Manuel d'utilisation de theses.fr

Version du 06-07-2011

Table des matières

1. Introduction à theses.fr 1
1.1. Présentation 1
1.2. Première version de l'interface web et développements à venir 1
1.3. Données disponibles et données à venir 2
2. L'interface www.theses.fr 3
2.1. Navigateurs conseillés 3
2.2. La page d'accueil 3
2.2.1. Le moteur de recherche 3
2.2.2. Explorer toute la base 6
2.2.3. S'abonner : les flux ATOM 6
2.2.4. Autres éléments de la page d'accueil 7
2.3. La page de recherche 7
2.3.1. Le moteur de recherche 7
2.3.2. Les facettes 10
2.3.3. Nombre de résultats et rappel des critères de recherche .16
2.3.4. La liste de résultats 16
2.3.5. Le bouton S'abonner : flux ATOM 20
2.3.6. Le bouton Partager : le module AddThis 20
2.4. La page d'une thèse 22
2.4.1. Le moteur de recherche 22
2.4.2. Les données décrivant la thèse 25
2.4.3. Accéder au texte intégral de la thèse 27
2.4.4. Comment citer la description de la thèse ? 28
2.4.5. Comment citer le document-thèse s'il est diffusé sur le web ? 28
2.4.6. Le bouton Partager 29
2.4.7. Signaler une erreur 30
2.4.8. Les boutons J'aime de Facebook et +1 de Google 31
2.4.9. Récupérer les références de la thèse dans Zotero 32
2.5. La page d'une personne 33
2.5.1. Le moteur de recherche 33
2.5.2. Le nuage de la personne 36
2.5.3. Les différents rôles de la personne et la liste des thèses liées 39
2.5.4. Le bouton Partager : le module AddThis 40

2.5.5. Signaler une erreur 42
2.6. La page d’un organisme 43
2.6.1. Le moteur de recherche 43
2.6.2. Le nuage de l’organisme 46
2.6.3. Les différents organismes et la liste des trente dernières thèses 48
2.6.4. Le bouton Partager 49
2.6.5. Signaler une erreur 51
3. Accéder aux données brutes de theses.fr : l’API XML 53
3.1. Introduction à l’API XML de theses.fr 53
3.2. Récupérer une liste de résultats 53
3.3. Récupérer les données d’une thèse 59
3.4. Récupérer les données d’une personne 60

Chapitre 1. Introduction à theses.fr

1.1. Présentation

theses.fr est le moteur de recherche des thèses de doctorat françaises, en préparation et soutenues, et des personnes et organismes en lien avec ces thèses.

theses.fr est une base de données dotée d'une interface web accessible à l'adresse http://www.theses.fr [http://www.theses.fr/] et d'une API [http://fr.wikipedia.org/wiki/Interface_de_programmation] XML.

1.2. Première version de l'interface web et développements à venir

La première version de theses.fr a ouvert en juillet 2011.

Nous vous encourageons à nous faire part des bugs et problèmes que vous constatez au fil de votre navigation, mais aussi de vos remarques et suggestions. Utilisez le guichet d'assistance ABESstp [https://stp.abes.fr/node/3?origine=thesesFr] accessible sur toutes les pages via le bouton 

L'interface web comprend les fonctionnalités de base à savoir :

- un moteur de recherche portant à la fois sur les données décrivant les thèses et sur le texte intégral des thèses
- une navigation à facettes
- une page par thèse
- une page par personne (sous réserve de l'identification univoque de la personne)
- une page par organisme (sous réserve de l'identification univoque de

l’organisme)

Les développements des prochains mois aboutiront à :

- un moteur de recherche dedié aux personnes
- un moteur de recherche dedié aux organismes
- la possibilité pour l’utilisateur d’accéder à un espace personnel lui offrant des services (sauvegarde des recherches, alertes..)

1.3. Données disponibles et données à venir

La première version de theses.fr ouverte en juillet 2011 donne accès aux thèses soutenues dans les établissements ayant choisi au moment de la soutenance d’abandonner le dépôt papier au profit de l’électronique.

Les données de theses.fr sont saisies et validées par les personnes des établissements de soutenance dans une application spécifique, STAR, interconnectée à theses.fr

Dans les prochains mois, les données relatives aux thèses en préparation seront versées dans theses.fr

A terme, theses.fr recensera toutes les thèses de doctorat actuellement en préparation et toutes les thèses soutenues en France depuis 1985. Des données relatives aux thèses plus anciennes seront également versées dans theses.fr

Chapitre 2. L’interface www.theses.fr

2.1. Navigateurs conseillés

L’interface de theses.fr est optimisée pour Firefox, Chrome, Safari et Opéra. Des lenteurs sont constatées avec Internet Exporer.

2.2. La page d’accueil

2.2.1. Le moteur de recherche

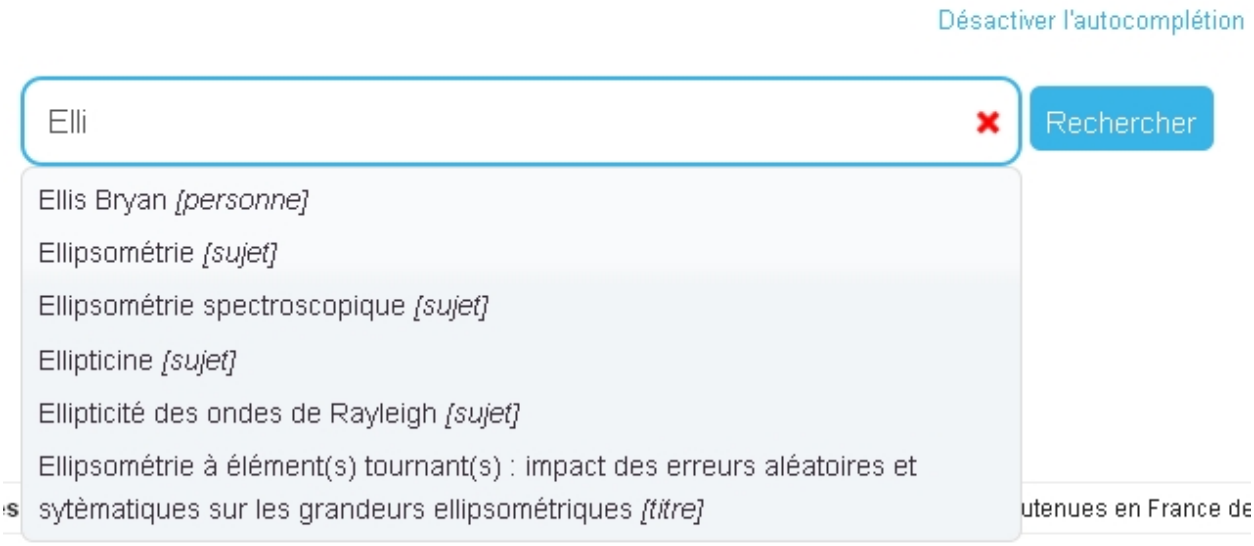
Sur quelles données porte la recherche ?

Le moteur de recherche de theses.fr porte sur les **informations de description de la thèse** (les noms et prénoms de l’auteur, du directeur de thèse, des membres du jury, les noms de l’établissement de soutenance, de l’école doctorale, des partenaires de recherche, le titre de la thèse, les résumés de la thèse, les mots clés…) **mais aussi le texte intégral de la thèse**.

L’auto complétion

Lorsque vous commencez à taper un mot, theses.fr vous suggère des réponses de trois types :

- des noms de personnes liées à des thèses
- des mots clés utilisés dans ces thèses
- des titres de thèse



Vous pouvez continuer votre saisie sans prêter attention à cette auto complétion ou bien sélectionnez avec la souris ou les flèches la suggestion qui vous intéresse.

L'auto complétion peut être désactivée par un clic sur le lien Désactiver l'autocomplétion et peut être réactivée par un clic sur Activer l'autocomplétion

Valider votre recherche

Vous pouvez soit appuyer sur la touche entrée de votre clavier soit cliquer sur le bouton Rechercher

Supprimer le terme recherché

Lorsque vous commencez à taper un caractère, une croix rouge apparait à droite de l'encart de recherche. Un clic sur cette croix rouge vide l'encart de recherche.



Recherche d'une expression exacte

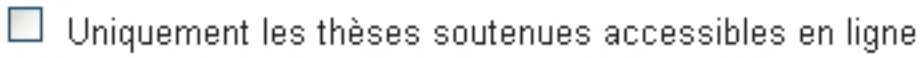
Si vous cherchez une expression exacte, vous devez la saisir entre guillemets.

Par exemple, si vous tapez droits fondamentaux, le moteur de recherche vous renverra toutes les thèses qui contiennent à la fois droits et fondamentaux.

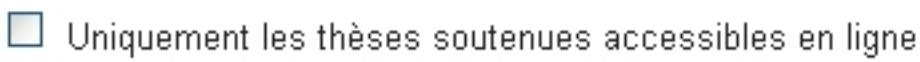
Si vous tapez « droits fondamentaux », vous aurez comme résultat toutes les thèses qui contiennent la chaine de caractère telle que vous l'avez écrite.

Le filtre « Uniquement les thèses soutenues accessibles en ligne »

Vous pouvez restreindre votre recherche aux seules thèses qui sont accessibles sur le web en cochant la case



Pour voir toutes les thèses soutenues accessibles en ligne, ne tapez rien dans l'encart de recherche, cochez la case



et cliquez sur



Mots vides, élision, lettres accentuées, casse et traitement de la chaine de caractères

Lorsque vous tapez une chaine de caractères dans le moteur de recherche, plusieurs traitements sont appliqués avant l'interrogation de la base de données.

Certains mots, appelés mots vides, ne sont pas pris en compte. Par exemple sont des mots vides : le, la les, ce, celle, de, des, un, une.. mais aussi et, ou,

à…

Un traitement des élisions est appliqué pour supprimer les l’, c’…

Les accents sont neutralisés, ainsi que la casse.

Enfin, tous les mots sont traités pour identifier leur racine et appliquer de fait une troncature permettant de chercher sur les formes plurielles ou féminines des mots.

Par exemple :

si vous tapez constitutionnalisation, la requête porte en fait sur constitutionnalis

si vous tapez classification phylogénétique, la requête porte en fait sur classif phylogenet

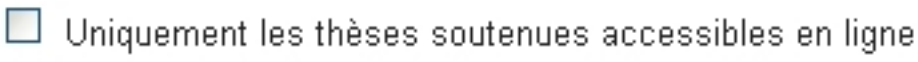
si vous tapez l’aéronautique, la requête porte en fait sur aeronaut

2.2.2. Explorer toute la base

Sans utiliser le moteur de recherche, vous pouvez voir le contenu de l’ensemble de la base de données de theses.fr en cliquant sur



Vous pouvez au préalable cocher le filtre



pour explorer les seules thèses accessibles en ligne.

2.2.3. S’abonner : les flux ATOM

Depuis la page d’accueil, vous pouvez vous abonner à plusieurs flux ATOM pour récupérer dans votre agrégateur

http://fr.wikipedia.org/wiki/Agr%C3%A9gateur] (Netvibes, Google reader…) les nouveautés de theses.fr

Vous serez automatiquement alerté de tout changement dans votre agrégateur sans avoir à refaire votre recherche dans theses.fr

Deux flux ATOM sont proposés :

les dernières thèses soutenues

les dernières thèses soutenues accessibles en ligne

2.2.4. Autres éléments de la page d'accueil

Au centre de la page d'accueil figure un bandeau de nouveautés, présent selon les annonces éventuelles.

La page d'accueil comprend également :

- en haut à droite, un accès à la documentation de theses.fr 
- en haut à droite, un lien vers le guichet d'assistance ABESstp [https://stp.abes.fr/node/3?origine=thesesFr] Assistance
- en bas à droite, une présentation de theses.fr derrière le lien à propos
- en bas à droite, une présentation des partenaires du projet accessible depuis partenaires

Ces éléments sont présents sur toutes les pages de theses.fr

2.3. La page de recherche

2.3.1. Le moteur de recherche

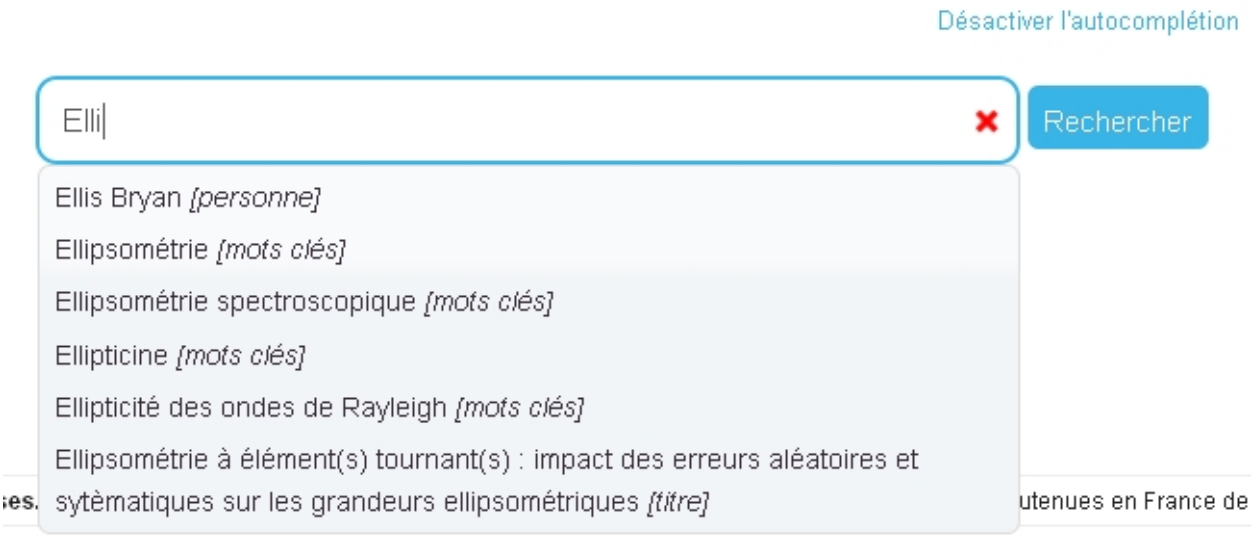
Sur quelles données porte la recherche ?

Le moteur de recherche de theses.fr porte sur les **informations de description de la thèse** (les noms et prénoms de l'auteur, du directeur de thèse, des membres du jury, les noms de l'établissement de soutenance, de l'école doctorale, des partenaires de recherche, le titre de la thèse, les résumés de la thèse, les mots clés…) **mais aussi le texte intégral de la thèse**.

L'auto complétion

Lorsque vous commencez à taper un mot, theses.fr vous suggère des réponses de trois types :

- des noms de personnes liées à des thèses
- des mots clés utilisés dans ces thèses
- des titres de thèse



Vous pouvez continuer votre saisie sans prêter attention à cette auto complétion ou bien sélectionnez avec la souris ou les flèches, la suggestion qui vous intéresse.

L'auto complétion peut être désactivée par un clic sur le lien Désactiver l'autocomplétion et peut être réactivée par un clic sur



Valider votre recherche

Vous pouvez soit appuyer sur la touche entrée de votre clavier soit cliquer sur le bouton 

Supprimer le terme recherché

Lorsque vous commencez à taper un caractère, une croix rouge apparait à droite de l'encart de recherche. Un clic sur cette croix rouge vide l'encart de recherche.



Recherche d'une expression exacte

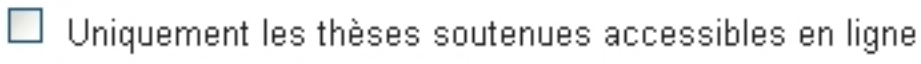
Si vous cherchez une expression exacte, vous devez la saisir entre guillemets.

Par exemple, si vous tapez droits fondamentaux, le moteur de recherche vous renverra toutes les thèses qui contiennent à la fois droits et fondamentaux.

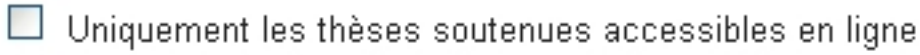
Si vous tapez « droits fondamentaux », vous aurez comme résultat toutes les thèses qui contiennent la chaine de caractère telle que vous l'avez écrite.

Le filtre « Uniquement les thèses soutenues accessibles en ligne »

Vous pouvez restreindre votre recherche aux seules thèses qui sont accessibles sur le web en cochant la case



Pour voir toutes les thèses soutenues accessibles en ligne, ne tapez rien dans l'encart de recherche, cochez la case



et cliquez sur 

Mots vides, élision, lettres accentuées, casse et traitement de la chaine de caractères

Lorsque vous tapez une chaine de caractères dans le moteur de recherche, plusieurs traitements sont appliqués avant l'interrogation de la base de données.

Certains mots, appelés mots vides, ne sont pas pris en compte. Par exemple sont des mots vides : le, la les, ce, celle, de, des, un, une.. mais aussi et, ou, à…

Un traitement des élisions est appliqué pour supprimer les l', c'…

Les accents sont neutralisés, ainsi que la casse.

Enfin, tous les mots sont traités pour identifier leur racine et appliquer de fait une troncature permettant de chercher sur les formes plurielles ou féminines des mots.

Par exemple :

si vous tapez constitutionnalisation, la requête porte en fait sur constitutionnalis

si vous tapez classification phylogénétique, la requête porte en fait sur classif phylogenet

si vous tapez l'aéronautique, la requête porte en fait sur aeronaut

2.3.2. Les facettes

Présentation des facettes

Dans le bandeau vertical situé à gauche de la page, sept facettes sont disponibles pour affiner une recherche.

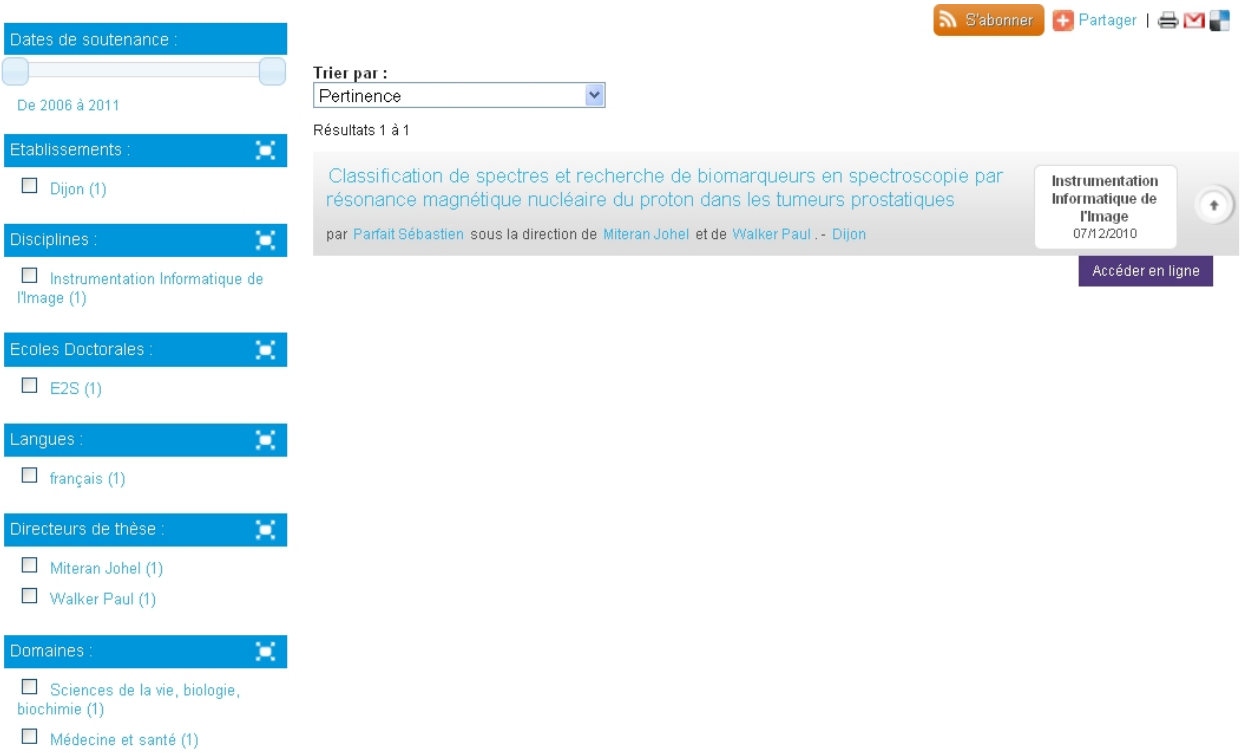
- les dates de soutenance
- les établissements de soutenance ou de cotutelle
- les disciplines du diplôme doctoral
- les écoles doctorales
- les langues du document-thèse
- les directeurs et codirecteurs de thèse
- les domaines scientifiques

Pour une thèse donnée, il existe toujours une date de soutenance, au moins un établissement (deux en cas de cotutelle), une discipline, une langue, au moins un directeur de thèse (plusieurs en cas de codirection), au moins un domaine scientifique. Par contre, la mention de l'école doctorale est facultative : cette facette peut être vide.

Quelle est la différence entre la discipline et le domaine ?

La discipline correspond à la spécialité qui figure sur le diplôme de doctorat alors que le domaine fait référence à une classification largement utilisée dans les bibliothèques universitaires (la classification décimale Dewey). Le libellé des disciplines est libre et infini alors que le libellé et le nombre de domaines est fixe (10 grands domaines incluant chacun 10 sous-domaines). Par conséquent, la discipline est souvent plus précise que le domaine ; le domaine permet des regroupements entre thèses traitant de sujets connexes.

Par exemple, la thèse www.theses.fr/2010DIJOS040 [http://www.theses.fr/2010DIJOS040] a pour discipline « Instrumentation Informatique de l'Image » et est rattachée à deux domaines « Sciences de la vie, biologie, biochimie » et « Médecine et santé »



Le curseur Dates de soutenance

Par défaut, le curseur sélectionne la période 2006-2011 car la base de données de theses.fr ne contient pas pour le moment de thèse antérieure à 2006.

Vous pouvez changer cette fourchette en cliquant-déplaçant le curseur de gauche puis le curseur de droite.



La fourchette sélectionnée apparait derrière la mention « Votre sélection » dans un tag. 

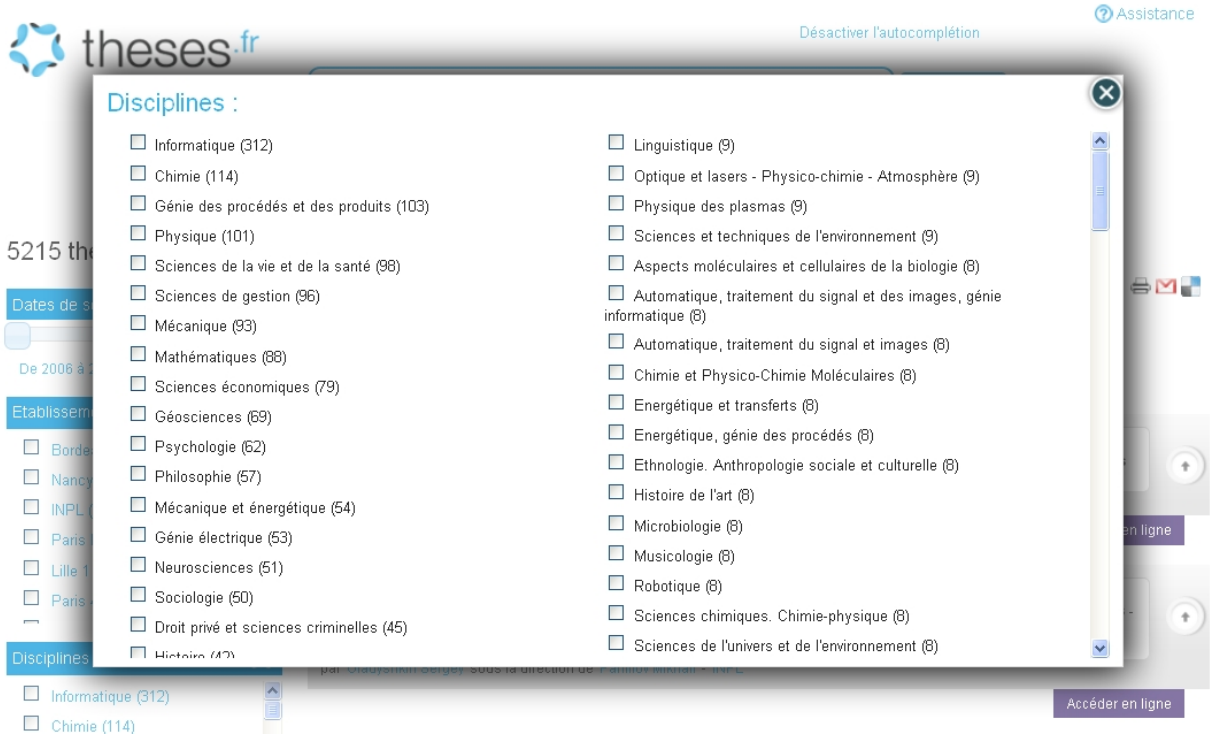
Le nombre de résultats et la liste de résultats se mettent automatiquement à jour.

Agrandir une facette, sélectionner/désélectionner les éléments des facettes

Par défaut, les facettes sont accessibles via le bandeau de gauche. Un ascenseur vertical permet de naviguer dans la liste des éléments d'une facette.

Il est également possible d'agrandir chaque facette en cliquant sur le bouton 

Les éléments de la facette s'ouvrent alors en surpression sur la page :



Pour fermer la facette ouverte en surimpression, il faut cliquer sur le bouton 

Pour sélectionner un ou plusieurs éléments d'une facette, cochez la case correspondante. L'élément coché apparait en haut de la page, à la suite de la

mention « Votre sélection » dans un tag. La liste de résultats et les autres facettes se mettent automatiquement à jour.

Pour désélectionner un ou plusieurs éléments d'une facette, vous pouvez :

soit supprimer le tag en cliquant sur la croix,

soit décocher les cases dans le bandeau de gauche,

soit ouvrir la facette en surimpression, décochez les cases et refermez la surimpression.

Il n'est pas possible de sélectionner ou désélectionner tous les éléments d'une facette en un seul clic.

Croiser les facettes

Si vous choisissez plusieurs éléments au sein d'une facette, l'opération booléen OU s'applique.

Si vous choisissez un ou plusieurs éléments au sein d'une seconde facette, l'opération booléen ET s'applique entre les deux facettes.

Par exemple, au sein de la facette « Etablissements », sélectionnez Montpellier 2 puis au sein de la facette « Disciplines », sélectionnez Neurosciences, puis au sein de la facette « Langues » sélectionnez « anglais ».

Votre sélection est la suivante :



Vous obtenez la liste de toutes les thèses de Montpellier 2 dont la discipline est neurosciences et qui sont écrites en anglais.

Autre exemple : Vous cherchez toutes les thèses ayant trait à la langue

anglaise ou à la littérature anglo-saxonne soutenues dans les universités de Tours ou Orléans. Au sein de la facette « Etablissements », sélectionnez Orléans puis Tours. Ensuite, au sein de la facette « Domaines », sélectionnez Langues anglaise. Anglo-saxon et Littérature anglaise et anglo-saxonne.

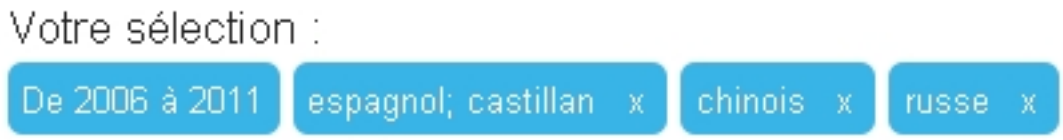
Votre sélection est la suivante :



Vous obtenez des thèses soutenues soit à Orléans soit à Tours qui appartiennent au domaine Langues anglaise. Anglo-saxon, ou au domaine Littérature anglaise et anglo-saxonne, ou à ces deux domaines.

Si vous croisez des facettes, vous constaterez peut-être avec étonnement la disparation de facettes sélectionnées précédemment. Ce comportement est normal : il est la conséquence de la mise à jour immédiate de la liste des résultats lors de la sélection d'une facette et des opérateurs booléens qui permettent de croiser les facettes.

Par exemple vous sélectionnez au sein de la facette « Langues » espagnol, chinois et russe. Vous obtenez :



Puis vous sélectionnez au sein de la facette « Domaines » sciences de l'ingénieur.

Vous obtenez



Le tag russe a disparu parce qu'au sein de l'ensemble de thèses écrites en espagnol, chinois et russe (opérateur OU au sein d'une facette), aucune de ces thèses n'appartient au domaine sciences de l'ingénieur (opérateur ET entre les facettes).

2.3.3. Nombre de résultats et rappel des critères de recherche

Le nombre de résultats apparait sur la gauche de l'écran, sous le logo de theses.fr

Si vous avez utilisé le moteur de recherche, les termes recherchés apparaissent :

dans le moteur de recherche

à la suite du nombre de résultats

en gras, dans un carré bleu appelé tag au sein de votre sélection

Dans votre sélection, vous trouverez le rappel de tous vos critères de recherche : termes recherchés par le moteur et facettes sélectionnés. Chaque tag peut être supprimé en cliquant sur la croix blanche, à l'exception de la fourchette de date.



2.3.4. La liste de résultats

Contenu de la liste de résultats

Chaque résultat est dans un pavé grisé qui comprend :

les références principales de la thèse sous la forme suivante :

Titre de la thèse par Nom du docteur, prénom du docteur, sous la direction de Nom du directeur, Prénom du directeur, et éventuellement des codirecteurs. – Etablissement de soutenance

Un clic sur le titre de la thèse renvoie à la page de la thèse.

Un clic sur le nom de l'auteur renvoie à la page de l'auteur.

Un clic sur le nom du directeur renvoie à la page du directeur.

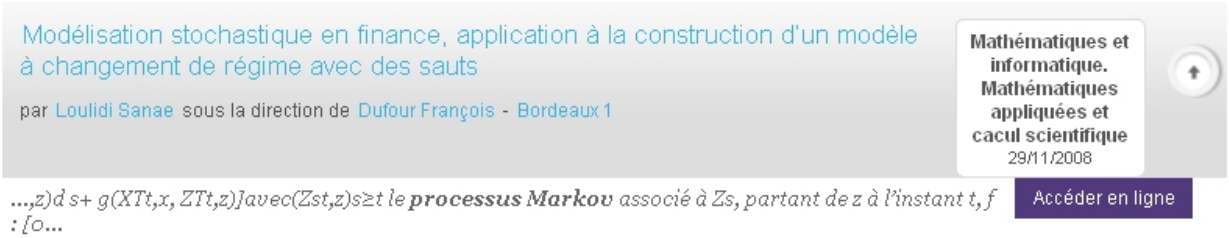
Un clic sur le nom de l'établissement renvoie à la page de l'établissement ;

un carré blanc indiquant la discipline de la thèse et la date de soutenance ;

éventuellement le bouton Accéder en ligne pour ouvrir directement le texte intégral de la thèse sans passer par la page donnant le détail de la description de la thèse ;

une flèche orientée vers le haut qui permet de revenir en haut de page ;

éventuellement, un court extrait du texte intégral où sont mis en gras les mots qui ont fait l'objet d'une recherche dans l'encart de recherche.



Naviguer dans la liste de résultats

Si votre recherche ramène plus de 10 résultats, plusieurs pages de résultats vous seront proposées.



Le numéro de la page active apparait en gras. Vous pouvez cliquer sur le numéro de la page de résultats qui vous intéresse, passer à la page suivante avec aller à la dernière page avec , revenir à la page précédente avec et revenir à la première page avec .

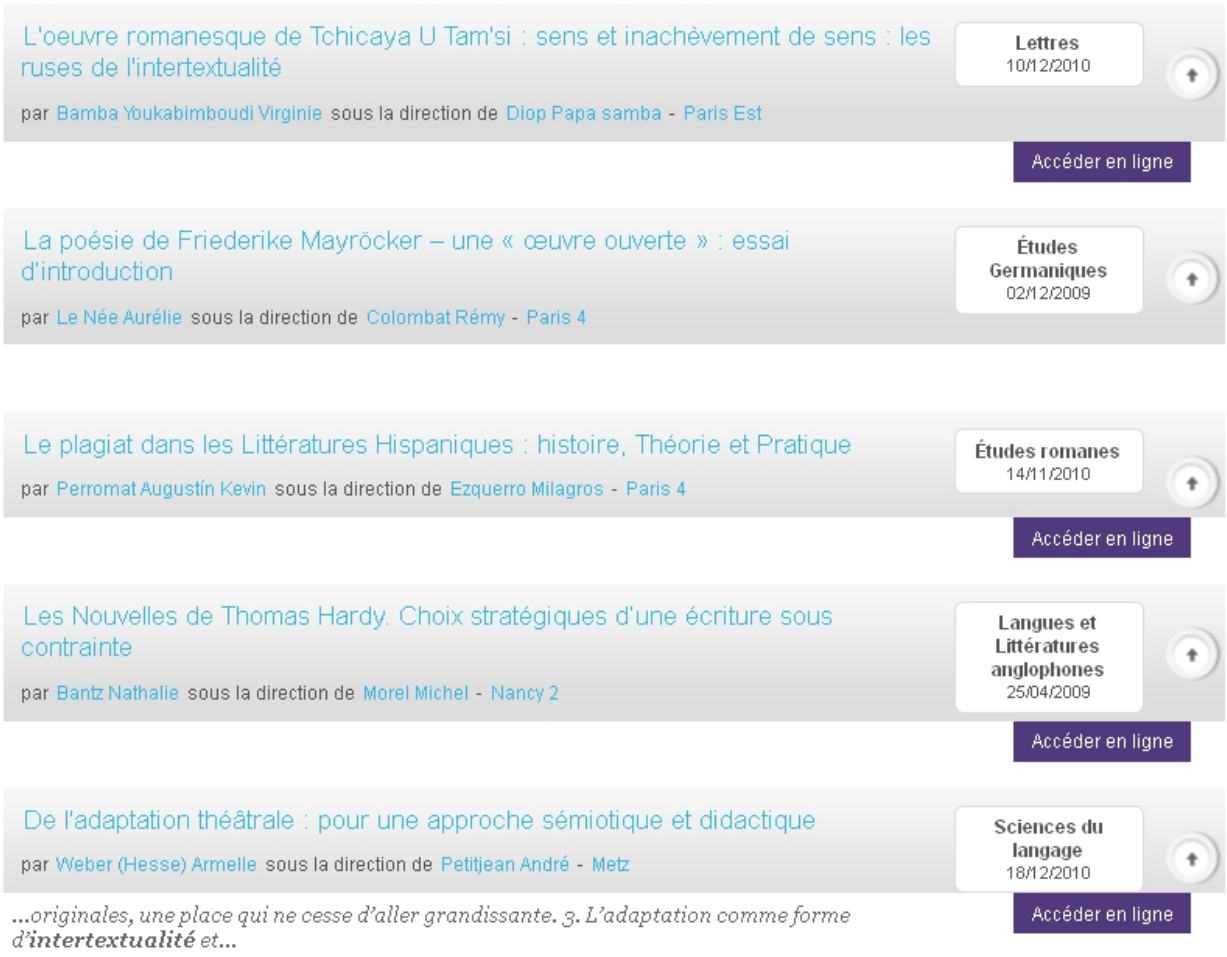
Tri de la liste de résultats

Par défaut la liste de résultats est triée par pertinence.

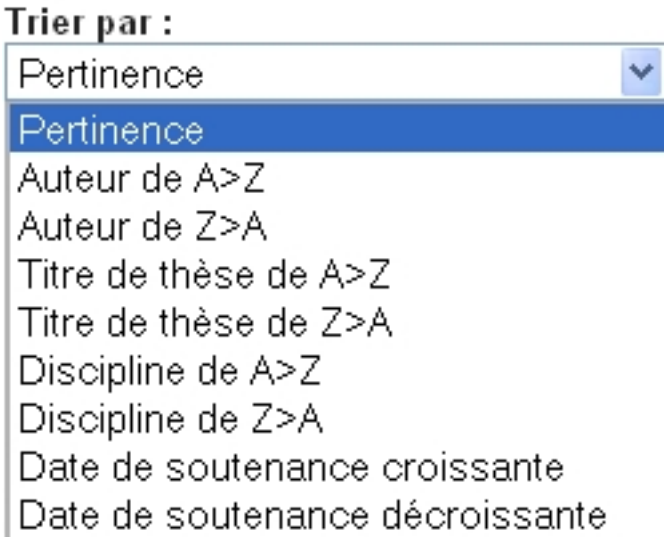
Qu’est-ce que la pertinence ?

theses.fr a été paramétré pour affecter un poids différent à chaque élément de description de la thèse et au texte intégral de la thèse. Ainsi, les mots du titre, le nom de l’auteur, celui du directeur ont plus de poids que les mots présents dans le texte intégral.

Par exemple, une recherche sur le mot intertextualité remontera d’abord les thèses contenant « intertextualité » dans le titre, puis dans les mots-clés ou le résumé, puis dans le texte intégral de la thèse.



Vous pouvez modifier le tri de la liste de résultats grâce au menu déroulant suivant :



Tous les mots, y compris les articles et les prépositions sont pris en compte. Une thèse dont le titre commence par Une sera trié à la lettre U.

2.3.5. Le bouton S'abonner : flux ATOM

En cliquant sur le bouton 

vous pouvez vous abonner au flux ATOM de la liste de résultats pour le récupérer dans votre agrégateur [http://fr.wikipedia.org/wiki/Agr%C3%A9gateur] (Netvibes, Google reader…).

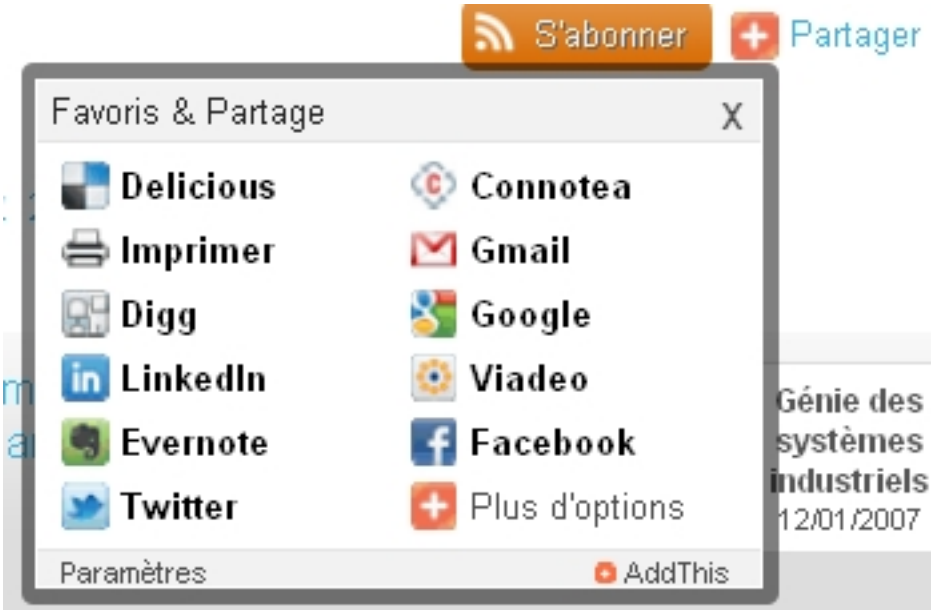
Vous serez automatiquement alerté de tout changement dans la liste de résultats dans votre agrégateur sans avoir à refaire votre recherche dans theses.fr

2.3.6. Le bouton Partager : le module AddThis

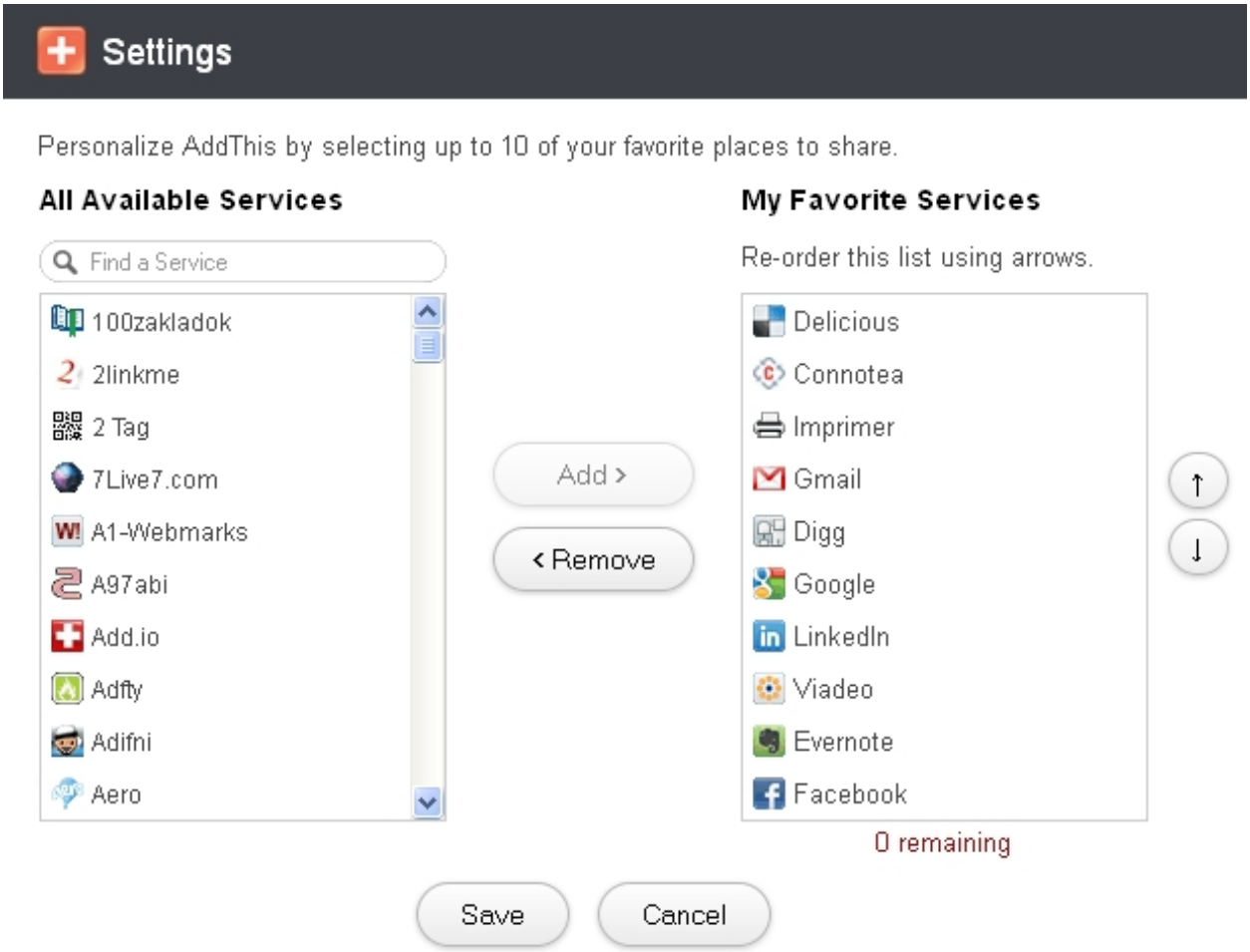
theses.fr intègre le module de partage de signets, AddThis.

AddThis est compatible avec des dizaines d'outils du web social.

Déplacez la souris sur le bouton Partager pour ouvrir le menu déroulant et voir les dix outils proposés par défaut.



Vous pouvez obtenir la liste exhaustive des outils disponibles en cliquant sur Plus d'options et vous pouvez paramétrer votre liste préférée en cliquant sur Paramètres.



Tous les liens disponibles sont proposés dans la fenêtre de gauche. Les dix sélectionnés passent dans la fenêtre de droite. Vous pouvez modifier leur ordre dans la fenêtre grâce aux flèches : l'icône des trois premiers apparaitra sur theses.fr à côté du bouton Partager. Enregistrez vos paramètres, fermez la fenêtre de paramétrage d'AddThis et rafraichissez la page de theses.fr (touche F5) pour que les modifications soient prises en compte.



Le paramétrage d'AddThis sera conservé au-delà de la session en cours sur

theses.fr mais il est lié au navigateur que vous utilisez sur un ordinateur donné.

2.4. La page d’une thèse

2.4.1. Le moteur de recherche

Sur quelles données porte la recherche ?

Le moteur de recherche de theses.fr porte sur les **informations de description de la thèse** (les noms et prénoms de l’auteur, du directeur de thèse, des membres du jury, les noms de l’établissement de soutenance, de l’école doctorale, des partenaires de recherche, le titre de la thèse, les résumés de la thèse, les mots clés…) **mais aussi le texte intégral de la thèse**.

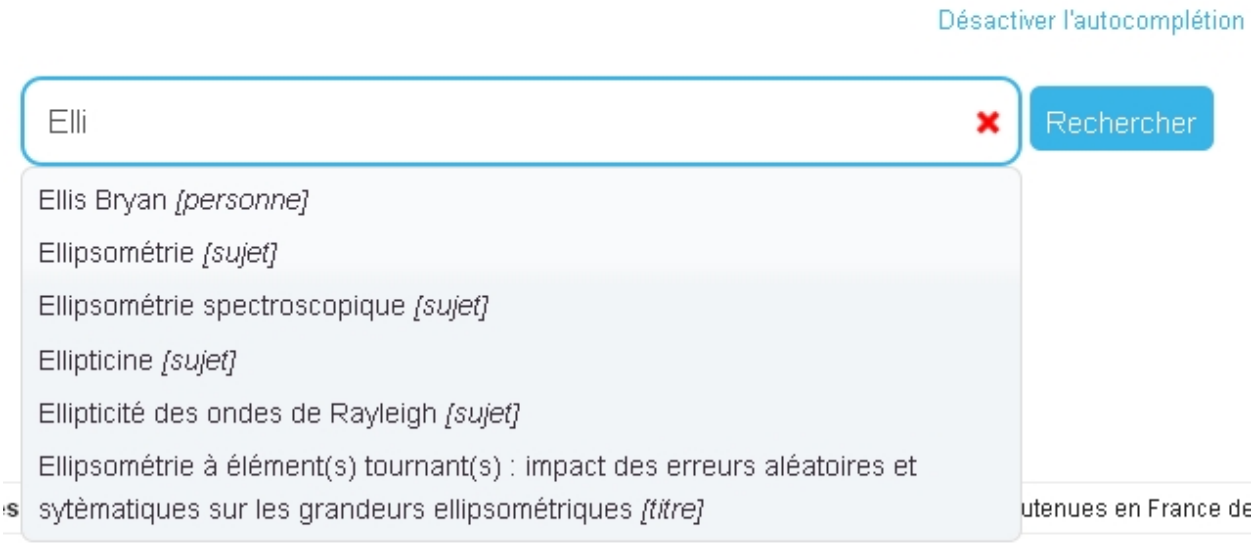
L’auto complétion

Lorsque vous commencez à taper un mot, theses.fr vous suggère des réponses de trois types :

des noms de personnes liées à des thèses

des mots clés utilisés dans ces thèses

des titres de thèse



Vous pouvez continuer votre saisie sans prêter attention à cette auto complétion ou bien sélectionnez avec la souris ou les flèches la suggestion qui vous intéresse.

L'auto complétion peut être désactivée par un clic sur le lien Désactiver l'autocomplétion et peut être réactivée par un clic sur Activer l'autocomplétion

Valider votre recherche

Vous pouvez soit appuyer sur la touche entrée de votre clavier soit cliquer sur le bouton Rechercher

Supprimer le terme recherché

Lorsque vous commencez à taper un caractère, une croix rouge apparait à droite de l'encart de recherche. Un clic sur cette croix rouge vide l'encart de recherche.



Recherche d’une expression exacte

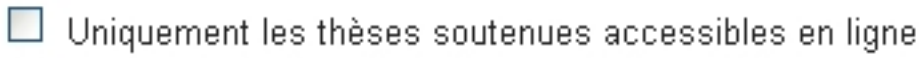
Si vous cherchez une expression exacte, vous devez la saisir entre guillemets.

Par exemple, si vous tapez droits fondamentaux, le moteur de recherche vous renverra toutes les thèses qui contiennent à la fois droits et fondamentaux.

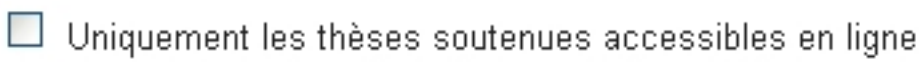
Si vous tapez « droits fondamentaux », vous aurez comme résultat toutes les thèses qui contiennent la chaine de caractère telle que vous l’avez écrite.

Le filtre « Uniquement les thèses soutenues accessibles en ligne »

Vous pouvez restreindre votre recherche aux seules thèses qui sont accessibles sur le web en cochant la case



Pour voir toutes les thèses soutenues accessibles en ligne, ne tapez rien dans l’encart de recherche, cochez la case



et cliquez sur



Mots vides, élision, lettres accentuées, casse et traitement de la chaine de caractères

Lorsque vous tapez une chaine de caractères dans le moteur de recherche, plusieurs traitements sont appliqués avant l’interrogation de la base de données.

Certains mots, appelés mots vides, ne sont pas pris en compte. Par exemple sont des mots vides : le, la les, ce, celle, de, des, un, une.. mais aussi et, ou,

à…

Un traitement des élisions est appliqué pour supprimer les l’, c’…

Les accents sont neutralisés, ainsi que la casse.

Enfin, tous les mots sont traités pour identifier leur racine et appliquer de fait une troncature permettant de chercher sur les formes plurielles ou féminines des mots.

Par exemple :

si vous tapez constitutionnalisation, la requête porte en fait sur constitutionnalis

si vous tapez classification phylogénétique, la requête porte en fait sur classif phylogenet

si vous tapez l’aéronautique, la requête porte en fait sur aeronaut

2.4.2. Les données décrivant la thèse

Le titre de la thèse, le nom de son auteur et la discipline sont les premières informations présentées.

Puis trois colonnes mentionnent :

le directeur de thèse, et les éventuels codirecteurs,

la date de soutenance, l’établissement de soutenance et éventuellement l’école doctorale et les partenaires de recherche impliqués dans la thèse (équipe de recherche, laboratoire, entreprise, fondation…)

s’ils sont connus, les membres du jury.


Psychopathie et troubles de la personnalité associés : recherche d'un effet particulier au trouble Borderline - Impact sur le risque de récidive.
par Nioche Angélique
Thèse de doctorat en Psychologie

Sous la direction de Réveillère Christian .

Soutenue le 30/06/2009
à Tours , dans le cadre de Santé, Sciences,Technologies , en partenariat avec SST/EA 2114 - Vieillissement et Développement Adulte ea 2114 (equipe de recherche) .

Le jury était composé de Fontaine Roger, Pham Thierry, Fernandez Lydia, Beaurepaire Christiane de.

Des liens hypertextes renvoient vers les pages des personnes ou des organismes cités.

Toutes les personnes et les organismes ont-ils une page dédiée dans theses.fr ? Non. Seules les personnes et les organismes identifiés de manière certaine, univoque et unique par un numéro ad hoc ont une page dédiée dans theses.fr Ce numéro, attribué par les bibliothécaires lors du traitement de la thèse correspond au numéro PPN de la notice d'autorité de la personne dans la base des autorités du catalogue Sudoc. Il permet de construire l'URL de la page de la personne ou de l'organisme sous la forme http://www.theses.fr/PPN

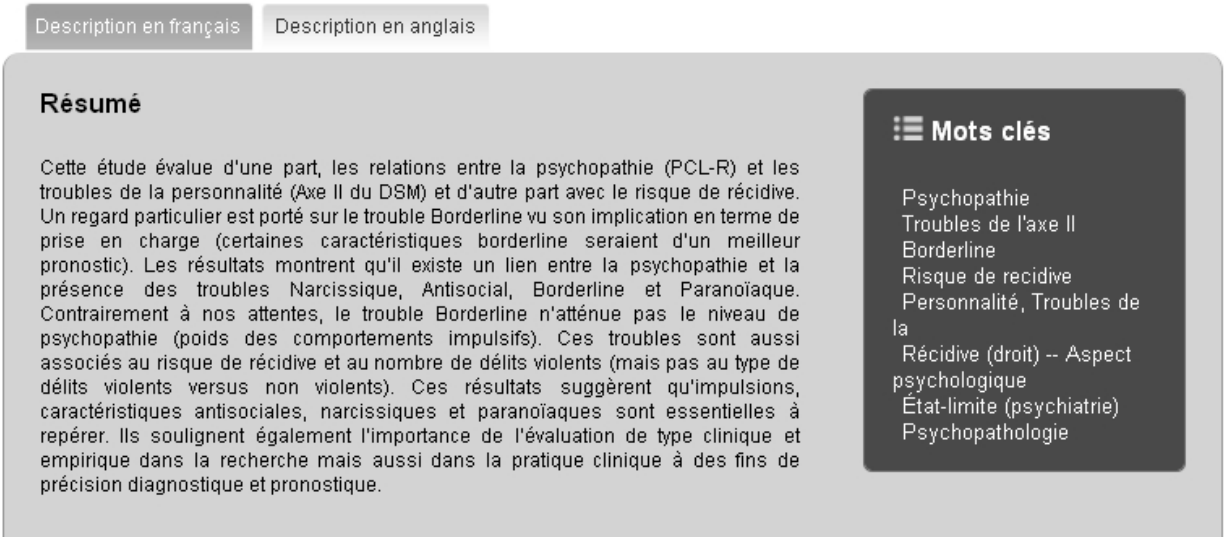
Exemple : http://www.theses.fr/122382781

Les auteurs de thèses, les directeurs de thèse et les établissements de soutenance ont toujours une page dédiée. Par contre, cela n'est pas systématique pour les présidents de jury, les rapporteurs, les établissements de cotutelle, les écoles doctorales et les partenaires de recherche.

En bas de page, les mots clés et les résumés sont présentés dans des onglets selon la langue de description.

Les mots clés en français sont soit des mots clés libres proposés par l'auteur lors du dépôt de son document pour l'organisation de la soutenance, soit une indexation effectuée par les bibliothécaires de l'établissement de soutenance (conforme à RAMEAU ou à FMESH).

Un clic sur un mot clé lance une nouvelle requête dans l'encart de recherche.



2.4.3. Accéder au texte intégral de la thèse

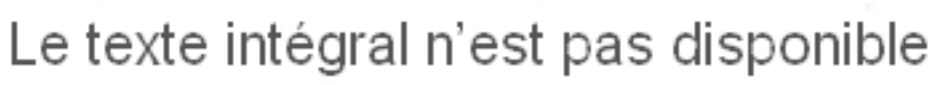
Au centre de la page de la thèse figurent les informations relatives à l'accès au texte intégral de la thèse.

Trois cas sont possibles :

1. Le docteur a accepté la diffusion sur le web. Le bouton 

permet d'ouvrir le document dans une nouvelle fenêtre.

2. Le docteur a refusé la diffusion sur le web. Vous voyez la mention



Consultez le Sudoc : il existe peut-être une version papier de la thèse qui serait consultable.

3. pour le moment, le texte intégral n'est pas diffusé sur le web, soit parce que le président de l'établissement de soutenance a prononcé une confidentialité, soit parce que le docteur souhaite conserver pour un temps

une exclusivité sur son document. La mention



précise la date à laquelle la thèse sera diffusée sur le web.

2.4.4. Comment citer la description de la thèse ?

Toutes les pages de theses.fr sont dotées d'une adresse URL spécifique construite sur le numéro national de thèse sous la forme http://www.theses.fr/NNT Le numéro national de thèse, attribué par la bibliothèque de l'établissement de soutenance comprend l'année de soutenance, un code d'identification de l'établissement, suivi d'un numéro séquentiel.

Exemples :

http://www.theses.fr/2010DIJOS001

http://www.theses.fr/2008METZ012L

Pour citer la page descriptive d'une thèse dans theses.fr, mentionnez cette URL.

2.4.5. Comment citer le document-thèse s'il est diffusé sur le web ?

Pour citer directement le texte intégral de la thèse, si celle-ci est accessible sur le web, utilisez l'URL de description de la thèse à laquelle vous ajoutez la mention /document : l'URL du document est donc sous la forme http://www.theses.fr/NNT/document

Exemples :

http://www.theses.fr/2010DIJOS001/document

http://www.theses.fr/2008METZ012L/document

2.4.6. Le bouton Partager

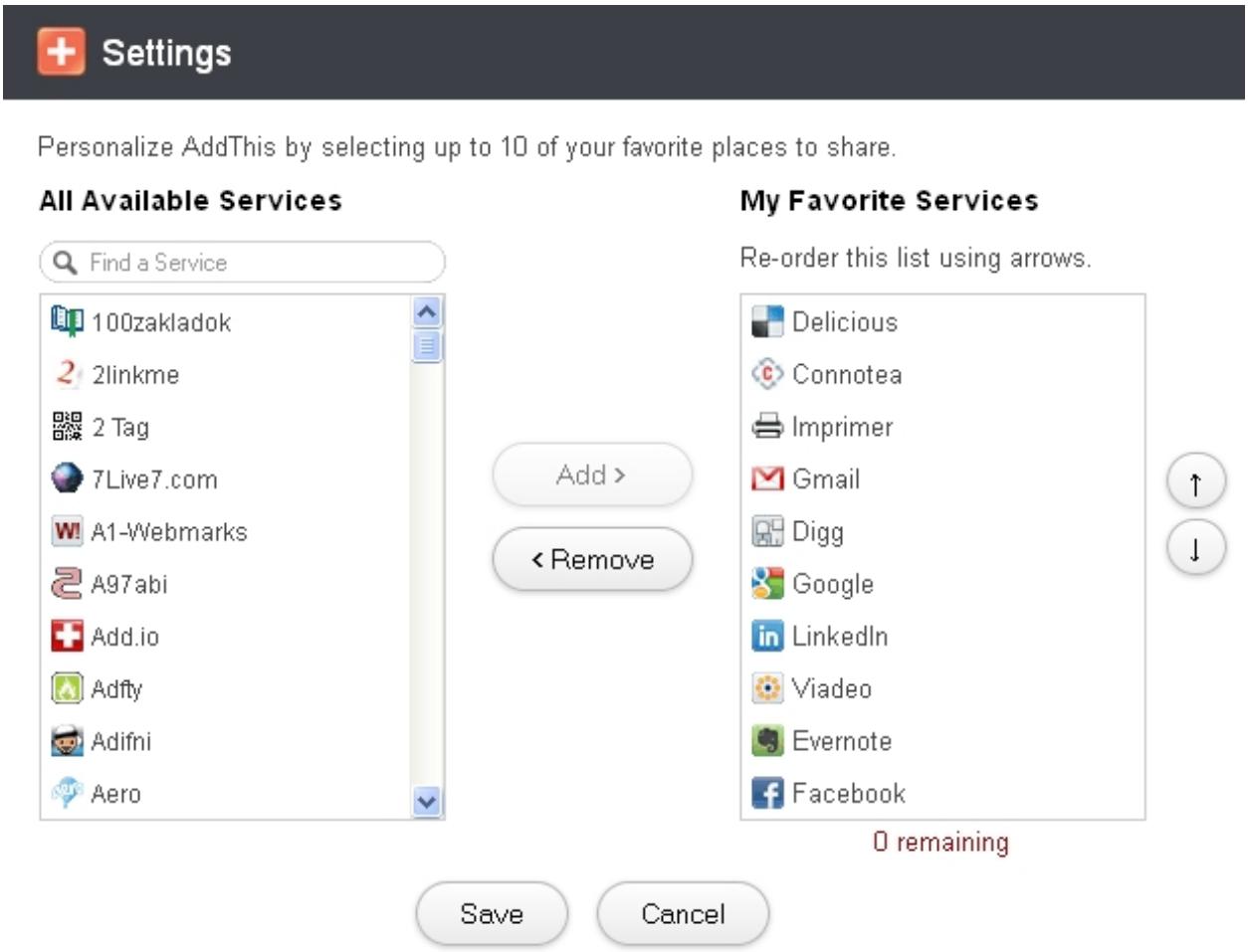
theses.fr intègre le module de partage de signets, AddThis.

AddThis est compatible avec des dizaines d'outils du web social.

Déplacez la souris sur le bouton Partager pour ouvrir le menu déroulant et voir les dix outils proposés par défaut.



Vous pouvez obtenir la liste exhaustive des outils disponibles en cliquant sur Plus d'options et vous pouvez paramétrer votre liste préférée en cliquant sur Paramètres.



Tous les liens disponibles sont proposés dans la fenêtre de gauche. Les dix sélectionnés passent dans la fenêtre de droite. Vous pouvez modifier leur ordre dans la fenêtre grâce aux flèches : l'icône des trois premiers apparaitra sur theses.fr à côté du bouton Partager. Enregistrez vos paramètres, fermez la fenêtre de paramétrage d'AddThis et rafraichissez la page de theses.fr (touche F5) pour que les modifications soient prises en compte.



Le paramétrage d'AddThis sera conservé au-delà de la session en cours sur theses.fr mais il est lié au navigateur que vous utilisez sur un ordinateur donné.

2.4.7. Signaler une erreur

Un clic sur ce bouton vous conduit sur le guichet d'assistance ABESstp [https://stp.abes.fr/assistance/domaines/1410/node/3?origine=thesesFr] dans le domaine Signaler une erreur. Vous pouvez décrire l'erreur rencontrée (faute de frappe, coquille…) en mentionnant l'URL de la page concernée ; votre demande sera transmise à la bibliothèque de l'établissement de soutenance pour correction.

2.4.8. Les boutons J'aime de Facebook et +1 de Google

Ces deux boutons fonctionnent de manière comparable. Ils vous permettent de signaler votre intérêt pour une thèse et de récupérer les références de cette thèse dans votre profil.

Lorsque vous cliquez sur le bouton J'aime de Facebook 

theses.fr communique à Facebook les références de la thèse, qui apparaissent dans votre profil Facebook. Si vous avez saisi un commentaire, celui-ci ainsi que les références de la thèse apparaitront sur votre mur.

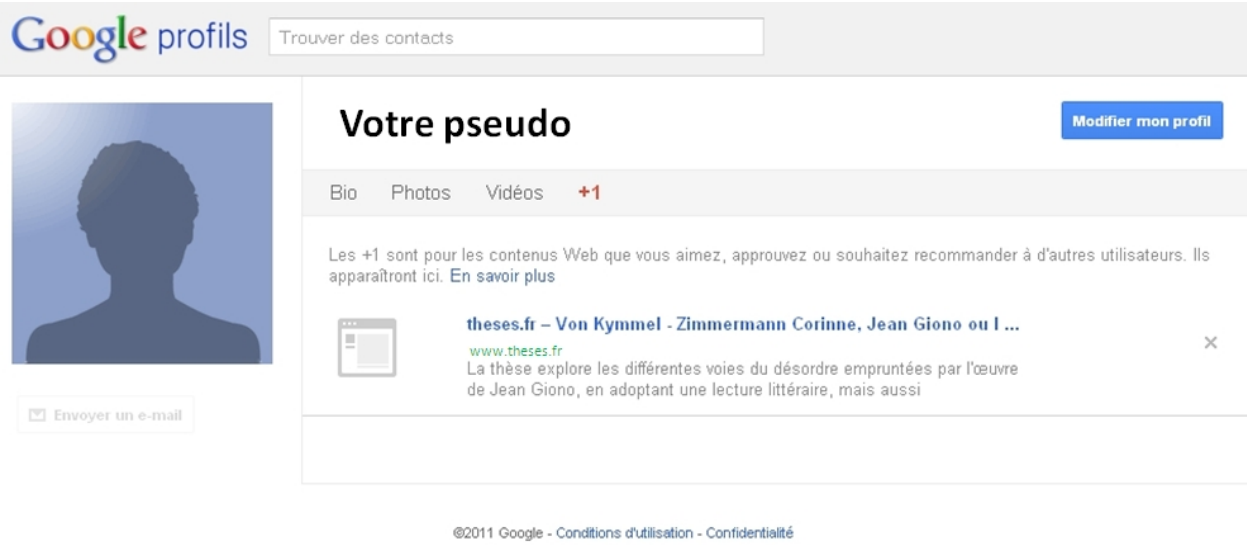


Lorsque vous cliquez sur le bouton +1 de Google, 

le nombre

de recommandations pour cette thèse augmente de 1.

Si vous allez ensuite dans votre profil Google, vous verrez apparaître les références de la thèse que vous avez marquée :



2.4.9. Récupérer les références de la thèse dans Zotero

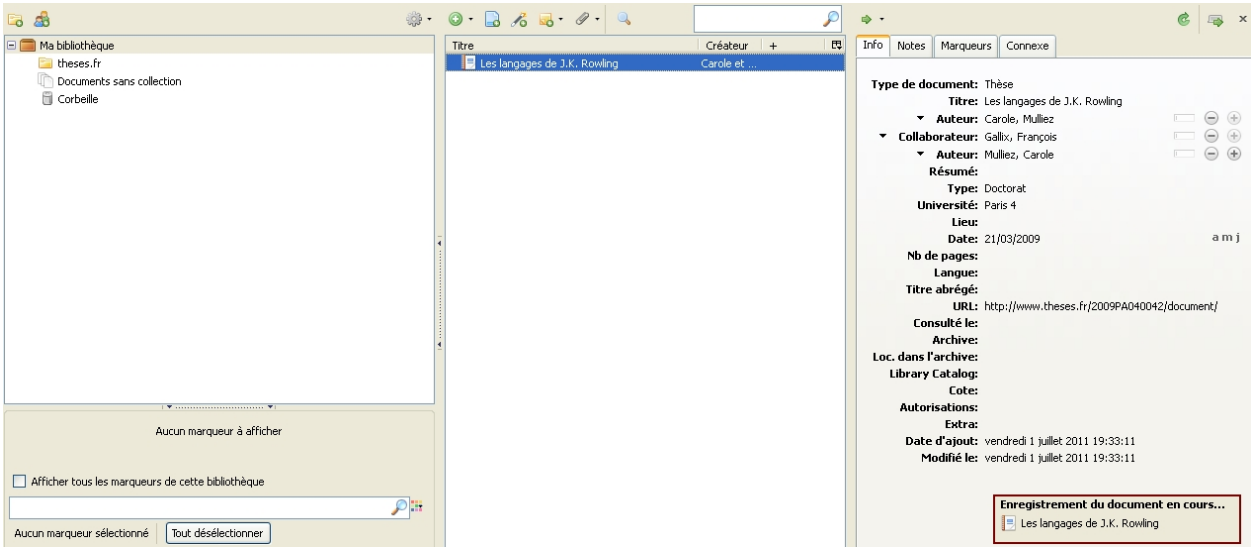
theses.fr est compatible avec l'outil en ligne de gestion des références bibliographiques Zotero [http://www.zotero.org/], plug-in à installer sur Firefox.

Les utilisateurs de Zotero peuvent constater dans toutes les pages de thèses la présence de l'icône de Zotero dans la barre d'adresse de Firefox pour récupérer les références bibliographiques de la thèse.



Cliquez sur cette icône de la barre d'adresse, puis ouvrez Zotero

Les références bibliographiques ont été chargées.



Si la thèse est accessible en ligne, l'URL chargée est l'URL du document-thèse, celle qui donne accès directement au texte intégral soit http://www.theses.fr/NNT/document

Si le texte intégral de la thèse n'est pas accessible, l'URL chargée est celle qui décrit la thèse soit http://www.theses.fr/NNT

2.5. La page d'une personne

Toutes les personnes ont-elles une page dédiée dans theses.fr ? Non. Seules les personnes identifiées de manière certaine, univoque et unique par un numéro ad hoc ont une page dédiée dans theses.fr Ce numéro, attribué par les bibliothécaires lors du traitement de la thèse correspond au numéro PPN de la notice d'autorité de la personne dans la base des autorités du catalogue Sudoc. Il permet de construire l'URL de la page de la personne sous la forme http://www.theses.fr/PPN

Exemple : http://www.theses.fr/122382781

Les auteurs de thèses et les directeurs de thèse ont toujours une page dédiée. Par contre, cela n'est pas systématique pour les présidents de jury et les rapporteurs.

2.5.1. Le moteur de recherche

Sur quelles données porte la recherche ?

Le moteur de recherche de theses.fr porte sur les **informations de description de la thèse** (les noms et prénoms de l'auteur, du directeur de thèse, des membres du jury, les noms de l'établissement de soutenance, de l'école doctorale, des partenaires de recherche, le titre de la thèse, les résumés de la thèse, les mots clés…) **mais aussi le texte intégral de la thèse**.

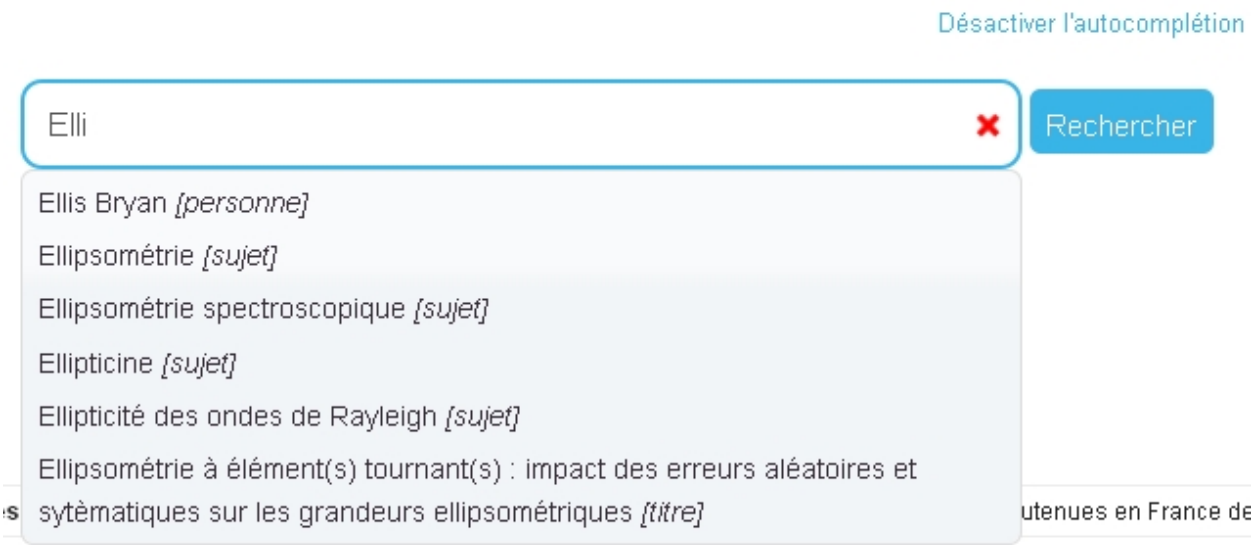
L'auto complétion

Lorsque vous commencez à taper un mot, theses.fr vous suggère des réponses de trois types :

des noms de personnes liées à des thèses

des mots clés utilisés dans ces thèses

des titres de thèse



Vous pouvez continuer votre saisie sans prêter attention à cette auto complétion ou bien sélectionnez avec la souris ou les flèches la suggestion qui vous intéresse.

L'auto complétion peut être désactivée par un clic sur le lien



Valider votre recherche

Vous pouvez soit appuyer sur la touche entrée de votre clavier soit cliquer sur le bouton 

Supprimer le terme recherché

Lorsque vous commencez à taper un caractère, une croix rouge apparait à droite de l'encart de recherche. Un clic sur cette croix rouge vide l'encart de recherche.



Recherche d'une expression exacte

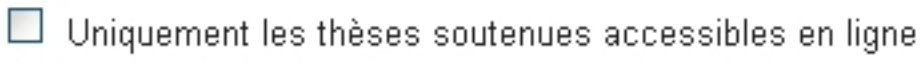
Si vous cherchez une expression exacte, vous devez la saisir entre guillemets.

Par exemple, si vous tapez droits fondamentaux, le moteur de recherche vous renverra toutes les thèses qui contiennent à la fois droits et fondamentaux.

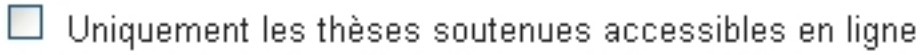
Si vous tapez « droits fondamentaux », vous aurez comme résultat toutes les thèses qui contiennent la chaine de caractère telle que vous l'avez écrite.

Le filtre « Uniquement les thèses soutenues accessibles en ligne »

Vous pouvez restreindre votre recherche aux seules thèses qui sont accessibles sur le web en cochant la case



Pour voir toutes les thèses soutenues accessibles en ligne, ne tapez rien dans l'encart de recherche, cochez la case



et cliquez sur 

Mots vides, élision, lettres accentuées, casse et traitement de la chaine de caractères

Lorsque vous tapez une chaine de caractères dans le moteur de recherche, plusieurs traitements sont appliqués avant l'interrogation de la base de données.

Certains mots, appelés mots vides, ne sont pas pris en compte. Par exemple sont des mots vides : le, la les, ce, celle, de, des, un, une.. mais aussi et, ou, à…

Un traitement des élisions est appliqué pour supprimer les l', c'…

Les accents sont neutralisés, ainsi que la casse.

Enfin, tous les mots sont traités pour identifier leur racine et appliquer de fait une troncature permettant de chercher sur les formes plurielles ou féminines des mots.

Par exemple :

si vous tapez constitutionnalisation, la requête porte en fait sur constitutionnalis

si vous tapez classification phylogénétique, la requête porte en fait sur classif phylogenet

si vous tapez l'aéronautique, la requête porte en fait sur aeronaut

2.5.2. Le nuage de la personne

La page de la personne présente un nuage de mots relatif à la personne.

Ce nuage de mots est construit à partir de toutes les thèses en lien avec la personne quel que soit le rôle que la personne a joué par rapport à la thèse (auteur, directeur, président de jury..)

Le nuage de mots est bâti sur les mots clés de toutes ces thèses, qu'il s'agisse des mots clés libres attribués par le docteur ou des mots clés contrôlés attribués par le bibliothécaire lors du traitement de la thèse dans les catalogues.

Le nuage de mots permet de visualiser en un coup d'œil les champs scientifiques où oeuvre la personne. La taille et la couleur des mots au sein du nuage est proportionnelle à leur nombre d'occurrence dans les thèses liées : plus un mot apparait en gros, plus il est présent dans les thèses en lien avec la personne.

Voici par exemple le nuage de mots de Cacaly Serge :


Nuage de Cacaly Serge
Analyse documentaire Analyse multi résolution Communication Culture Descripteur mémoriel Documentalistes Développement durable Forme sémiotique interne Gouvernance ascendante Indexation Indexation qualitative Information scientifique Information technique Informatique documentaire Intelligence territoriale Intelligence économique Lobbying Modélisation de l'information Métadonnées Ondelettes Patrimoine Patrimoine culturel Phénomène d'influence Plan de veille personnalisé Recherche d'information Recherche de l'information Ressources territoriales Rôle du professionnel de l'information Systèmes d'information Systèmes d'information Gestion Systèmes de Recherche d'Information Sémiotique Transformées en ondelettes Veille Veille technologique

Trois niveaux représentés par trois couleurs et taille de police apparaissent au sein de ce nuage. Le terme « Recherche de l'information » est celui qui apparaît le plus souvent dans les mots clés des thèses liées.

Le nuage des mots est évolutif : à chaque fois qu'une nouvelle thèse est référencée dans theses.fr le nuage de mots des personnes et organismes liés à cette thèse est modifié.

Il est possible d'agrandir le nuage de mots pour qu'il s'affiche en surimpression sur l'écran en cliquant sur 

2.5.3. Les différents rôles de la personne et la liste des thèses liées

Les thèses liées à la personne sont présentées à la suite du nuage de mots. Les relations qu'une personne entretient avec une thèse peuvent être les suivantes :

- auteur d'une thèse
- directeur de thèse
- président d'un jury de thèse
- rapporteur d'une thèse lors de la soutenance

La liste des thèses concernées est affichée selon ces quatre rôles.

Par exemple Chemat Farid est lié à cinq thèses dans trois rôles différents :



2.5.4. Le bouton Partager : le module AddThis

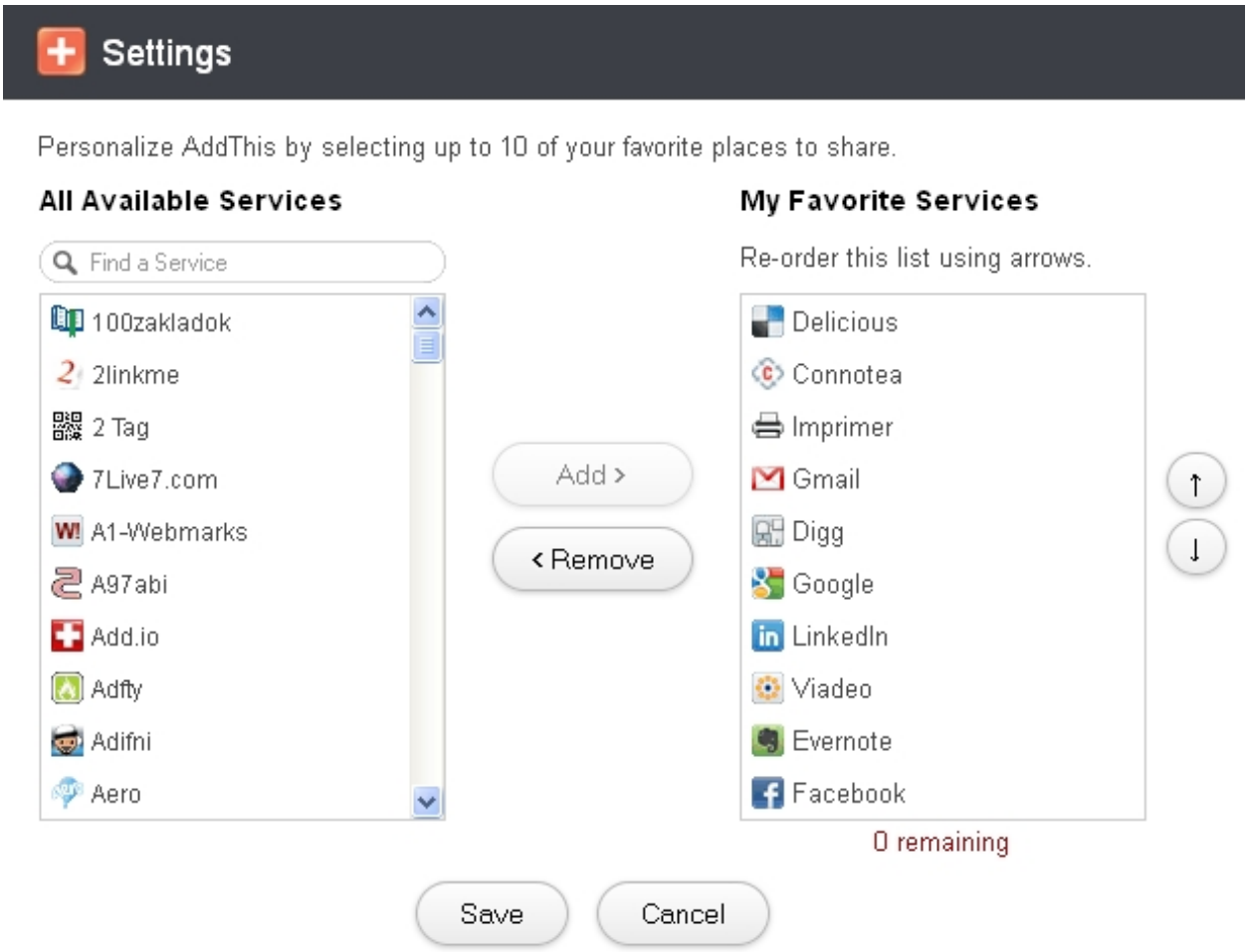
theses.fr intègre le module de partage de signets, AddThis.

AddThis est compatible avec des dizaines d'outils du web social.

Déplacez la souris sur le bouton Partager pour ouvrir le menu déroulant et voir les dix outils proposés par défaut.



Vous pouvez obtenir la liste exhaustive des outils disponibles en cliquant sur Plus d'options et vous pouvez paramétrer votre liste préférée en cliquant sur Paramètres.



Tous les liens disponibles sont proposés dans la fenêtre de gauche. Les dix sélectionnés passent dans la fenêtre de droite. Vous pouvez modifier leur ordre dans la fenêtre grâce aux flèches : l'icône des trois premiers apparaitra theses.fr à côté du bouton Partager. Enregistrez vos paramètres, fermez la fenêtre de paramétrage d'AddThis et rafraichissez la page de theses.fr (touche F5) pour que les modifications soient prises en compte.



Le paramétrage d'AddThis sera conservé au-delà de la session en cours sur theses.fr mais il est lié au navigateur que vous utilisez sur un ordinateur donné.

2.5.5. Signaler une erreur

Un clic sur ce bouton vous conduit sur le guichet d'assistance ABESstp [https://stp.abes.fr/assistance/domaines/1410/node/3?origine=thesesFr] dans le domaine Signaler une erreur. Vous pouvez décrire l'erreur rencontrée (faute de frappe, coquille…) en mentionnant l'URL de la page concernée ; votre demande sera transmise à la bibliothèque de l'établissement de soutenance pour correction.

2.6. La page d'un organisme

Toutes les organismes ont-ils une page dédiée dans theses.fr ?

Non. Seules les organismes identifiés de manière certaine, univoque et unique par un numéro ad hoc ont une page dédiée dans theses.fr Ce numéro, attribué par les bibliothécaires lors du traitement de la thèse correspond au numéro PPN de la notice d'autorité de l'organisme dans la base des autorités du catalogue Sudoc. Il permet de construire l'URL de la page de l'organisme sous la forme http://www.theses.fr/PPN

Exemple : http://www.theses.fr/026403412

Les établissements de soutenance français ont toujours une page dédiée. Par contre, cela n'est pas systématique pour les établissements de co-tutelle, les écoles doctorales et les partenaires de recherche.

2.6.1. Le moteur de recherche

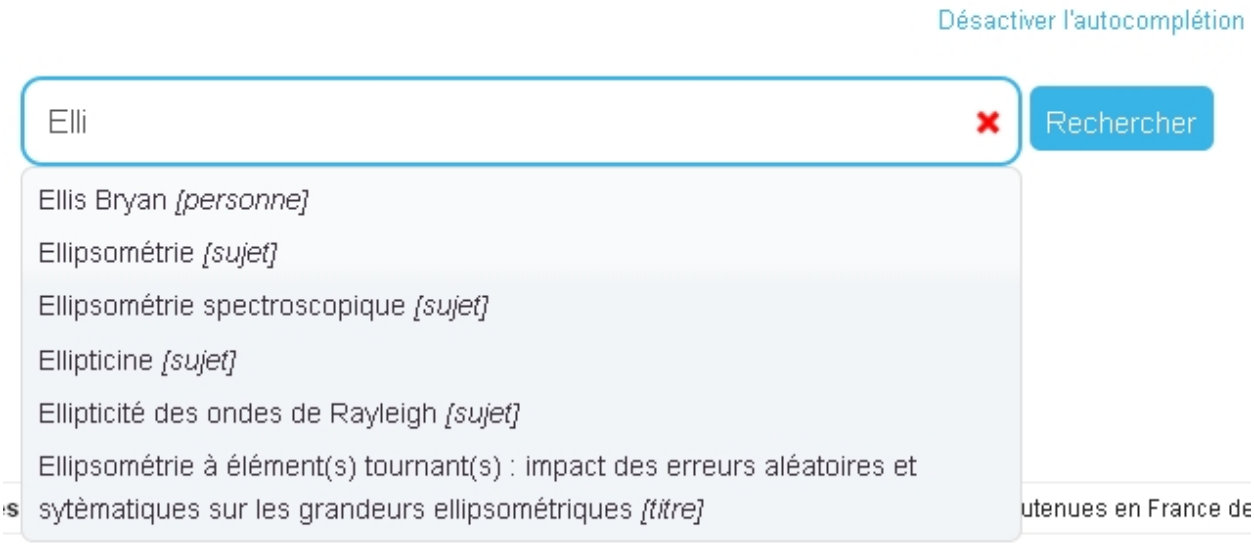
Sur quelles données porte la recherche ?

Le moteur de recherche de theses.fr porte sur les **informations de description de la thèse** (les noms et prénoms de l'auteur, du directeur de thèse, des membres du jury, les noms de l'établissement de soutenance, de l'école doctorale, des partenaires de recherche, le titre de la thèse, les résumés de la thèse, les mots clés…) **mais aussi le texte intégral de la thèse**.

L'auto complétion

Lorsque vous commencez à taper un mot, theses.fr vous suggère des réponses de trois types :

- des noms de personnes liées à des thèses
- des mots clés utilisés dans ces thèses
- des titres de thèse



Vous pouvez continuer votre saisie sans prêter attention à cette auto complétion ou bien sélectionnez avec la souris ou les flèches la suggestion qui vous intéresse.

L'auto complétion peut être désactivée par un clic sur le lien Désactiver l'autocomplétion et peut être réactivée par un clic sur Activer l'autocomplétion

Valider votre recherche

Vous pouvez soit appuyer sur la touche entrée de votre clavier soit cliquer sur le bouton Rechercher

Supprimer le terme recherché

Lorsque vous commencez à taper un caractère, une croix rouge apparait à droite de l'encart de recherche. Un clic sur cette croix rouge vide l'encart de recherche.



Recherche d'une expression exacte

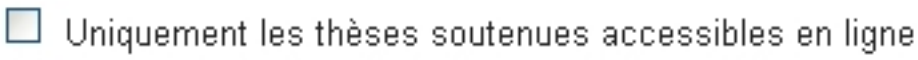
Si vous cherchez une expression exacte, vous devez la saisir entre guillemets.

Par exemple, si vous tapez droits fondamentaux, le moteur de recherche vous renverra toutes les thèses qui contiennent à la fois droits et fondamentaux.

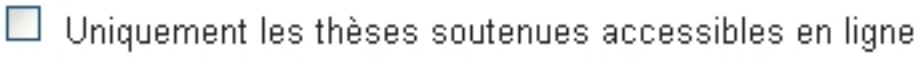
Si vous tapez « droits fondamentaux », vous aurez comme résultat toutes les thèses qui contiennent la chaine de caractère telle que vous l'avez écrite.

Le filtre « Uniquement les thèses soutenues accessibles en ligne »

Vous pouvez restreindre votre recherche aux seules thèses qui sont accessibles sur le web en cochant la case



Pour voir toutes les thèses soutenues accessibles en ligne, , ne tapez rien dans l'encart de recherche, cochez la case



et cliquez sur 

Mots vides, élision, lettres accentuées, casse et traitement de la chaine de caractères

Lorsque vous tapez une chaine de caractères dans le moteur de recherche, plusieurs traitements sont appliqués avant l'interrogation de la base de données.

Certains mots, appelés mots vides, ne sont pas pris en compte. Par exemple sont des mots vides : le, la les, ce, celle, de, des, un, une.. mais aussi et, ou, à…

Un traitement des élisions est appliqué pour supprimer les l', c'…

Les accents sont neutralisés, ainsi que la casse.

Enfin, tous les mots sont traités pour identifier leur racine et appliquer de fait une troncature permettant de chercher sur les formes plurielles ou féminines des mots.

Par exemple :

si vous tapez constitutionnalisation, la requête porte en fait sur constitutionnalis

si vous tapez classification phylogénétique, la requête porte en fait sur classif phylogenet

si vous tapez l'aéronautique, la requête porte en fait sur aeronaut

2.6.2. Le nuage de l'organisme

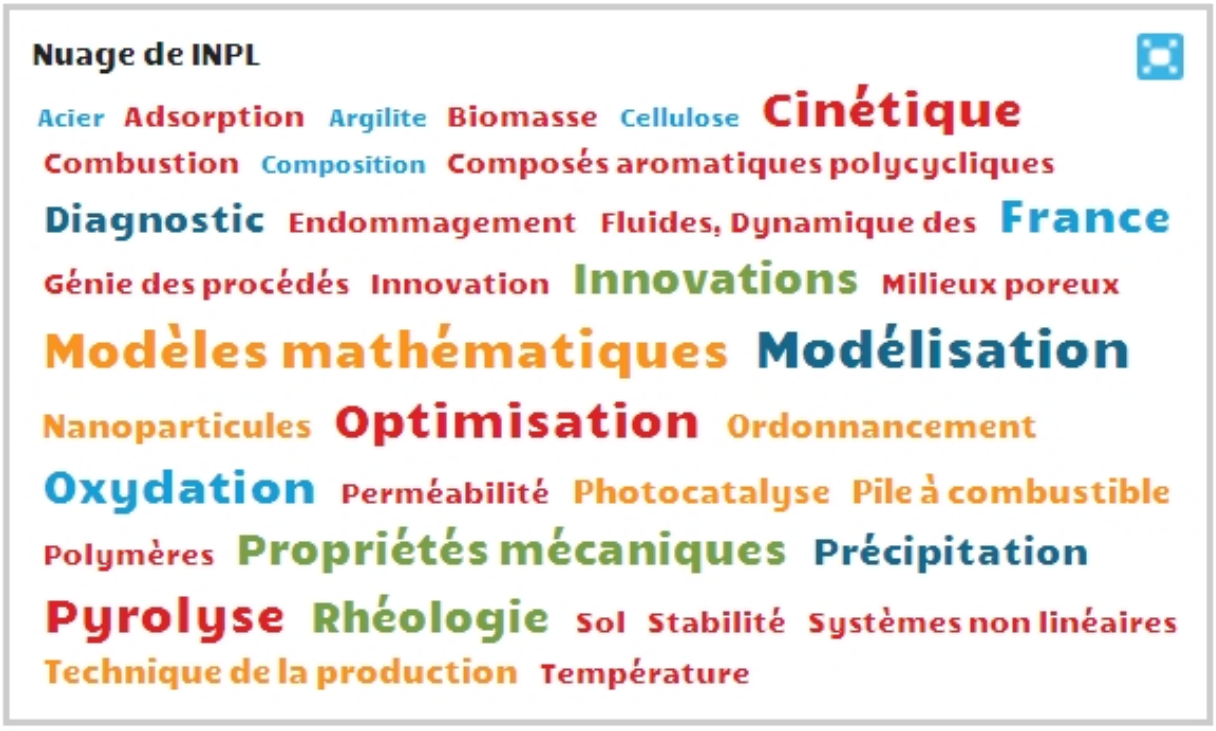
La page de l'organisme présente un nuage de mots relatif à l'organisme.

Ce nuage de mots est construit à partir de toutes les thèses liées à l'organisme. Le nuage de mots est bâti sur les mots clés de ces thèses, qu'il s'agisse des mots clés libres attribués par le docteur ou des mots clés contrôlés attribués par le bibliothécaire lors du traitement de la thèse dans les catalogues.

Le nuage de mots permet de visualiser en un coup d'œil les champs scientifiques de l'organisme. La taille et la couleur des mots au sein du nuage est proportionnelle à leur nombre d'occurrence dans les thèses liées : plus un

mot apparait en gros, plus il est présent dans les thèses en lien avec l'organisme.

Voici par exemple le nuage de mot pour l'INPL :

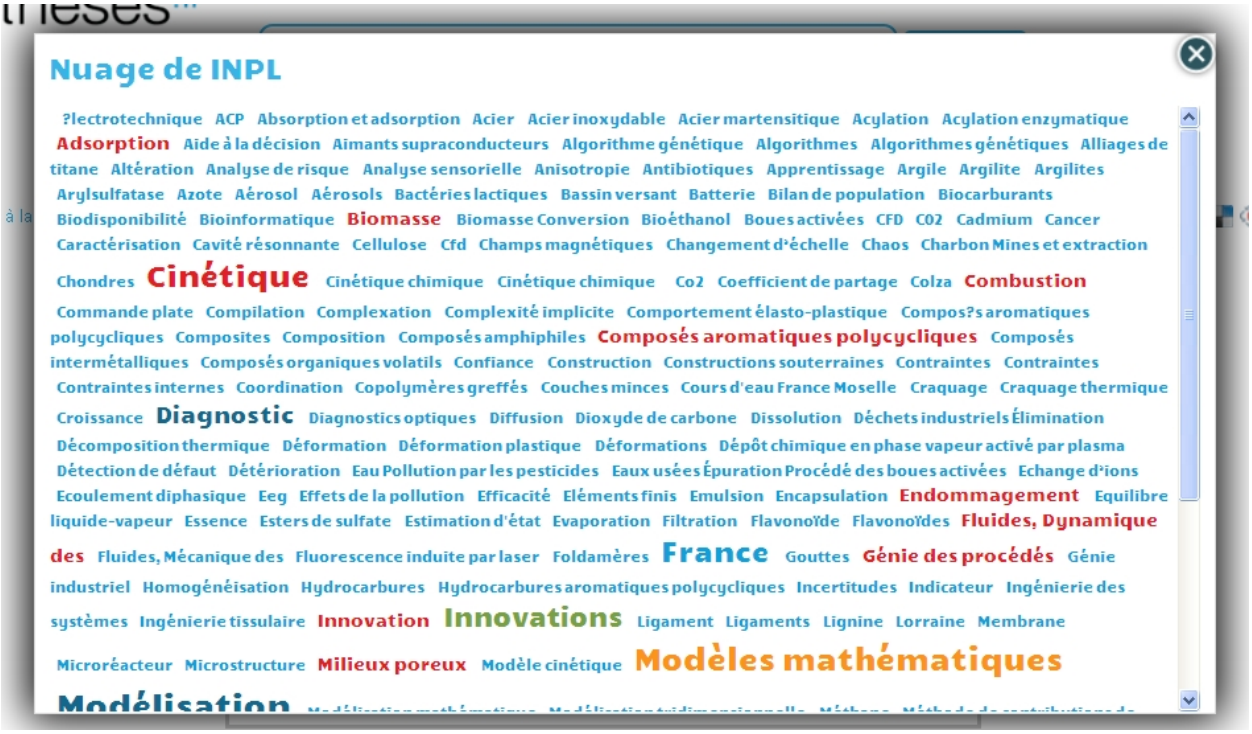


Huit niveaux représentés par huit couleurs et tailles de police apparaissent au sein de ce nuage. Le terme « Modélisation » est celui qui apparaît le plus souvent dans les mots clés de l'ensemble des thèses liées.

Le nuage des mots est évolutif : à chaque fois qu'une nouvelle thèse est référencée dans theses.fr le nuage de mots des personnes et organismes liés à cette thèse est modifié.

Il est possible d'agrandir le nuage de mots pour qu'il s'affiche en surimpression sur l'écran en cliquant sur  Le nuage est alors plus fourni : vous pouvez balayer tous les mots clés en utilisant l'ascenseur vertical.

Voici par exemple le nuage de mot pour l'INPL :



2.6.3. Les différents organismes et la liste des trente dernières thèses

Les différents organismes dans theses.fr sont :

- les établissements. Il s'agit des établissements de soutenance et des établissements de cotutelle
- les écoles doctorales. Le doctorant était affilié à cette école doctorale le temps de sa thèse.
- les partenaires de recherche. Ce sont des organismes ayant participé à la réalisation de la thèse par la mise à disposition de moyens, notamment en terme de financement de la recherche doctorale. Les partenaires de recherche peuvent être :
- des laboratoires de recherche
- des équipes de recherche

- des entreprises
- des fondations
- ou autres

En dessous du nuage de mots, vous trouverez le nombre de thèses liées à l'organisme et un extrait des thèses liées (trente au maximum) :

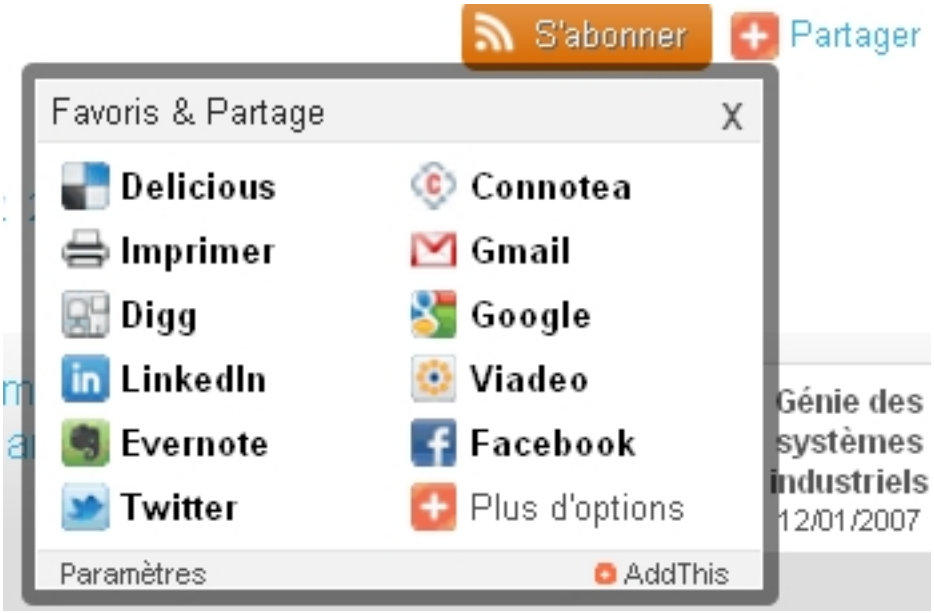


2.6.4. Le bouton Partager

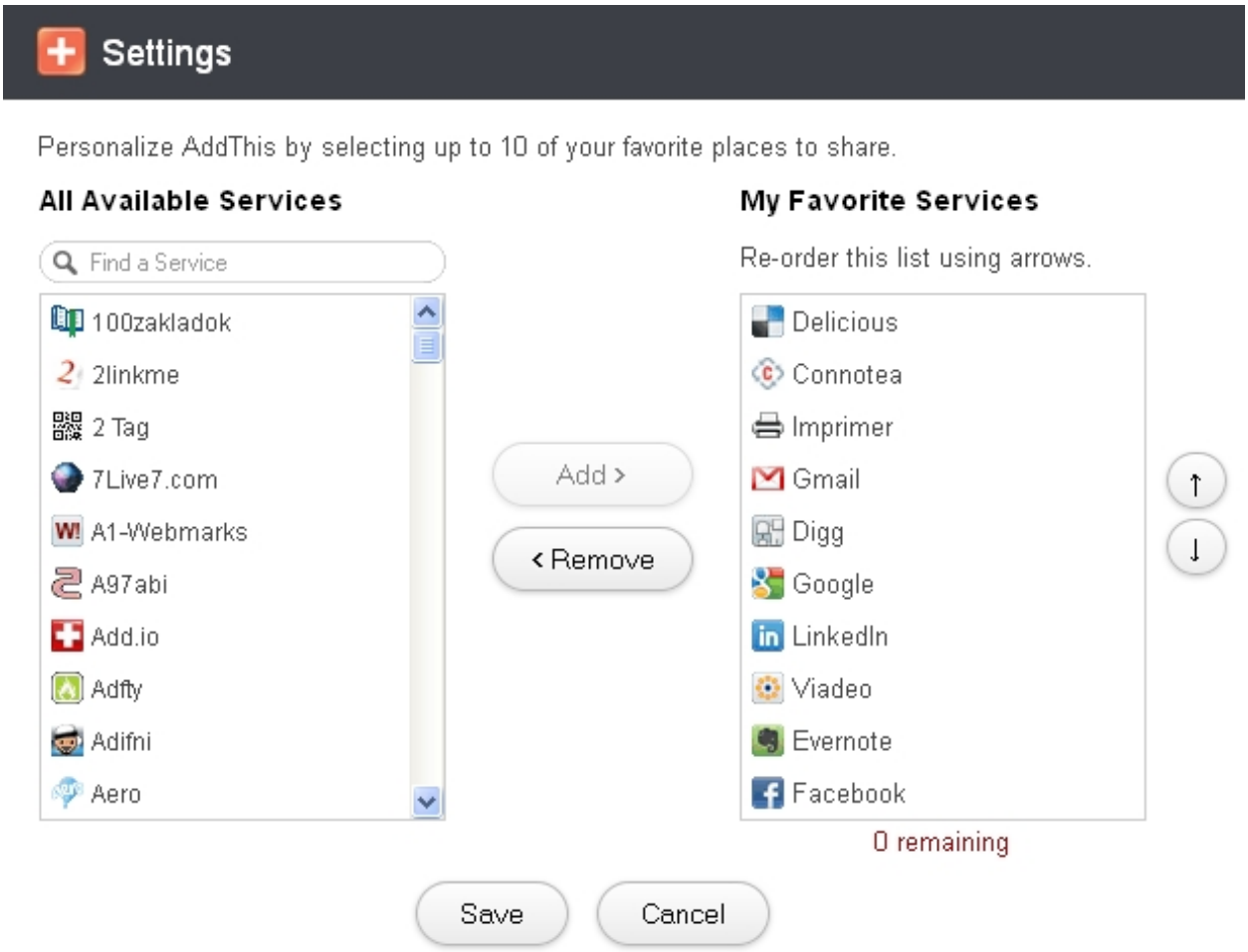
theses.fr intègre le module de partage de signets, AddThis.

AddThis est compatible avec des dizaines d’outils du web social.

Déplacez la souris sur le bouton Partager pour ouvrir le menu déroulant et voir les dix outils proposés par défaut.



Vous pouvez obtenir la liste exhaustive des outils disponibles en cliquant sur Plus d’options et vous pouvez paramétrer votre liste préférée en cliquant sur Paramètres.



Tous les liens disponibles sont proposés dans la fenêtre de gauche. Les dix sélectionnés passent dans la fenêtre de droite. Vous pouvez modifier leur ordre dans la fenêtre grâce aux flèches : l'icône des trois premiers apparaitra theses.fr à côté du bouton Partager. Enregistrez vos paramètres, fermez la fenêtre de paramétrage d'AddThis et rafraichissez la page de theses.fr (touche F5) pour que les modifications soient prises en compte.



Le paramétrage d'AddThis sera conservé au-delà de la session en cours sur theses.fr mais il est lié au navigateur que vous utilisez sur un ordinateur donné.

2.6.5. Signaler une erreur

Un clic sur ce bouton vous conduit sur le guichet d’assistance ABESstp [https://stp.abes.fr/assistance/domaines/1410/node/3?origine=thesesFr] dans le domaine Signaler une erreur. Vous pouvez décrire l’erreur rencontrée (faute de frappe, coquille…) en mentionnant l’URL de la page concernée ; votre demande sera transmise à la bibliothèque de l’établissement de soutenance pour correction.

Chapitre 3. Accéder aux données brutes de theses.fr : l'API XML

3.1. Introduction à l'API XML de theses.fr

L'architecture technique de theses.fr permet grâce à des servlets de délivrer le contenu de la base de données sous d'autres formats que le format HTML utilisé par l'interface www.theses.fr [http://www.theses.fr/]

Les API REST XML peuvent être utilisés par qui le souhaite pour récupérer des données. Vous pouvez donc accéder aux données brutes telles que renvoyées par SolR la plateforme de recherche opensource qu'utilise theses.fr

Le principe est le suivant :

Utilisez l'interface www.theses.fr [http://www.theses.fr/] pour construire votre requête

Récupérez l'URL dans la barre d'adresse de votre navigateur

Ajoutez à cette URL un paramètre pour obtenir les données brutes dans le format de vos choix

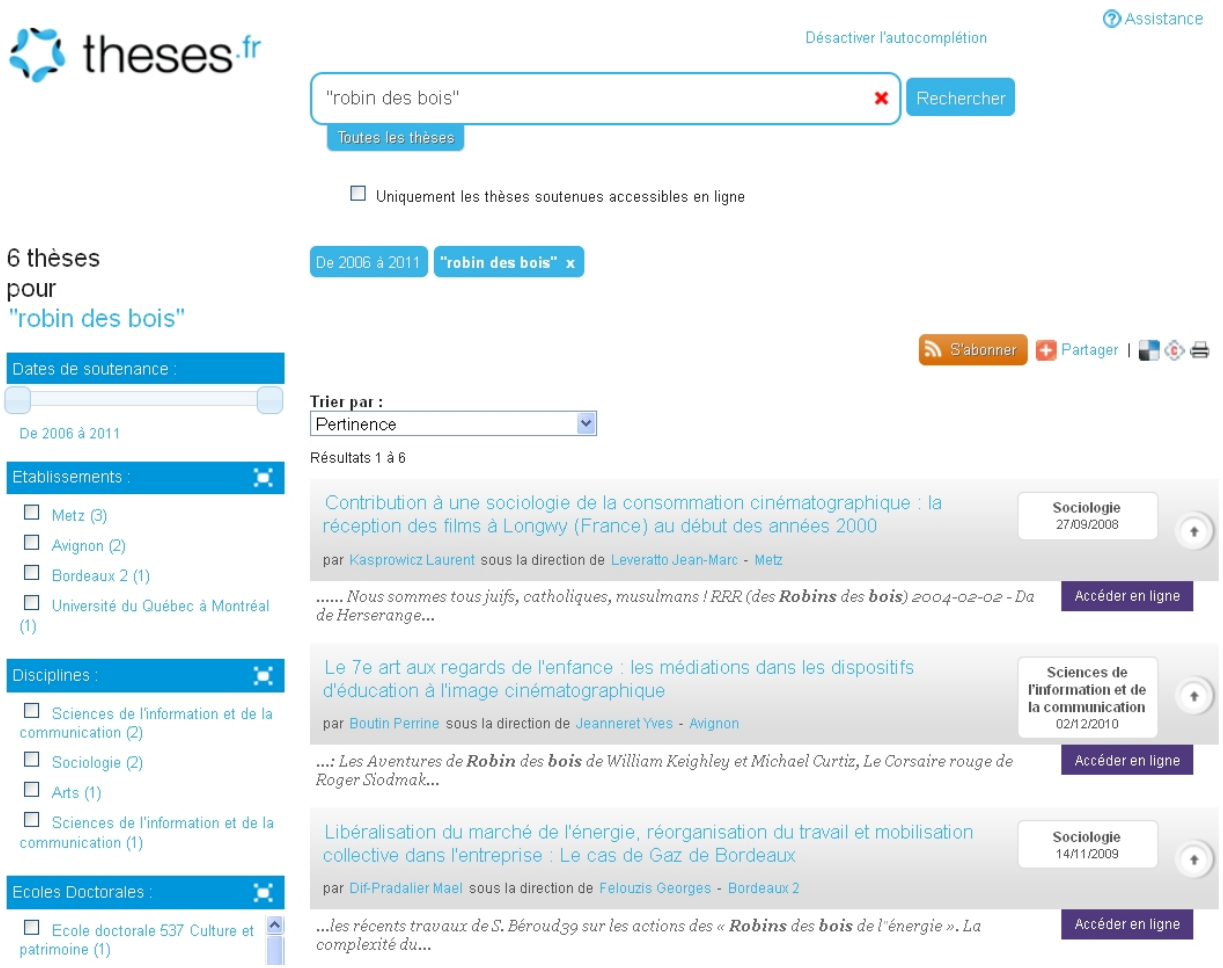
Dans un premier temps, peu de formats de récupération des données sont proposés. Mais l'ABES est à l'écoute de vos besoins et examinera avec attention toute demande d'évolution. Utilisez le guichet d'assistance ABESstp [https://stp.abes.fr/assistance/domaines/1415/node/3?origine=thesesFr] pour nous faire part de vos remarques et commentaires.

3.2. Récupérer une liste de résultats

L'API de theses.fr vous permet de récupérer une liste de résultats dans plusieurs formats possibles.

Exemple :

Vous voulez récupérer les résultats de la recherche « robin des bois ». L'URL de la requête est http://www.theses.fr/?q="robin+des+bois"



Format xml

Pour récupérer les données au format xml, ajoutez le paramètre &format=xml à la suite de l'URL de la requête.

Vous obtenez le fichier xml de réponse de SolR.

En début de fichier, vous verrez les paramètres SolR par défaut :

qf : pondération entre les index appliquée dans theses.fr

facet, facet.mincount et facet.limit : paramètres des facettes

rows : nombre de résultats renvoyés par la réponse. Ce nombre est fixé à 10 par défaut.

fl : liste des index qui doivent être affichés dans la réponse

start : numéro du résultat à partir duquel la réponse s'affiche. Par défaut, l'affichage commence au résultat 0.

qt : dismax ce paramètre appelle la pondération des index précisée dans qf

q : la requête

La réponse SolR remonte entre chaque balise <doc> les détails suivants de chaque thèse :

accessible : oui si le document est diffusé sur le web ; non, s'il ne l'est pas

auteur : nom et prénom de l'auteur

auteurPpn : numéro de la notice d'autorité de l'auteur dans le Sudoc et dans IdRef

dateSoutenance : date de soutenance

directeurThese : nom et prénom des directeurs de thèse sous forme de liste

directeursThese : nom et prénom du premier directeur de thèse, nom et prénom du deuxième directeur de thèse etc.

directeurThesePpn : numéro de la notice d'autorité du directeur de thèse dans le Sudoc et dans IdRef

discipline : discipline de la thèse

etabSoutenance : nom de l'établissement de soutenance

etabSoutenancePpn : numéro de la notice d'autorité de l'établissement de soutenance dans le Sudoc et dans IdRef

etablissement : nom de l’établissement de soutenance et de l’établissement de cotutelle sous forme de liste

id : numéro interne dans theses.fr

langueThese : langues de la thèse sous forme de liste

num : pour les thèses soutenues, numéro national de thèse

oaiSetSpec : codes de domaine issus de la classification décimale Dewey et utilisés par la recommandation TEF sous forme de liste

status : soutenue pour les thèses soutenues

titre : titre de la thèse

La réponse SolR remonte dans le bloc <facet_fields> les informations sur les facettes suivantes :

etablissement : facette par établissement : nom de l’établissement et nombre d’occurrences

discipline : facette par discipline : libellé de la discipline et nombre d’occurrences

ecoleDoctorale : facette par école doctorale : libellé de l’école doctorale et nombre d’occurrences

langueThese : facette par langue : code de la langue sur deux caractères et nombre d’occurrences

directeurThese : facette par directeur de thèse : nom et prénom du directeur et nombre d’occurrences

oaiSetSpec : facette par domaine : codes de domaine issus de la classification décimale Dewey et utilisés par la recommandation TEF

La plupart de ces paramètres sont en dur dans l’API XML. Vous pouvez cependant jouer sur :

le paramètre start pour afficher les résultats suivants

le paramètre sort pour trier les résultats

Exemple :

Vous cherchez toutes les thèses soutenues depuis 2010, accessibles en ligne, en français, dans le domaine de l'informatique et des sciences de l'ingénieur qui parlent de la chaine de markov.

Pour construire votre requête dans l'interface :

Vous cochez le filtre « Uniquement les thèses soutenues accessibles en ligne »,

Vous tapez chaine de markov

Puis vous affinez la liste de résultat avec :

le curseur date de soutenance de 2010 à 2011

la facette langue : français

la facette domaine : informatique et sciences de l'ingénieur

Votre URL de recherche est http://www.theses.fr/?q=chaine%20de%20markov&fq=dateSoutenance:[2010-01-01T23:59:59Z%2BTO%2B2011-12-31T23:59:59Z]&checkedfacets=langueThese=fr;oaiSetSpec=ddc:620;oaiSetSpec=ddc:004;&start=0&status=&access=accessible:oui

Ajoutez le paramètre &format=xml

http://www.theses.fr/?q=chaine%20de%20markov&fq=dateSoutenance:[2010-01-01T23:59:59Z%2BTO%2B2011-12-31T23:59:59Z]&checkedfacets=langueThese=fr;oaiSetSpec=ddc:620;oaiSetSpec=ddc:004;&start=0&status=&access=accessible:oui**&format=xml**

Modifiez le paramètre &start= pour afficher les 10 résultats suivants

http://www.theses.fr/?q=chaine%20de%20markov&fq=dateSoutenance:[2010-01-01T23:59:59Z%2BTO%2B2011-12-31T23:59:59Z]&checkedfacets=la&status=&access=accessibl e:oui&fo rmat=xm ngueThese=fr;oaiSetSpec=ddc:620;oaiSetSpec=ddc:004;**&start=10**l

Ajoutez le format &sort= pour trier différemment, par exemple par date de soutenance de la plus récente à la plus ancienne

http://www.theses.fr/?q=chaine%20de%20markov&fq=dateSoutenance:[2010-01-01T23:59:59Z%2BTO%2B2011-12-31T23:59:59Z]&checkedfacets=langueThese=fr;oaiSetSpec=ddc:620;oaiSetSpec=ddc:004;&start=0&status=&access=accessible:oui&**sort=dateSoutenance+desc**&format=xml

Format JSON

Pour récupérer les données au format json, ajoutez le paramètre &format=json à la suite de l'URL de la requête.

Format ATOM

Pour récupérer les données au format ATOM, ajoutez le paramètre &format=atom à la suite de l'URL de la requête.

SolR renvoie les résultats triés selon le critère choisi dans l'interface (pertinence, auteur, titre, discipline, date de soutenance).

Les données renvoyées sont conforme à ATOM et à une extension d'ATOM : OpenSearch
http://www.opensearch.org/Specifications/OpenSearch/1.1#OpenSearch_res [ponse_elements] Opensearch permet au flux ATOM de se comporter comme une liste de résultats : un client peut appeler la suite des résultats et reconstruire la pagination des résultats. Grâce à Opensearch, le flux ATOM de theses.fr peut être utilisé comme un web service de recherche.

3.3. Récupérer les données d'une thèse

Les données d'une thèse sont disponibles au format RDF [http://fr.wikipedia.org/wiki/Resource_Description_Framework], en format XML.

Repérez grâce à l'interface l'URL de la page de la thèse et ajoutez à l'URL le paramètre .rdf ou .xml, vous obtiendrez les données brutes concernant cette thèse.

Exemple :

Vous voulez récupérer les données concernant la thèse d'ethnologie de François Oudin. L'URL de la page est http://www.theses.fr/2009METZ015L



Ajoutez .rdf à cette URL. Vous obtenez un fichier xml en rdf.

L'ABES a fait le choix de réutiliser des vocabulaires existants notamment BIBO (Bibliographic Ontology [http://bibliontology.com/]) pour identifier les thèses, ISBD pour décrire des documents (ce vocabulaire est en cours de normalisation par l'IFLA), FOAF (Friend of a friend [http://www.foaf-project.org/]) pour identifier les personnes et les organismes, MARC RELATORS PROPERTIES [http://dublincore.org/usage/documents/relators/] pour caractériser les liens entre les personnes et organismes et les thèses, et Dublin core [http://dublincore.org/documents/dces/].


```
- <rdf:RDF>
  - <!--
      Informations sur la notice / Information about the record
    -->
  - <bibo:Document rdf:about="http://www.theses.fr/2009METZ015L">
      <foaf:primaryTopic rdf:resource="http://www.sudoc.fr/2009METZ015L/id"/>
      <dcterms:creator rdf:resource="http://www.idref.fr/033702462/id"/>
      <dcterms:created rdf:datatype="http://www.w3.org/2001/XMLSchema#date">2010-09-28T12:56:29</dcterms:created>
      <dcterms:modified rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2010-12-06T12:17:45Z</dcterms:modified>
    </bibo:Document>
  - <bibo:Thesis rdf:about="http://www.theses.fr/2009METZ015L/id">
    - <!--
        Informations sur la thèse / Information about the thesis
      -->
      <rdf:type rdf:resource="http://purl.org/ontology/bibo/Thesis"/>
    - <dc:title>
        Ethnologie du quotidien des "passionnés de moto" : construction d'un être au monde par l'investissement d'une machine
      </dc:title>
      <!--Date de soutenance / Thesis defense date-->
      <dc:date>2009</dc:date>
      <dcterms:dateAccepted>2009-06-16</dcterms:dateAccepted>
      <!--ISBD + DCMI types-->
      <dc:type>Electronic Thesis or Dissertation</dc:type>
    - <isbd:P1001>
      - <skos:Concept rdf:about="http://iflastandards.info/ns/isbd/terms/contentform/T1009">
          <skos:prefLabel xml:lang="en">text</skos:prefLabel>
        </skos:Concept>
      </isbd:P1001>
      <dc:type>Text</dc:type>
    - <isbd:P1003>
      - <skos:Concept rdf:about="http://iflastandards.info/ns/isbd/terms/mediatype/T1002">
          <skos:prefLabel xml:lang="en">electronic</skos:prefLabel>
        </skos:Concept>
      </isbd:P1003>
      <dcterms:language rdf:resource="http://id.loc.gov/vocabulary/iso639-1/fr"/>
    - <dcterms:abstract xml:lang="fr">
        Nous affirmons que la perspective consistant à aborder l'usage de la motocyclette comme figure de la prise de risque détourne du sens
```


3.4. Récupérer les données d'une personne

Les données relatives à une personne et aux thèses liées sont disponibles au format RDF [http://fr.wikipedia.org/wiki/Resource_Description_Framework], en format XML.

Repérez grâce à l'interface l'URL de la page de la personne et ajoutez à l'URL le paramètre .rdf ou .xml, vous obtiendrez les données brutes de theses.fr concernant cette personne.

Vous ne pourrez pas récupérer par ce biais le nuage de mot de la personne mais seulement la liste des thèses liées à cette personne.

Exemple :

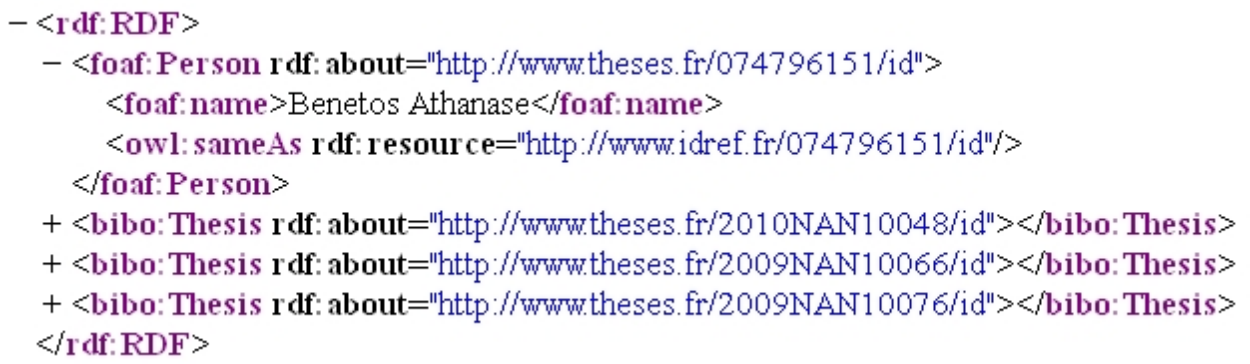
Vous voulez récupérer les données concernant Benetos Athanase. L'URL de la page est http://www.theses.fr/074796151



Ajoutez .rdf à cette URL. Vous obtenez un fichier xml en rdf.

L'ABES a fait le choix de réutiliser des vocabulaires existants, notamment FOAF (Friend of a friend [http://www.foaf-project.org/]) pour identifier les personnes, BIBO (Bibliographic Ontology [http://bibliontology.com/]) pour identifier les thèses et MARC RELATORS PROPERTIES [http://dublincore.org/usage/documents/relators/] pour caractériser les liens entre les personnes et organismes et les thèses.

Les copies d’écran qui suivent vont vous aider à interpréter les résultats. Dans celle qui suit les éléments bibo:Thesis sont fermés.



La personne Benetos Athanase possède deux identifiants, l’un issu de theses.fr et l’autre d’IdRef.

Trois thèses sont liées à cette personne. Chaque thèse est identifiée grâce à une URL unique et pérenne attribuée par theses.fr

Dans la copie d’écran qui suit les éléments décrivant la première thèse sont ouverts.

```
- <rdf:RDF>
  - <foaf:Person rdf:about="http://www.theses.fr/074796151/id">
      <foaf:name>Benetos Athanase</foaf:name>
      <owl:sameAs rdf:resource="http://www.idref.fr/074796151/id"/>
    </foaf:Person>
  - <bibo:Thesis rdf:about="http://www.theses.fr/2010NAN10048/id">
    - <dc:title>
        Nouvelles approches méthodologiques pour l'évaluation du vieillissement des gros troncs artériels par l'étude de la distensibilité artérielle et de l'analyse de la courbe de la pression artérielle chez l'homme
      </dc:title>
    - <marcrel:aut>
      - <foaf:Person rdf:about="http://www.theses.fr/148321186/id">
          <foaf:name>Salvi Paolo</foaf:name>
        </foaf:Person>
      </marcrel:aut>
    - <marcrel:ths>
      - <foaf:Person rdf:about="http://www.theses.fr/074796151/id">
          <foaf:name>Benetos Athanase</foaf:name>
        </foaf:Person>
      </marcrel:ths>
    - <marcrel:ths>
      - <foaf:Person rdf:about="http://www.theses.fr/091581443/id">
          <foaf:name>Rossignol Patrick</foaf:name>
        </foaf:Person>
      </marcrel:ths>
    - <marcrel:dgg>
      - <foaf:Organization rdf:about="http://www.theses.fr/026403390/id">
          <foaf:name>Nancy 1</foaf:name>
        </foaf:Organization>
      </marcrel:dgg>
      <dc:date>2010</dc:date>
      <dc:identifier>http://www.theses.fr/2010NAN10048/document</dc:identifier>
    </bibo:Thesis>
  + <bibo:Thesis rdf:about="http://www.theses.fr/2009NAN10066/id"></bibo:Thesis>
  + <bibo:Thesis rdf:about="http://www.theses.fr/2009NAN10076/id"></bibo:Thesis>
  </rdf:RDF>
```


Le titre de la thèse est en dc:title.

L’auteur de la thèse est dans un élément marcrel:aut contenant son nom et son identifiant dans theses.fr

Le ou les directeurs de thèse sont dans un élément marcrel:ths.

L’établissement de soutenance est dans un élément marcrel:dgg

La date de soutenance est dans dc:date

L’identifiant du document est dans dc:identifier ; cet identifiant correspond à l’URL de theses.fr suivi du numéro national de thèse et du suffixe /document uniquement si la thèse est accessible sur le web.