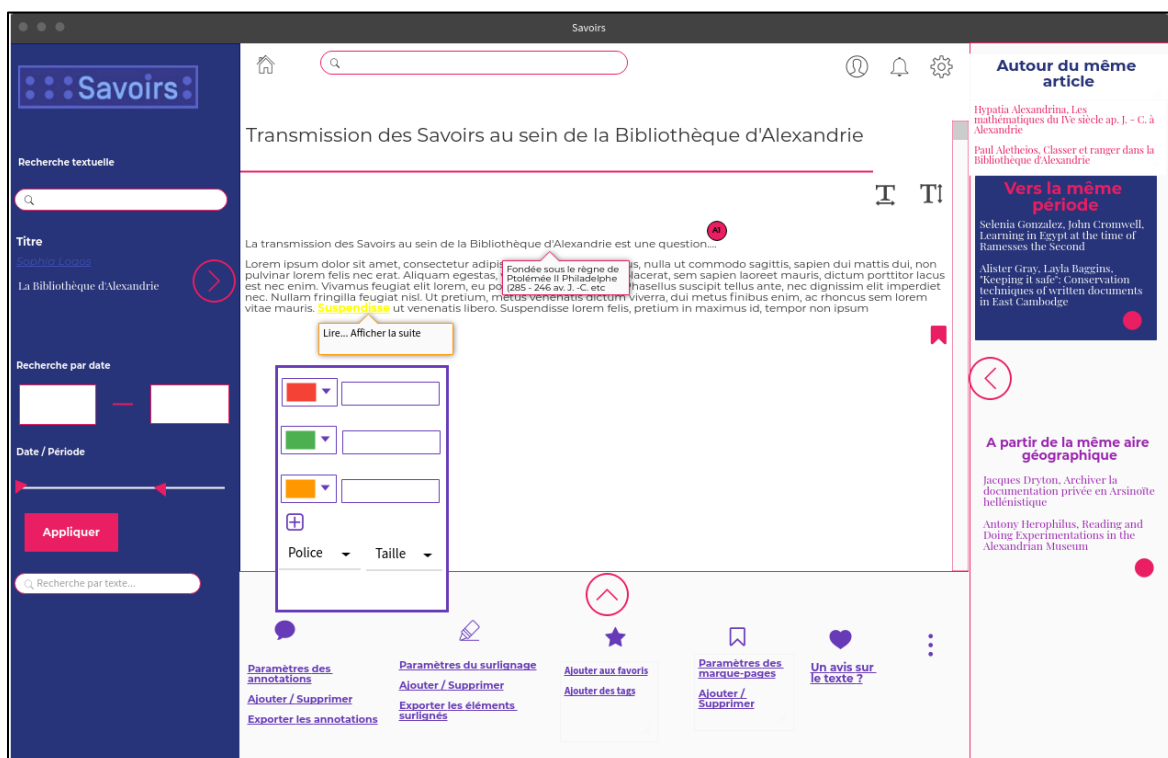


Figure 9 : Maquette de l'interface d'annotation, de surlignage...

5. Le calendrier de développement

5.1. Ordre à suivre pour le développement des fonctionnalités

En ce qui concerne la conception de l’espace profil, nous envisageons de procéder de la façon suivante, par ordre d’importance et dans une logique d’interdépendance de certaines fonctionnalités – certaines ne peuvent pas être implémentées si d’autres n’existent pas au préalable. Tout d’abord, il s’agirait de construire la partie profil, ensuite les collections et enfin les éléments liés aux textes. Dans la page profil, il faudrait d’abord développer les paramétrages profils, et relier le profil au fil d’Ariane. Par la suite, il s’avérerait nécessaire alors d’implémenter le questionnaire utilisateur.

Dans l’espace collections, il faudrait, en premier lieu, créer la possibilité de faire les différents bureaux en intégrant les boutons pour y avoir accès dans la page profil. Suite à cela, il serait envisageable de créer l’historique des parcours de lecture ainsi que la représentation spatiale. A partir de là, il serait réalisable d’embrayer sur l’implémentation de l’historique.

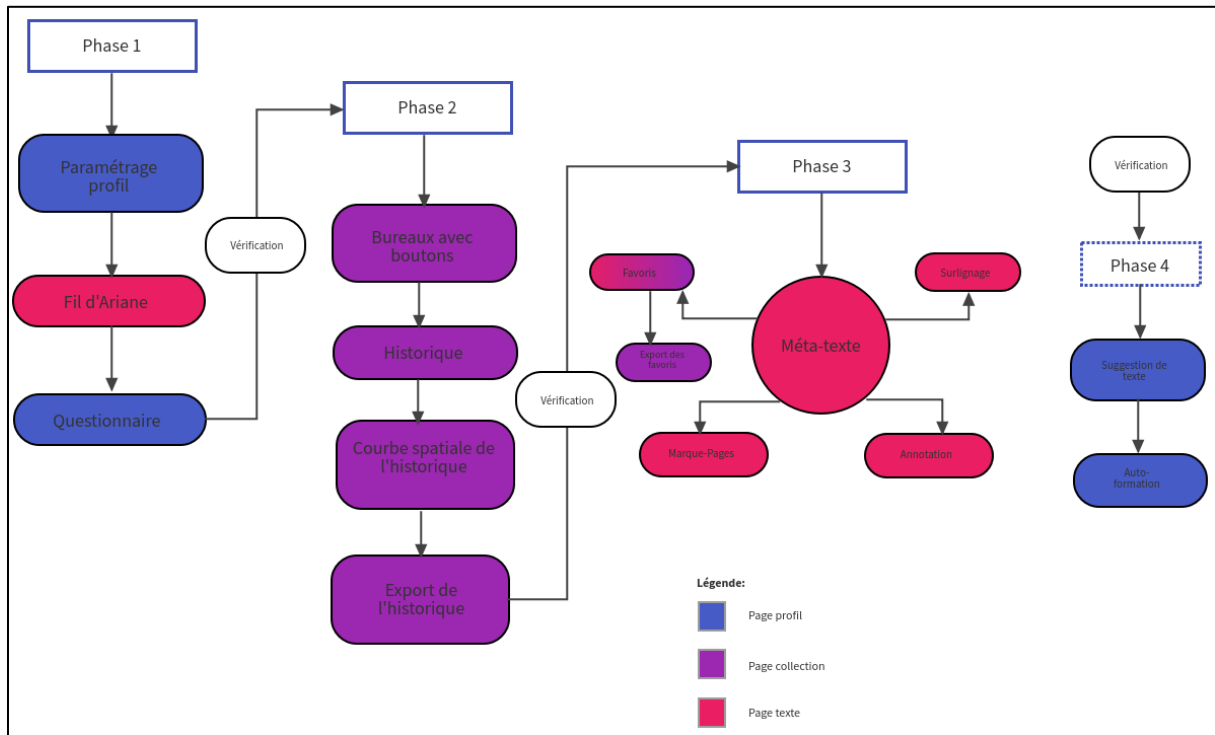
Pour le texte, la priorité est de créer l’espace métatexte puisque les fonctionnalités imaginées pour l’espace texte en dépendent. Les favoris, situés sur ce volet, devront être reliés aux collections puisque la création et la composition de ces dernières en dépendent directement.

La suggestion de texte, bien qu’elle dépende de l’espace profil, peut tout à fait être développée vers la fin, puisqu’il s’agit d’une fonctionnalité bien plus optionnelle que les autres.

Il est même possible d'envisager de la laisser de côté pour le moment afin de l'implémenter dans une phase future du projet.

L'autoformation, quant à elle, sera développée à la fin puisqu'il est impératif d'avoir l'ensemble des fonctionnalités pour qu'elle révèle toute son utilité. Evidemment, elle sera susceptible d'être modifiée au cours des étapes futures du développement de la plateforme Savoirs.

Figure 10 : Phases de développement de l'espace de profil



5.2. Le calendrier détaillé

Il nous est difficile définir un temps pour la création de l'intégration des fonctionnalités, mais nous suggérons une semaine par partie dans la conception, à laquelle nous pouvons ajouter une semaine supplémentaire afin de vérifier l'ensemble des fonctionnalités. Cela pourrait donc représenter un mois de développement continu.

Cette vérification pourrait être réalisée à l'issue de chaque partie. L'idéal serait de faire des tests avec plusieurs scénarios afin de voir comment la plateforme réagit. Ces scénarios seraient imaginés à la fois par le développeur et par l'équipe Savoirs ; ainsi, les deux parties pourraient déterminer si une tâche est réussie et éventuellement passer à la suivante, en prenant en compte, bien-sûr, que tout doit pouvoir fonctionner simultanément. Il serait également intéressant de faire tester la plateforme à des personnes extérieures, ayant différents niveaux informatiques et ayant des manières de travailler différentes, afin d'identifier d'éventuels bugs et de savoir ce qu'il faudrait améliorer en termes d'ergonomie.