



Cours de programmation web

PAUL-ANTOINE BISGAMBIGLIA

2023-2024 V5, V1 DE 2016-2017

[HTTP://PAUL-ANTOINE-BISGAMBIGLIA.UNIV-CORSE.FR/](http://paul-antoine-bisgambiglia.univ-corse.fr/)

[@PABISGAMBIGLIA](#)



Licence



- ▶ Cours en licence libre
- ▶ Certaines images ont été récupérées sur google image avec le filtre « réutilisation utilisée sans but commercial »
- ▶ Certains exemples de code sont issus de sites web référencés en fin de chapitre via leurs urls

Référencement

**« AMÉLIORER LE
POSITIONNEMENT ET LA
VISIBILITÉ DE VOTRE SITE »**

Les bases

- ▶ Le référencement naturel reste un levier d'acquisition de trafic déterminant.
- ▶ Avec 93 % des expériences en ligne qui démarrent sur un moteur de recherche, et 75 % des internautes qui ne cliquent pas au-delà de la 1ère page des SERP ([source](#)).
- ▶ Pratique complexe, dans la mesure où ses règles changent tout le temps, contraignant les acteurs du SEO à s'adapter en continu.

<https://www.journaldunet.com/solutions/seo-referencement/1489251-seo-en-2020-les-grandes-tendances-du-referencement-naturel/>

Historique

- ▶ 1991, Tim Berners-Lee publie le premier site internet au monde
 - ▶ 1994 naissance des moteurs de recherche notamment Yahoo
 - ▶ 1997 naissance de Google
 - ▶ 2006 et 2009 professionnalisation du SEO (Google Analytics)
 - ▶ à partir de 2010 « Content is king »
 - ▶ Maintenant la vie privée sur le net est au cœur du débat (**Google vous connaît par cœur... et ça fait peur**)
 - ▶ Alternatives moins intrusive (Qwant ou DuckduckGo)
- ▶ Source <https://www.webmarketing-com.com/2018/02/01/69974-evolutions-seo-passe-present-futur>
 - ▶ Google : « volonté des moteurs de recherche de répondre instantanément aux besoins de l'utilisateur. Si possible, l'objectif ultime serait même de le devancer sur sa quête pour une solution. »

Référencement

- ▶ Le terme « **référencement** » (en anglais **SEO** pour *Search Engine Optimization*, traduisez Optimisation des moteurs de recherche) désigne l'ensemble des techniques permettant d'améliorer la visibilité d'un site web, à savoir :
 - ▶ **soumission** (en anglais **submission**) consistant à faire connaître le site auprès des outils de recherche ;
 - ▶ **positionnement** (en anglais **ranking**) consistant à positionner les pages d'un site en bonne position dans les pages de résultat pour certains mots-clés ;

<http://www.commentcamarche.net/contents/1267-referencement-naturel-seo-guide-pratique-complet#definition>

Référencement

- ▶ Tout d'abord, il ne sert à rien d'optimiser quoi que ce soit sur votre site si les moteurs de recherche ne peuvent pas le voir. Pour qu'un site apparaisse sur un moteur de recherche comme Google, il doit être crawlé et indexé par lui. C'est pourquoi la crawlabilité et l'indexabilité d'un site sont les éléments clés qui peuvent anéantir tous les autres efforts SEO.
- ▶ Pour améliorer la navigation et la compréhension à la fois pour les utilisateurs et les bots de crawl, vous devez construire une architecture de site bien organisée. Un site bien fait pour le SEO équivaut ici à un site agréable pour l'utilisateur. Pour parvenir à ce résultat, vous devez rationaliser la structure de votre site, et vous assurer que le contenu le plus efficace est bien disponible (4 clics maximum à partir de la page d'accueil).

Référencement

► <https://www.google.fr/trends/topcharts/>. && <https://trends.google.fr/trends/yis/2022/FR/>

Toutes les questions

- 1 Cas contact que faire ?
- 2 Pourquoi la guerre en Ukraine ?
- 3 Pourquoi la Russie attaque l'Ukraine ?
- 4 Comment voter aux élections législatives ?
- 5 Comment vérifier son inscription sur les listes électorales ?

disgambiglia@univ-corse.fr

Les "pourquoi"

- 1 Pourquoi la guerre en Ukraine ?
- 2 Pourquoi la Russie attaque l'Ukraine ?
- 3 Pourquoi les Corn Flakes ont été inventées ?
- 4 Pourquoi Nikos ne présente plus The Voice ?
- 5 Pourquoi 21 coups de canon ?



?

Les "où"

- 1 Où trouver de l'essence ?
- 2 Où voter ?
- 3 Où se trouve le Qatar ?
- 4 Où se trouve l'Ukraine ?
- 5 Où trouver de la moutarde ?

Les jeux-vidéo

- 1 Lost Ark

Les "c'est quoi"

- 1 NUPES c'est quoi ?
- 2 NFT c'est quoi ?
- 3 L'OTAN c'est quoi ?
- 4 Pass vaccinal c'est quoi ?
- 5 49.3 c'est quoi ?

Paroles des chansons

- 1 Et bam paroles
- 2 Suavemente paroles
- 3 FADE UP paroles

Les définitions

- 1 Ludique définition
- 2 NFT définition
- 3 Résilience définition
- 4 Reine consort définition
- 5 Jubilé définition

Les recettes

- 1 Clafoutis recette
- 2 Taboulé recette
- 3 Salade César recette
- 4 Recettes gombo
- 5 Recette moutarde

Référencement

- ▶ La stratégie web n'est pas seulement **un beau site, bien conçu**. Elle passe aussi par un bon contenu optimisé grâce à des mots-clés, des vrais, des bons mots-clés, ceux tapés par vos visiteurs, et référencé par les SE.
- ▶ En quoi consiste cette étape ?
- ▶ Il faut faire une recherche de mots-clés répondant aux caractéristiques suivantes :
 - ▶ tapés en volume suffisant pour justifier un travail spécifique de référencement,
 - ▶ en adéquation avec des pages ou des rubriques STABLES de votre site web,
 - ▶ laissant supposer une recherche liée à la thématique de votre site,
 - ▶ dont le travail des autres référenceurs laisse espoir d'une bonne place.

Référencement

Fonctionnent les robots d'indexation des moteurs de recherche (appelés aussi Web spiders ou crawlers ou Bots)

- ▶ ils commencent par ils vont regarder votre .htaccess puis télécharger et analyser le fichier `http://www.monsite.com/robots.txt`
- ▶ puis ils téléchargent la **racine** du site, c'est à dire l'URL `http://www.monsite.com/`
- ▶ ils analysent le contenu de cette page et en extrait notamment la liste des liens internes qu'elle contient.
- ▶ tous ces liens internes sont à leur tour téléchargés, et leur liens internes extraits
- ▶ récursivement tous ces liens sont téléchargés et analysés (s'ils sont nouveaux), jusqu'à ce que le robot n'en trouve plus de nouveau.

Référencement

Le fichier le Robots.txt

- ▶ Les explorateurs (robots) envoyés par les moteurs de recherche vont tenter d'indexer le plus d'informations possible présentées sur votre site internet.
 - ▶ dans un premier temps, ils vont regarder votre .htaccess
 - ▶ puis identifier la présence du fichier robots.txt
- ▶ Le Robots.txt est un fichier que vous pouvez produire au moyen de programmes comme Bloc-notes ou Notepad. Pour parfaire son exécution, ce document nécessite d'être placé à la racine de votre site internet.

<http://www.digitude.fr/les-tutoriels-marketing/les-tutos-seo-seo/robots-txt-rol-e-joue-t-il-pour-le-seo/>

Référencement

Le fichier le Robots.txt

- ▶ Ce fichier inclut des lignes de commande devant respecter le protocole: **RES** (Robots Exclusion Standard)
- ▶ Il contient des instructions sur les répertoires que vous souhaitez autoriser ainsi que les catégories de robots que vous voulez cibler.
- ▶ Lisible à l'adresse www.votre-site.fr/robots.txt

Il est possible de tester le fonctionnement d'un robots.txt en utilisant l'outil de test du fichier disponible dans la **Google Search Console**

Il existe un générateur <http://tools.seobook.com/robots-txt/generator/>

<http://www.digitude.fr/les-tutoriels-marketing/les-tutos-seo-seo/robots-txt-role-joue-t-il-pour-le-seo/>

Référencement

Le fichier le Robots.txt un exemple :

USER-AGENT: *

DISALLOW: /WP-LOGIN.PHP

DISALLOW: /WP-ADMIN

DISALLOW: /WP-INCLUDES

DISALLOW: /WP-CONTENT

ALLOW: /WP-CONTENT/UPLOADS

DISALLOW: */TRACKBACK

DISALLOW: */FEED

DISALLOW: */COMMENTS

DISALLOW: /CGI-BIN

DISALLOW: /*.PHP\$

DISALLOW: /*.PDF\$

DISALLOW: /*.CSS\$

DISALLOW: /*.JS\$

SITEMAP: HTTP://WWW.VOTRE-SITE.COM/SITEMAP.XML

SITEMAP: HTTP://WWW.VOTRE-SITE.COM/SITEMAP.XMLGZ

<http://www.digitude.fr/les-tutoriels-marketing/les-tutos-seo-seo/robots-txt-rol-ejoue-t-il-pour-le-seo/>

Référencement

- ▶ Un **sitemap** est un fichier qui contient une liste enrichie d'URLs internes à un site web. Ce fichier est au format XML. Il permet de fournir aux principaux moteurs de recherche des informations sur les URLs d'un site web qu'ils ne pourraient pas découvrir par eux-mêmes.
- ▶ Un sitemap peut contenir des URLs "classiques" (HTML, PDFs...), des URLs d'images et même des URLs de vidéos.
- ▶ Il y a 3 façons de générer un fichier sitemap :
 - ▶ Manuellement
 - ▶ En développant une fonction spécifique
 - ▶ En utilisant un outil de génération automatique, exemple <https://www.yakaferci.com/genere-sitemap/>

<https://www.yakaferci.com/genere-sitemap/>

Référencement

Comment informer Google de votre sitemap ?

- ▶ Contrairement au fichier Robots.txt, il ne suffit pas de mettre un sitemap sur votre site pour qu'un robot le trouve.
- ▶ En effet il n'y a pas de convention de nommage précise sur ce fichier (il peut y en avoir plusieurs, le fichier peut être zippé etc...). Vous devez par conséquent le déclarer explicitement aux moteurs de recherche auxquels vous vous intéressez particulièrement.
- ▶ Dans le cas de Google par exemple, il faut faire cette déclaration grâce aux [Webmaster tools](#).

<https://www.yakaferci.com/generer-sitemap/>

Référencement

- ▶ Les bons **mots-clés** sont les fondations de votre **référencement naturel**
- ▶ Le choix des **mots-clés** se fait page par page
- ▶ Ou les indiquer ?
 - ▶ La **balise title** est la plus importante, c'est le titre de la page que l'on retrouve en haut du navigateur, c'est également le nom du lien que l'on voit dans les résultats de recherche de google.
 - ▶ **<title>** *Ma page à moi que j'aime* **</title>**

<http://www.referencement-naturel-white-hat.fr/prestations/mots-clés-strategie-referencement/>

Référencement

- ▶ La **balise meta description** sert à décrire le site, y glisser des termes clés est recommandé.
- ▶ `<meta name="description" content="Une description de mon site contenant quelques termes clés" />`
- ▶ La **balise keywords** contient les mots clés importants du site permettant également d'être trouvé dans les moteurs de recherche
- ▶ `<meta name="keywords" content="Mot1, Mot2, Mot3, Mot4, Mot5, Mot6" />`

<http://tekcollab.imdeo.com/optimiser-son-code-html-pour-le-referencement-rappel-des-bases/>

Référencement

<META http-equiv="Content-Type" content="text/html; charset=iso-8859-1">

<META NAME="description" CONTENT=" ... ">

<META NAME="keywords" CONTENT=" ... ">

<META NAME="robots" CONTENT="x"> ou **x** peut être :

- ▶ All : carte blanche au robot qui visitera la totalité des fichiers.
- ▶ None : pas d'indexage ni de suivi des liens : le robot se limite à la page d'accueil.
- ▶ Index : Uniquement indexage de la page d'accueil OU de la rubrique.
- ▶ NoIndex : pas d'indexage : au cas où un robot passerait "par hasard" sur le site.
- ▶ Follow : assure un suivi des liens ... à condition qu'ils ne soient pas brisés !
- ▶ NoFollow : ne réalise pas de suivi des liens.

Référencement

<META NAME="Content-language" CONTENT="french">

<META NAME="Author" CONTENT=" ... ">

<META NAME="Copyright" CONTENT="Nom du copyright">

<META HTTP-EQUIV="Reply-to" CONTENT="moi@monjolisite.com">

<META NAME="Distribution" CONTENT="**Global or Local**">

<META NAME="Rating" CONTENT="**General or Mature or Restricted**">

<META HTTP-EQUIV="Refresh" CONTENT="durée en millisecondes;
URL=http://www.monsite.com/">

Référencement

► Et ensuite !

- Les balises de titre sont les <h1>, <h2>, elles ne servent pas seulement à donner une mise en forme différente aux titres, elles servent surtout à définir l'importance du titre.
 - **Attention** : vous pouvez utiliser plusieurs balises de titre H2 ou H3 ou H4 mais il est conseillé d'utiliser une seule balise H1
- Les liens doivent eux même avoir un nom pertinent, surtout si ce sont des liens internes comme pour le menu. Eviter les liens finissant par .php, mieux vaud utiliser la réécriture d'url pour avoir un .html
- Les images, La balise strong, les URL

Référencement

► Le contenu

- Soigner au maximum le contenu des premiers paragraphes des pages en utilisant des termes pertinents, la balise strong pour les mots importants. N'ayez pas 2 pages avec le même contenu. Structurer chaque page comme suit :

```
<h1>Mon titre principal</h1>
```

```
<h2><strong>Mon sous-titre</strong></h2>
```

```
<p>mon paragraphe de texte avec des <strong>mots importants</strong>  
pour améliorer le référencement</p>
```

```
<p>et on met la suite de mon texte</p>
```

Référencement

► Mots-clés :

- Le premier travail consiste à déterminer les mots-clés sur lesquels on souhaite positionner les pages de son site. Les mots-clés que l'on a en tête ne correspondent pas toujours aux mots-clés utilisés par les visiteurs, car ceux-ci ont tendance à utiliser des termes les plus courts possibles ou encore à faire des fautes d'orthographe.
- Il existe des outils permettant de comparer le volume de recherche d'un mot-clé par rapport à un autre et donnant des suggestions :
 - <http://www.google.com/trends/?hl=fr>
- Enfin, il existe des sites permettant de connaître les mots-clés des sites concurrents :
 - SEMRush.com

Référencement

- ▶ L' étape suivante consiste à faire en sorte que les principaux moteurs de recherche et en particulier Google (car il est le plus utilisé) identifient le site et viennent le parcourir régulièrement.
- ▶ Pour ce faire, il existe des formulaires en ligne permettant de soumettre son site web :
 - ▶ <https://www.google.com/webmasters/tools/submit-url?pli=1>
 - ▶ <http://www.bing.com/toolbox/webmaster>

Référencement

- ▶ **SEO : les critères les plus importants selon 140 référenceurs**
- ▶ Src :
<http://www.journaldunet.com/solution/s/seo-referencement/1195455-seo-les-criteres-les-plus-importants-selon-140-referenceurs/>
- ▶ De Virgile Juhan pour le JDN

- ▶ Comment hisser sa page web tout en haut des résultats naturels de Google ?

Référencement

Les critères qui favorisent le plus le référencement

Ra ng	Critère	Score (moyenne des notes sur 100)	Indice de confi ance (max : 100)
1	De nombreux liens provenant de sites qui font autorité pointent vers la page	92	91
2	Mots clés pertinents dans la balise Title (précisions : un contenu ou un mot clé "pertinent" répond à la requête de l'internaute)	89	95
3	Le contenu de la zone centrale de la page est pertinent : il donne une bonne réponse à la requête	81	83
4	Nombreux liens provenant de sites du même secteur / de la même thématique	81	83
5	De nombreux liens pointent vers la page	77	83

Référencement

6	Mots clés pertinents en H1	77	84
7	Fraîcheur des infos de la page sur un sujet qui le requiert	76	78
8	Site adapté au mobile (selon l'outil de Google)	74	80
9	Rapidité d'affichage du site sur mobile	73	77
10	Présence d'une balise H1	72	81
11	De nombreux liens pointent vers le site	72	81
12	Ancienneté du site (sur la même thématique)	72	78
13	Le vocabulaire employé est varié et soutenu	71	76
14	Variété des domaines pointant vers le site	70	65
15	Rapidité d'affichage du site sur desktop	68	65

Référencement

16	Présence de contenu enrichi (avec des images, des vidéos...)	68	70
17	Dans les résultats de Google, le taux de clic de la page est supérieur à la moyenne	66	62
18	Mots clés pertinents en H2, Hn	60	70
19	Présence de données structurées (issues de schema.org) sur la page	58	74
20	Présence de H2, Hn	56	69
21	Une page plus ancienne au contenu comparable redirige (via redirection 301) vers la page à référencer	54	58
22	Mots clés pertinents dans le reste de l'URL	53	63
23	La page contient des liens internes pointant vers des sites faisant autorité	53	59
24	Page en https	51	55
25	Mots clés pertinents dans le nom de domaine	48	65
26	De nombreux liens provenant des réseaux sociaux pointent vers la page (nombreux partages Facebook ou nombreux tweets pointant vers la page)	46	63
27	Mots clés pertinents dans la meta description. Rappel : c'était l'impact sur les positions dans Google qu'il faut noter, pas sur le trafic organique	38	51
28	Part importante de liens dofollow (très peu de liens nofollow). Ici il était possible de noter l'impact positif ou négatif	36	29
29	Présence de la balise "author"	22	66

Référencement

Les critères qui pénalisent le plus le référencement

Rang	Critère	Score (moyenne des notes sur 100)	Indice de confiance (max : 100)
1	Le site est considéré par Google comme piraté ou hébergeant du contenu piraté	89	80
2	Le contenu présenté au moteur est étrangement différent de celui présenté aux internautes	85	70
3	La majorité du contenu est dupliqué, pris ailleurs, sur un autre site (duplicate externe)	80	79
4	Une proportion élevée de liens issus de pages, de forums ou de blogs douteux et sans trafic pointent vers la page	75	66
5	Le contenu peut être considéré comme insuffisant, ou manquant de substance	66	72
6	Présence de publicité cachant le contenu principal	66	54

Référencement

7	Le temps de réponse du site est long (plus que la moyenne ou celui des concurrents)	64	67
8	La majorité du contenu est dupliqué, pris ailleurs, en interne, sur une autre page du site (duplicate interne)	63	59
9	Retour rapide sur les résultats de Google (l'internaute visite la page, mais revient vite sur les résultats de Google pour cliquer sur un autre résultat)	62	52
10	Taux de retour sur les résultats de Google élevé (beaucoup d'internautes reviennent sur les résultats de Google après avoir visité la page, et ils cliquent ensuite sur un autre résultat)	62	51
11	La page pèse lourd (plus que la moyenne ou celle des concurrents)	59	65
12	Une proportion élevée de liens issus de sites sans rapport thématique avec la page pointent vers la page	57	54
13	Rebond rapide des visiteurs sur la page	56	51
14	Liens non naturels ou achetés pointant vers le site	50	36

Référencement

15	Taux de rebond élevé sur la page	48	49
16	De nombreux liens avec une ancre trop optimisée pour être naturelle pointent vers la page. Ici on peut noter l'impact positif ou négatif...	44	45
17	Les mots clés visés sont excessivement utilisés dans le texte. Ici on peut aussi noter l'impact positif ou négatif.	33	56

Référencement

Référencement gratuit ou payant ?

selon [SEO Round Table](#), 80 % des utilisateurs continuent d'ignorer les annonces et des liens payants.

40 erreurs techniques

- ▶ SRC : <https://www.webmarketing-com.com/2017/08/30/61965-les-40-erreurs-techniques-les-plus-graves-en-seo>
- ▶ Quels sont les problèmes SEO les plus rencontrés en 2017 ?
- ▶ Quelles sont les erreurs plus graves ?
- ▶ 40 erreurs les plus communes de SEO classées en trois groupes :
 - ▶ crawlabilité et architecture de site,
 - ▶ SEO on-page
 - ▶ SEO technique.

40 erreurs techniques

40 ERREURS DE SEO TECHNIQUE ÉTUDE SEMRUSH 2017

Nous avons recueilli des données anonymes sur 100 000 sites web et 450 millions de pages en utilisant l'outil SEMrush Audit de site pour déterminer les erreurs et les problèmes SEO et techniques on-site les plus communs.

Vous pouvez voir ci-dessous le pourcentage de sites examinés faisant face à un certain type de problème.



Chaque problème a un niveau de gravité allant de 1 à 5, où 5 est le plus haut niveau de gravité.

CRAWLABILITÉ

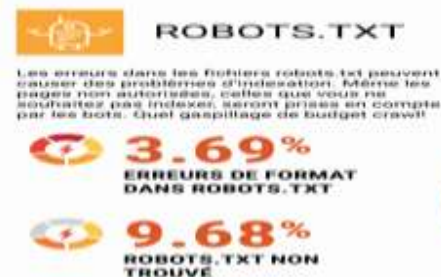
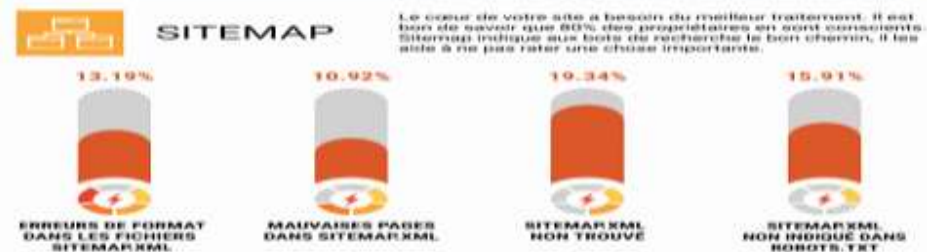
16.98%
DOMAINE WWW MAL CONFIGURÉ

5.49%
CHAÎNES DE REDIRECTION ET BOUCLES



LIENS & REDIRECTIONS

Les problèmes avec les liens et les redirections sont très divers. Certains peuvent être très dommageables, d'autres ne peuvent pas vraiment être considérés comme des erreurs. Assurez-vous que toutes les pages principales de votre site sont accessibles et n'ont pas de problèmes de redirections, car il semble qu'un site sur quatre a trop d'erreurs de lien.



OPTIMISATION ON-PAGE



CONTENU

"Toutes les pages de votre site n'exigent pas un texte parfaitement écrit" - c'est ce que les propriétaires de site peuvent penser, et considérer ces problèmes comme secondaires. Mais vous devez faire attention à ces erreurs sur les pages importantes de votre site ! Par ailleurs, un ratio texte/HTML bas peut être considéré comme un indicateur de la qualité de votre code.



53.99%
META
DESCRIPTION
DUPLIQUÉE

63.28%
META
DESCRIPTION
MANQUANTE



META DESCRIPTION

Les moteurs de recherche n'y accordent pas d'attention, les propriétaires de site non plus, c'est ce qui rend les problèmes de meta description si fréquents. Mais il est préférable de décider vous-même à quoi votre snippet va ressembler sur la SERP et de fournir une meta description pertinente, plutôt que de laisser les moteurs de recherche s'en charger.



BALISE TITLE

Si ça n'a pas une influence notable sur les rankings, ce n'est pas important. Il semblerait que les balises de titres soient victime de cette conception. Mais n'oubliez pas l'importance des balise title pour le taux de clics.

10.53%



61.01%



53.23%



26.33%



BALISE H1

Les balises H1 représentent un des indices que le propriétaire de site peut donner aux utilisateurs à propos du contenu des pages. Dommage que les balises H1 soient trop souvent oubliées.



63.25%
BALISE H1
MANQUANTE

35.11%
CONTENU
DUPLIQUÉ DANS
LA H1 ET LE TITRE

59.27%
TITRES H1
MULTIPLES



IMAGES

Beaucoup de propriétaires de site oublient d'optimiser les images. Cela peut s'expliquer: les problèmes d'image n'ont pas d'impact sur le positionnement ou la santé globale d'un site. Mais vos utilisateurs risquent d'être déçus, pensez-y !



SEO TECHNIQUE



VITESSE DE LA PAGE

1,2,3,4... Personne ne veut attendre. Si la vitesse de chargement d'une page est trop lente, dites au revoir à vos utilisateurs et aux rankings.



23.16%

CHARGEMENT DE PAGE LENT



0.98%

GRANDE TAILLE DE PAGE HTML



TECHNOLOGIE ANCIENNE

Vous ne pouvez pas lire le latin, les moteurs de recherche ne peuvent pas lire la technologie ancienne, ce qui peut conduire à des problèmes d'indexation. Certaines vieilles habitudes doivent être abandonnées.

0.38%



UTILISATION DE CADRES

5.78%



CONTENU FLASH UTILISÉ

9.03%



DOCTYPE NON DÉCLARÉ



MOBILE

La recherche mobile est en plein boom et il est agréable de constater que peu de sites ont des problèmes de SEO mobile. Mais sachez quand même que tous les sites n'ont pas une version mobile et encore moins des pages AMP.



0.08%

BALISES CANONIQUES MANQUANTES DANS LES PAGES AMP



0.66%

VIEWPORT NON CONFIGURÉ

- ▶ La vitesse de page est un facteur important de ranking, elle dépend à la fois du serveur et de la performance de la page. Elle exerce évidemment une grosse influence sur le taux de rebond. Vous devez donc optimiser le HTML, réduire les scripts et les styles, et faire en sorte de réduire le poids de la page à son minimum.
- ▶ **HTML surchargé** – 0,98 %. Une des manières d'optimiser le code HTML est d'utiliser les outils de compression comme gzip ou deflate. La compression du HTML, CCS et Javascript peut produire un excellent effet sur la vitesse de chargement, mais il y a des inconvénients : le réglage est compliqué et il y a parfois des problèmes avec certains navigateurs.

20 conseils pour optimiser votre contenu et être sur la première page de Google

SOURCE :

[HTTP://WWW.JOURNALDUNE
T.COM/SOLUTIONS/EXPERT/6
8237/20-CONSEILS-POUR-
OPTIMISER-VOTRE-CONTENU-
ET-ETRE-SUR-LA-PREMIERE-
PAGE-DE-GOOGLE.SHTML](http://www.journaldunet.com/solutions/expert/68237/20-conseils-pour-optimiser-votre-contenu-et-etre-sur-la-premiere-page-de-google.shtml)

20 conseils

1. Rafraîchissez votre contenu existant
2. Créer un [blog](#)
3. Infographies et chiffres
4. "Checklist"- Listes de contrôle - Listes de vérification
5. Documents imprimables
6. Interviews -Quiz
7. Livres blancs- en anglais "White papers »
8. Vidéos
9. Chronologie
10. Appâts de l'égo !

1. Répondre aux questions et aux commentaires des gens
2. Organiser un concours
3. Créer des cartes annotées (map)
4. Profiter des événements et des actualités
5. Journée dans la vie
6. Nouvelles et actualités
7. Guide pratique :l'expression magique "Comment »
8. Traits, caractéristiques ou particularités réguliers (ères)
9. bulletine-mail- newsletters
10. Expérience /recherche

20 conseils

1. **Rafraîchissez votre contenu existant**
2. Créer un [blog](#)
3. Infographies et chiffres
4. "Checklist"- Listes de contrôle - Listes de vérification
5. Documents imprimables
6. Interviews -Quiz
7. Livres blancs- en anglais "White papers »
8. Vidéos
9. Chronologie
10. Appâts de l'égo !

Les critères à respecter seront classés sous 2 catégories :

- ▶ Pour l'internaute :Contenu riche et intéressant,
 - ▶ Contenu sous forme de paragraphes avec des petits titres attractifs,
 - ▶ Belle imagerie,
 - ▶ Bonne [ergonomie](#),
 - ▶ Liens claires pour que l'internaute trouve facilement ce qu'il recherche.
- ▶ Pour les moteurs de recherche :
 - ▶ Contenu riche et relativement grand,
 - ▶ Contenu original et unique (non-dupliqué ou copié des autres sites web),
 - ▶ Présence des balises "H",
 - ▶ Optimisation des titres et des descriptions sur chaque page à part une formation référencement naturel vous sera très utile afin d'apprendre toutes les techniques d'optimisation et du "ON Page",
 - ▶ Cohérence totale entre titres, descriptions et contenu de la page,
 - ▶ Des images bien nommées (n'oubliez pas les attributs ALT)

20 conseils

- « Comment » est l'expression ou le mot magique tapé environ 400 000 000 (quatre cents millions) de fois par mois. Déjà, cet article commence par "comment" !

1. Répondre aux questions et aux commentaires des gens
2. Organiser un concours
3. Créer des cartes annotées (map)
4. Profiter des événements et des actualités
5. Journée dans la vie
6. Nouvelles et actualités
7. **Guide pratique : l'expression magique "Comment »**
8. Traits, caractéristiques ou particularités réguliers (ères)
9. bulletine-mail- newsletters
10. Expérience /recherche

Bilan

- ▶ Tout marche bien
- ▶ Un bon contenu mis en valeur via
 - ▶ Balise Titre, meta, h1, h2, h3, strong
 - ▶ Des mots clés biens sélectionnés
- ▶ Un fichier robot.txt
- ▶ Un site map
- ▶ Misez sur des mots clés en longue traîne!
- ▶ N'oubliez pas l'intelligence artificielle !
- ▶ Le SEO Mobile a pris le dessus !
- ▶ Un classique : le temps de chargement de votre site
- ▶ Cependant, n'oubliez pas d'appliquer certaines règles bien connues.

Source : <http://www.journaldunet.com/solutions/expert/68116/5-techniques-seo-a-mettre-absolument-en-oeuvre-pour-2018.shtml>

Règles bien connues.

- ▶ Allégez le poids de vos images en les compressant, réduisez les dimensions et préférez le JPEG pour vos photos et le PNG
- ▶ Comprimez les fichiers JavaScript et le CSS
- ▶ En résumé, votre site internet doit être simple et se charger rapidement.

Tendances 2019

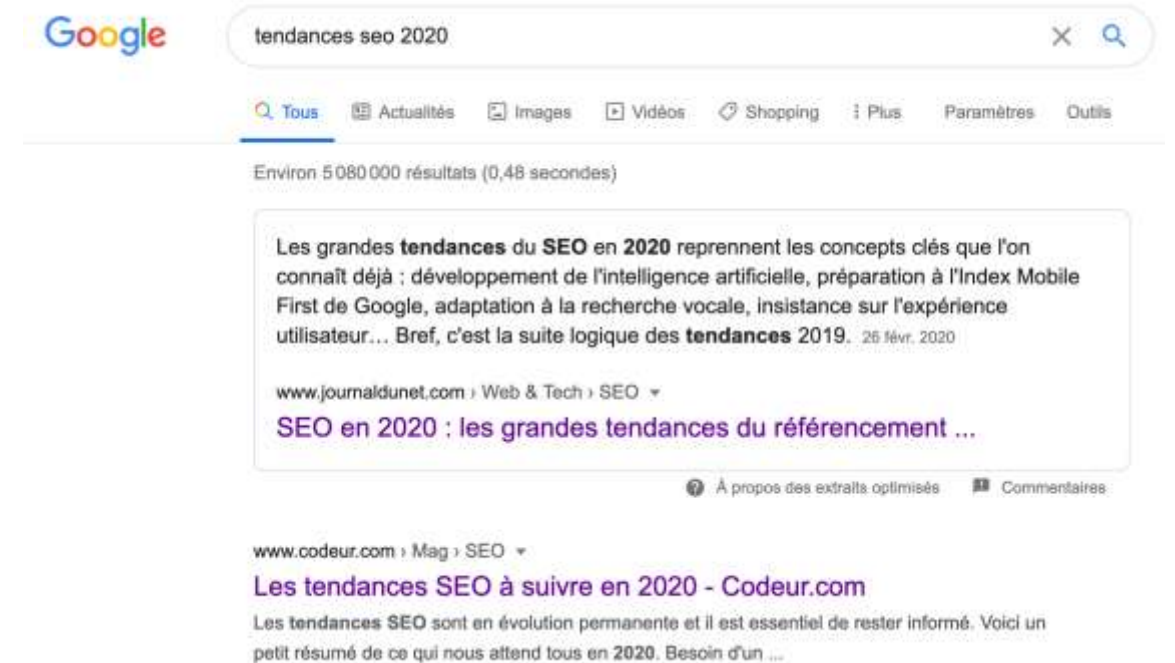
- ▶ <https://www.journaldunet.com/solutions/seo-referencement/1420823-seo-ce-que-google-vous-reserve-pour-2019-et-comment-s-y-preparer/>
- ▶ SEO : ce que Google vous réserve pour 2019 et comment s'y préparer
 - ▶ recherche visuelle, résultats personnalisés, JavaScript et mobile first indexing
 - ▶ recherche visuelle, vocale (<https://schema.org/>), toujours plus orienté smartphone (AMP, la PWA et le [JavaScript](#))

Tendances 2020

- ▶ <https://www.journaldunet.com/solutions/seo-referencement/1489251-seo-en-2020-les-grandes-tendances-du-referencement-naturel/>
- ▶ tendances du SEO en 2020 reprennent les concepts clés que l'on connaît déjà : développement de l'intelligence artificielle, préparation à l'Index Mobile First de Google, adaptation à la recherche vocale, insistance sur l'expérience utilisateur

Tendances 2020

- <https://www.journauldunet.com/solutions/seo-referencement/1489251-seo-en-2020-les-grandes-tendances-du-referencement-naturel/>



Protocole HTTP/3

- ▶ <https://www.codeur.com/blog/tendances-seo-2020/>
- ▶ Il existe 2 protocoles majeurs, le protocole http 1.1 et le protocole http/2, qu'il est obligatoire pour les professionnels du référencement de connaître (et utiliser). Pourtant, seulement 35 % des sites Web l'utilisent à l'heure actuelle. Http/2 permet de charger simultanément des objets comme des images ou des scripts, alors que l'ancien http 1.1 charge tout l'un après l'autre. Cela signifie que le nouveau protocole est beaucoup plus rapide dans le chargement des sites Web.

Le format WebP

- ▶ <https://www.codeur.com/blog/tendances-seo-2020/>
- ▶ JPG, PNG, GIF sont des formats dépassés et ont tous été remplacés par un seul : [WebP](#), qui est développé par Google.
- ▶ WebP est souvent plus de 50% plus léger que les images compressées des autres formats.

LIENS

<http://www.commentcamarche.net/contents/1267-referencement-naturel-seo-guide-pratique-complet#definition>

<http://www.referencement-naturel-white-hat.fr/prestations/mots-cles-strategie-referencement/>

<http://tekcollab.imdeo.com/optimiser-son-code-html-pour-le-referencement-rappel-des-bases/>

<http://docanski.free.fr/valastuc/balises1.htm>

<https://www.webmarketing-com.com/2017/08/30/61965-les-40-erreurs-techniques-les-plus-graves-en-seo>

<https://www.codeur.com/blog/tendances-seo-2020/>

TD 1 : moteur d'indexation

- Réalisation d'un moteur d'indexation :
 - Il faut lire une url, télécharger le contenu, et en extraire :
 - un tableau d'url
 - un tableau de mot
 - Il faut pouvoir évaluer le temps de téléchargement, et
 - Compter les occurrences des différents mots (mots clés)
 - Récursivement, il faut appliquer la même fonction à toutes les adresses du tableau d'URL
- <https://www.php.net/manual/fr/book.curl.php>

```
import requests
from bs4 import BeautifulSoup

url = 'http://www.blog.pythonlibrary.org/'

def get_articles():
    """
    Get the articles from the front page of the blog
    """
    req = requests.get(url)
    html = req.text
    soup = BeautifulSoup(html, 'html.parser')
    pages = soup.findAll('h2')

    articles = {}
    for i in pages:
        article = i.a
        s = '{title}: {url}'.format(
            title=article.text.strip(),
            url=article['href'])
        articles[article] = s
    print(s)

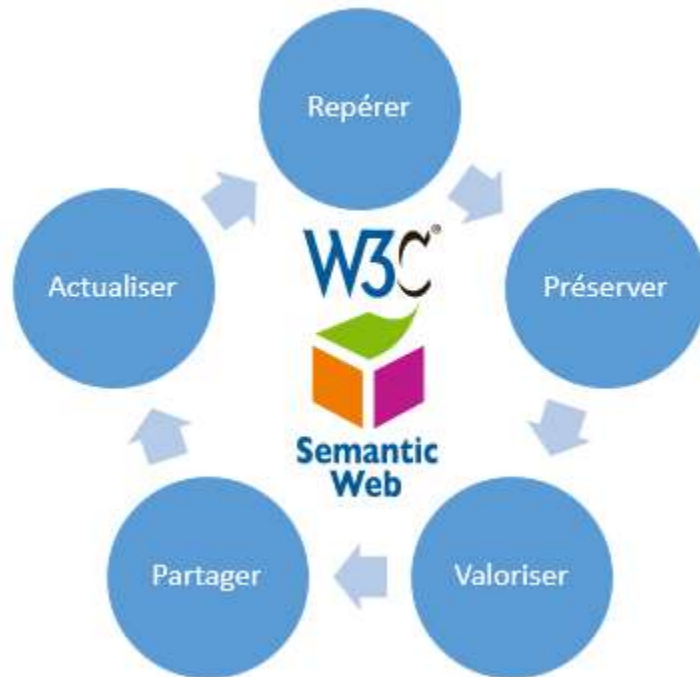
    return articles

if __name__ == '__main__':
    articles = get_articles()
```

PARTIE

WEB SÉMANTIQUE

Définition



- Le Web sémantique (plus techniquement appelé « le Web de données ») permet aux machines de comprendre la sémantique, la signification de l'information sur le Web.

- Src : <http://jplu.developpez.com/tutoriels/web-semantique/introduction/>

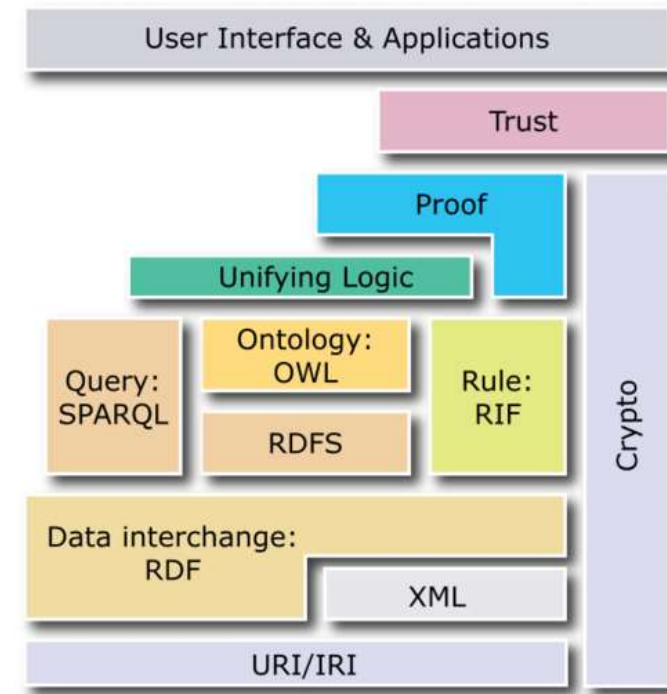
Src : <http://www.cotechnoe.com/blog-web-semantique/2016/05/24/sept-raisons-pour-mettre-la-technologie-du-web-semantique-au-centre-dune-politique-du-numerique/>

Définition

- ▶ Src : <http://jplu.developpez.com/tutoriels/web-semantique/introduction/>
- ▶ Le WS étend le réseau des hyperliens entre des pages Web classiques par un réseau de lien entre données structurées permettant ainsi aux agents automatisés d'accéder plus intelligemment aux différentes sources de données contenues sur le Web et, de cette manière, d'effectuer des tâches (recherche, apprentissage, etc.) plus précises pour les utilisateurs. Le terme a été inventé par **Tim Berners-Lee**, co-inventeur du Web et directeur du W3C, qui supervise l'élaboration des propositions de standards du Web sémantique.

Les techno

- ▶ RDF (Ressource Description Framework)
- ▶ RDFS (Ressource Description Framework Schema)
- ▶ Ontologies
- ▶ OWL (Web Ontology Language)
- ▶ SPARQL



Src : https://fr.wikipedia.org/wiki/Web_s%C3%A9mantique

Une ontologie c'est quoi ?

https://www.canal-u.tv/video/inria/introduction_aux_ontologies.20327

Historique

<http://jplu.developpez.com/tutoriels/web-semantic/introduction/>

- 1994 : naissance du terme « web sémantique »
- 1997 : premier document de recommandation
- 1998 : document de référence (Tim Berners-Lee)

*I have a dream for the Web [in which computers]
become capable of analyzing all the data on the
Web*

Objectifs

<http://jplu.developpez.com/tutoriels/web-semantic/introduction/>

- Permettre aux utilisateurs d'utiliser la totalité du potentiel du Web !
 - les pages Web actuelles sont conçues pour être lisibles par des êtres humains
- Pour découpler les usages, faciliter et automatiser certaines tâches
 - Les machines doivent pouvoir lire (et comprendre !) les pages Web

☹ critiques

<http://jplu.developpez.com/tutoriels/web-semantic/introduction/>

- Plus de données, mieux documentées = plus d'information notamment sur l'utilisateur (vie privé)
- Lourd !
 - les langages utilisés pour le Web sémantique sont très verbeux, car dérivés du XML
 - l'écriture d'ontologies est souvent très problématique, car elle exige une spécialisation dans un domaine particulier et, lorsque l'on ne maîtrise pas ce domaine, elle devient très difficile à créer

(rapide intro aux) Ontologies

<http://jplu.developpez.com/tutoriels/web-semantic/introduction/>

L'ontologie est la base de ce que l'on appelle la représentation des connaissances. Ce domaine est né de la volonté des chercheurs de représenter diverses connaissances du monde actuel, de façon à ce qu'elles soient utilisables par des ordinateurs, pour qu'ils puissent effectuer des raisonnements sur ces connaissances. Ces connaissances sont exprimées sous forme de symboles auxquels on donne une « sémantique » (un sens).

(rapide intro aux) Ontologies

<http://jplu.developpez.com/tutoriels/web-semantic/introduction/>

- *Recherche sémantique :*
- ▶ on construit ce que l'on appelle des bases de connaissances qui sont constituées :
 - ▶ d'une **ontologie** : une collection de classes et de relations (que nous appellerons propriété pour les mêmes raisons que concept) entre ces classes ;
 - ▶ de **règles** : une expression de contraintes sur les propriétés et les classes de l'ontologie ;
 - ▶ de **faits** : des instances de l'ontologie.

(rapide intro aux) Ontologies

<http://jplu.developpez.com/tutoriels/web-semantique/introduction/>

- A lire http://www.technolangue.net/imprimer.php3?id_article=280

C'est quoi

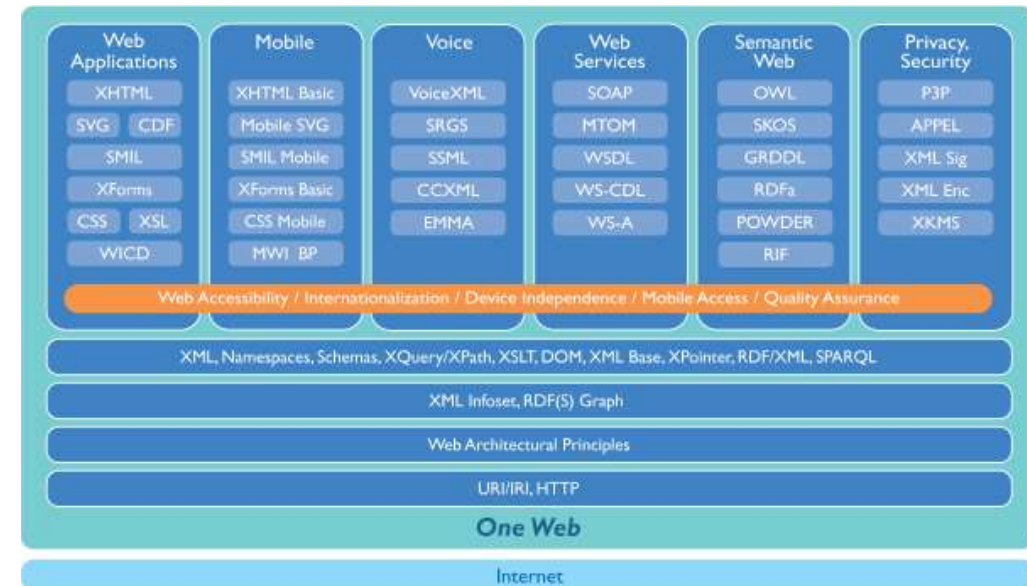
<http://www.cotechnoe.com/blog-web-semantique/2016/05/24/sept-raisons-pour-mettre-la-technologie-du-web-semantique-au-centre-dune-politique-du-numerique/>

1. Le web sémantique est une technologie mure, éprouvée ouverte, non-propriétaire, gratuite, libre de droits et sécuritaire
2. Les informations, les données et les métadonnées sont interoperables
3. Elle génère des bénéfices puisqu'elle se fonde sur le principe de réutilisabilité des vocabulaires
4. Le web sémantique une technologie au centre de l'économie du savoir
5. Le web sémantique offre les principes de gouvernance aux données ouvertes
6. Le web sémantique est un médium de valorisation d'une culture
7. Web sémantique est une plateforme d'Intelligence artificielle et moteur de l'innovation

Le web sémantique est une technologie ?

- ▶ « le web et le web sémantique vont s'imposer de plus en plus dans nos vies »
- ▶ technologie normalisée par le W3C
- ▶ les vocabulaires RDF, RDFS, OWL, SKOS font partie des standards qui permettent d'exploiter la puissance du web.

<https://www.w3.org/2004/10/RecsFigure-Smaller.png>



<http://www.cotechnoe.com/blog-web-semantique/2016/05/24/sept-raisons-pour-mettre-la-technologie-du-web-semantique-au-centre-dune-politique-du-numerique/>

Les informations, les données et les métadonnées sont interopérables

- ▶ **Interopérable** : lire des données de manière indépendante des techno (apache, etc.), et des logiciels (IE, Chrome, etc.)
- ▶ Données, informations, connaissances
- ▶ Les notations interopérables à la base du web sémantique sont : le Resource Description Framework (**RDF**), le Resource Description Framework Schema (**RDFS**) et le Web Ontology Language (**OWL**)

« La ville de Montréal est une ville-intelligente »

Est-ce une donnée utilisable par un programme ?

Comment faire ?

<http://www.cotechnoe.com/blog-web-semantique/2016/05/24/sept-raisons-pour-mettre-la-technologie-du-web-semantique-au-centre-dune-politique-du-numerique/>

Les informations, les données et les métadonnées sont interoperables

- ▶ L'information doit être codée avec une notation utilisable par une machine
- ▶ Le **signifiant** (Tag) doit aussi être codé avec une notation utilisable par une machine
- ▶ Les notations utilisées pour coder l'information et son **signifiant** doivent être interoperables.

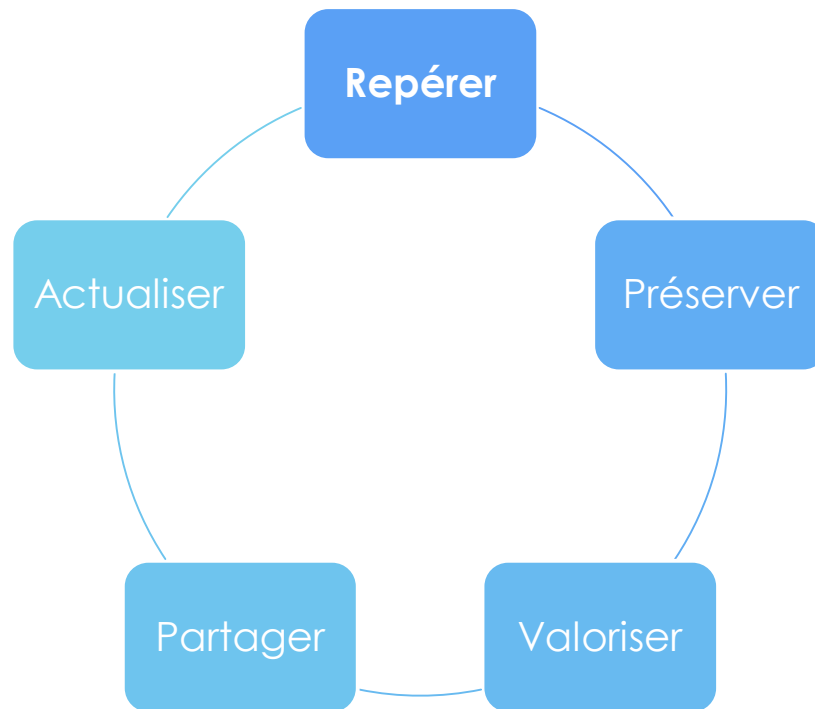
<http://www.cotechnoe.com/blog-web-semantique/2016/05/24/sept-raisons-pour-mettre-la-technologie-du-web-semantique-au-centre-dune-politique-du-numerique/>

Principe de réutilisabilité des vocabulaires

- ▶ le fichier qui entrepose « un tag » est nommé « *Ontologie* »
 - ▶ L'ontologie est pour ainsi dire la base de données du web,
 - ▶ elle déclare les données et la structure des données qu'il utilise.
- ▶ Le web est un espace de partage de point de vue, il est possible de diffuser des ontologies qui permettent de coder la signification du point de vue. C'est ce qui est appelé un '**vocabulaire**'.

<http://www.cotechnoe.com/blog-web-semantic/2016/05/24/sept-raisons-pour-mettre-la-technologie-du-web-semantic-au-centre-dune-politique-du-numerique/>

L'économie du savoir

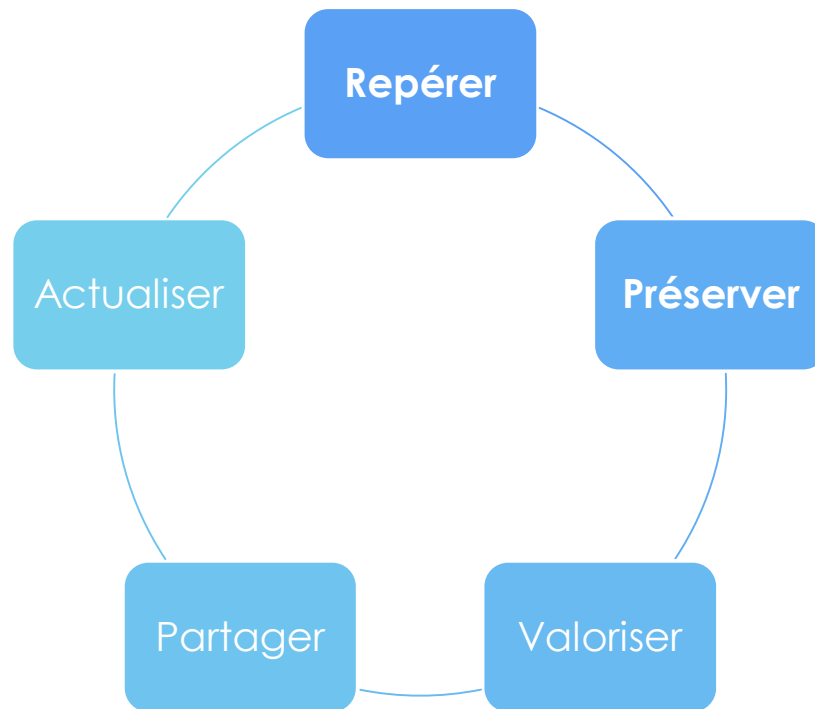


► Repérer

Le web sémantique abonde en outils qui permettent de localiser l'information. Google, Bing, Yahoo, etc. sont des outils de recherches qui utilisent entre autres le web sémantique pour indexer le web

<http://www.cotechnoe.com/blog-web-semantique/2016/05/24/sept-raisons-pour-mettre-la-technologie-du-web-semantique-au-centre-dune-politique-du-numerique/>

L'économie du savoir

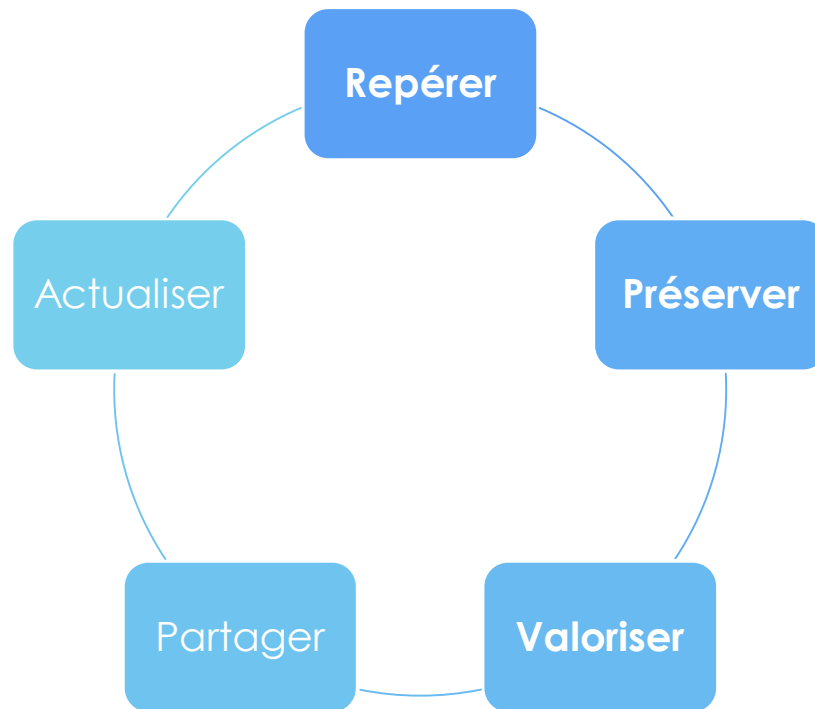


► Préserver

La préservation de la connaissance est assurée par l'ontologie. Comme il a été indiqué plus haut, l'ontologie emmagasine les données décrivant quelque chose ainsi que la structure qui permet de décrire la donnée.

<http://www.cotechnoe.com/blog-web-semantique/2016/05/24/sept-raisons-pour-mettre-la-technologie-du-web-semantique-au-centre-dune-politique-du-numerique/>

L'économie du savoir

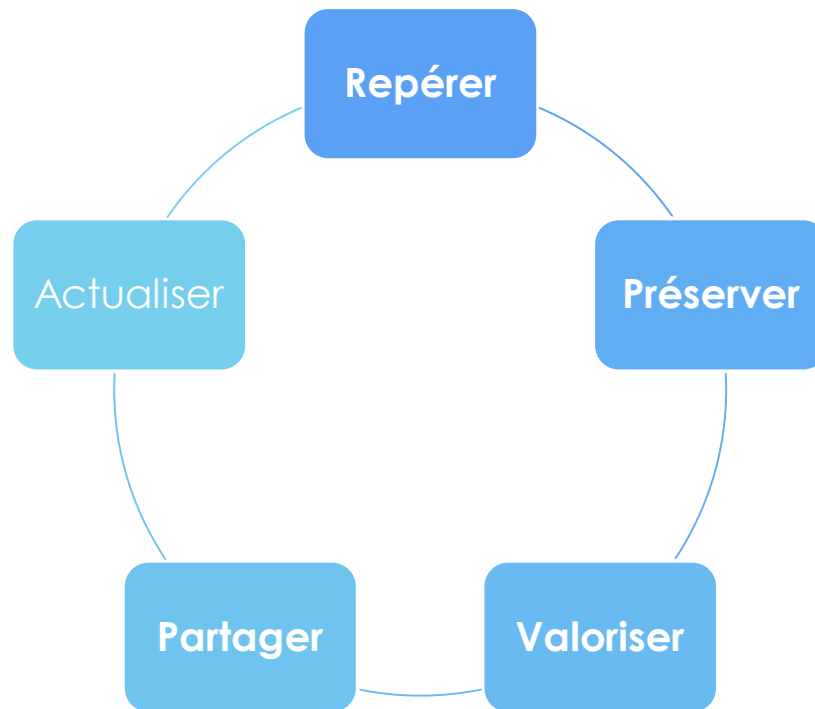


► Valoriser

L'interopérabilité de l'ontologie permet à des clients web de venir capturer les données diffusées et de réaliser des traitements spécifiques qui sont guidés par la sémantique de la donnée qui est codée dans l'ontologie

<http://www.cotechnoe.com/blog-web-semantique/2016/05/24/sept-raisons-pour-mettre-la-technologie-du-web-semantique-au-centre-dune-politique-du-numerique/>

L'économie du savoir

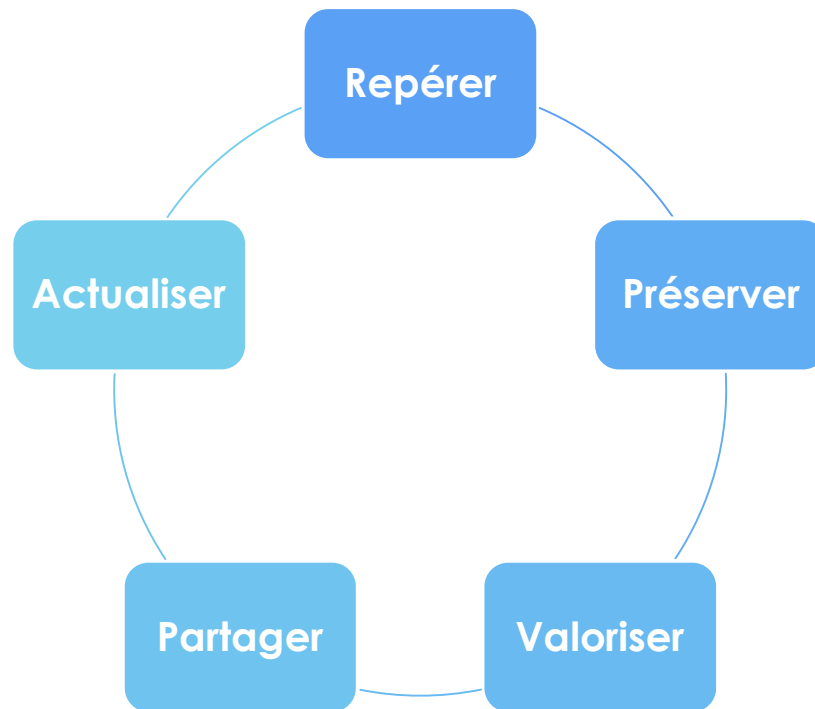


► Partager

L'ontologie qui contient la connaissance (donnée + structure de données) est partagée sur le web avec l'aide d'un serveur web (par ex. : Apache ou IIS) auquel on ajoute des services web particuliers pour faciliter l'accès aux données par des requêtes web (voir le langage SPARQL) ou faciliter la présentation de contenu de l'ontologie

<http://www.cotechnoe.com/blog-web-semantique/2016/05/24/sept-raisons-pour-mettre-la-technologie-du-web-semantique-au-centre-dune-politique-du-numerique/>

L'économie du savoir



► Actualiser

Tous changements de contenu est visible du web. Les agents du web qui perçoivent un changement dans un contenu local s'ajustent à ce contenu. De même, un contenu local s'adaptera à un changement de contenu dans le web. Une politique du numérique qui incorpore le web sémantique stimule l'économie du savoir en stimulant les interconnexions entre les données du web.

<http://www.cotechnoe.com/blog-web-semantique/2016/05/24/sept-raisons-pour-mettre-la-technologie-du-web-semantique-au-centre-dune-politique-du-numerique/>

Données ouvertes



Une donnée web est ouverte et liée si elle répond aux cinq conditions :

- ▶ License ouverte (OL-Open license). La donnée est libre de droits.
- ▶ Lisible par un ordinateur (RE -Machine readable). La donnée est lisible et traitable par un logiciel (par exemple un chiffrier Excel)
- ▶ Format de données ouvert (OF Open format). La structure de la donnée est dans une notation interopérable.
- ▶ URI pour désigner les ressources. La donnée est accessible à partir d'un *Uniform Resource Identifier* nécessaire pour localiser les ressources du web
- ▶ Lier les données (Linked data). La donnée est diffusée à partir d'un serveur web qui offre les services nécessaires à son utilisation par les agents du web.

<http://www.cotechnoe.com/blog-web-semantic/2016/05/24/sept-raisons-pour-mettre-la-technologie-du-web-semantic-au-centre-dune-politique-du-numerique/>

Plateforme d'intelligence artificielle

- L'ontologie entrepose la donnée et le schéma de la donnée pour une diffusion dans le web => base de connaissances
- traitée par un moteur d'inférence ça permet de conclure à de nouvelles connaissances



Si

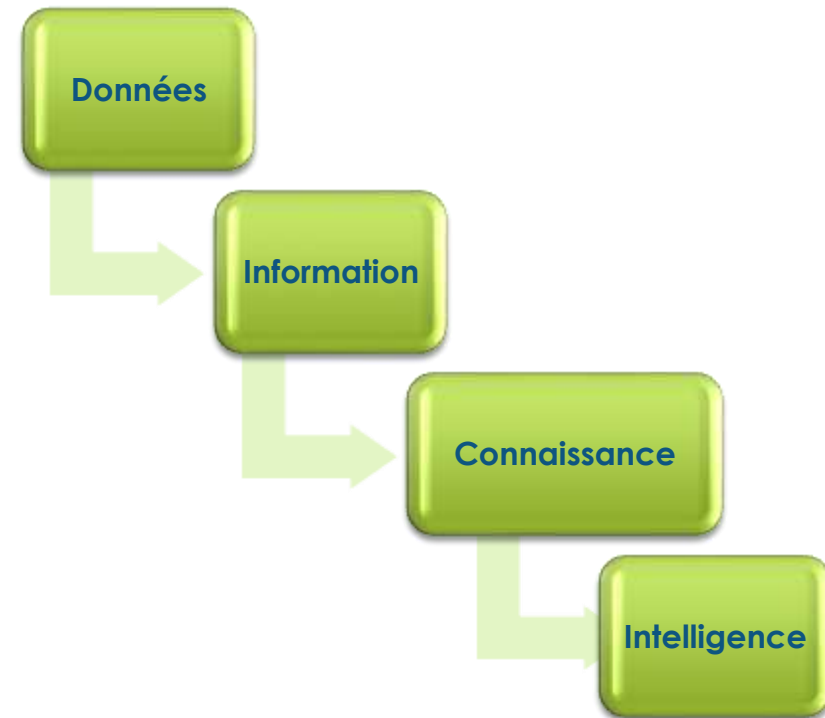
- « Snoppy est un Beagle; et qu'un Beagle est une sorte de Chien »

Alors

- « Shoppy est un Chien »

L'internet des objets

- Le web sémantique permettra d'interopérationaliser les données fournies par les objets connectés



Introduction « audio »

<https://www.youtube.com/watch?v=ZbtZqFG7LG0>

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
<html xmlns="http://www.w3.org/1999/xhtml" xml:lang="en">
  <head>
    <title>Titre</title>
  </head>
  <body>
    <div id="entete">
      <ul id="menu">
        <li><a href="/">Page d'accueil</a></li>
        <li><a href="/contact/">Contact</a></li>
      </ul>
    </div>
    <h1>Titre de l'article</h1>
    <h2>Un article de quelqu'un</h2>
    <p>
      Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nunc non sem in erat vehicula ultrices. Donec nibh felis, hendrerit
      non faucibus vel,
      Mauris id eros felis. Sed porttitor, massa at iaculis venenatis, quam quam tristique ante, ut porttitor eros enim in elit. Donec porta
      nisl lectus.
      Nunc eleifend auctor enim quis vestibulum. Cras tincidunt luctus risus eget vulputate.
    </p>
    <p>
      Le contenu de cette page est sous licence, <a href="/licence/">plus de détails.</a>
    </p>
  </body>
</html>
```

RDF

<http://tcuvelier.developpez.com/tutoriels/web-semanticue/rdfa/introduction/>

- ▶ Un humain voit très bien ce qui est un titre, qui est l'auteur, ce qu'est le menu du site, quelle est la licence de cette page. Il voit qu'il y a un menu, qui va l'aider à naviguer sur le site ; qu'il visionne la page qui porte un certain titre ; que cette page a été rédigée par quelqu'un dont le nom est écrit ; que cette page est sous une certaine licence. Il faudrait que le navigateur ait aussi accès à ces informations.
- ▶ Le navigateur voit une liste, un titre, un titre de niveau inférieur, un long paragraphe, un paragraphe plus court. Cela n'a aucun sens pour lui.
- ▶ un moteur de recherche propose une fonctionnalité de recherche de documents en fonction de leur licence. Comment donner explicitement cette information de licence ?

Pour donner ces informations complémentaires, il faut utiliser une certaine manière d'annoter le contenu : RDFa

RDF

<http://tcuvelier.developpez.com/tutoriels/web-semanticue/rdfa/introduction/>

- ▶ RDFa est un sigle signifiant RDF in HTML attributes, RDF signifiant **Resource Description Framework**. RDFa mettra son contenu dans des attributs HTML de la page.
- ▶ RDFa n'est défini que pour le XHTML. Mais la majorité des parseurs RDFa pourront reconnaître les attributs dans des pages HTML. Il n'empêche que l'ajout d'attributs dans des pages HTML ne produira pas une page conforme aux standards HTML.
- ▶ RDFa est intégré dans le standard XHTML 2.
- ▶ Plusieurs vocabulaires RDFa sont disponibles ; chaque vocabulaire permet de définir de nouveaux attributs pour marquer les données de la page. Par exemple, on peut aussi utiliser RDFa pour le réseautage social, à l'aide du vocabulaire FOAF (Friend Of A Friend).

RDF

<http://tcuvelier.developpez.com/tutoriels/web-semantic/rdfa/introduction/>

► Avant :

`<p> Le contenu de cette page est sous licence, plus de détails. </p>`

► Après

`<p> Le contenu de cette page est sous licence, plus de détails. </p>`

RDF

<http://tcuvelier.developpez.com/tutoriels/web-semantic/rdfa/introduction/>

► Avant :

```
<h1>Titre de l'article</h1>
```

```
<h2>Un article de quelqu'un</h2>
```

► Après

```
<div xmlns:dc="http://purl.org/dc/elements/1.1/">
```

```
<h1 property="dc:title">Titre de l'article</h1>
```

```
<h2 property="dc:creator">Un article de quelqu'un</h2>
```

```
</div>
```


RDF

<http://tcuvelier.developpez.com/tutoriels/web-semantic/rdfa/introduction/>

- ▶ On peut évidemment utiliser d'autres attributs sémantiques dans la même ontologie, en voici quelques-uns :
 - ▶ dc:contributor : indique un contributeur ;
 - ▶ dc:date : indique une date ;
 - ▶ dc:description : indique un texte descriptif (chapeau d'un article, par exemple).

RDF

<http://tcuvelier.developpez.com/tutoriels/web-semantic/rdfa/introduction/>

- On peut définir des informations sémantiques sur à peu près n'importe quoi, pour autant qu'une ontologie adaptée soit disponible. Si aucune n'est disponible, libre à vous d'en créer une ! On peut mélanger autant d'ontologies que nécessaire par page Web, d'où l'importance d'utiliser des espaces de noms pour stocker les ontologies, afin d'éviter que deux éléments de sémantiques différentes aient exactement la même dénomination.

RDF

<http://tcuvelier.developpez.com/tutoriels/web-semantic/rdfa/introduction/>

► Pour aller plus loin

- <http://tcuvelier.developpez.com/tutoriels/web-semantic/rdfa/introduction/>
- https://jena.apache.org/tutorials/rdf_api.html

Modèle RDF

Définition

- ▶ Un graphe RDF est un ensemble de triplets :
 - ▶ Modèles
 - ▶ Déclarations ou ressources
 - ▶ Propriétés

Exemple basé sur l'ontologie vCard

```
<rdf:RDF>  
  <rdf:Description  
    rdf:about="http://somewhere/JohnSmith/">  
    <vCard:FN>John Smith</vCard:FN>  
    <vCard:N rdf:parseType="Resource">  
      <vCard:Family>Smith</vCard:Family>  
      <vCard:Given>John</vCard:Given>  
    </vCard:N>  
  </rdf:Description>
```

<https://www.w3.org/TR/vcard-rdf/>

Modèle RDF

- ▶ Chaque arc dans un modèle RDF s'appelle une **déclaration (statements)**. Elle décrit une ressource, et comporte trois parties:
 1. Le sujet est la ressource cible (source de l'arc)
 2. Le prédicat est la propriété qui étiquette l'arc
 3. L'objet est la ressource ou le littéral visée/pointée par l'arc
- ▶ Une déclaration est parfois appelée triplet, en raison de ses trois parties.

Lien RDFa 1.1 Primer - Third Edition

► <https://www.w3.org/TR/xhtml-rdfa-primer/#using-rdfa> à lire

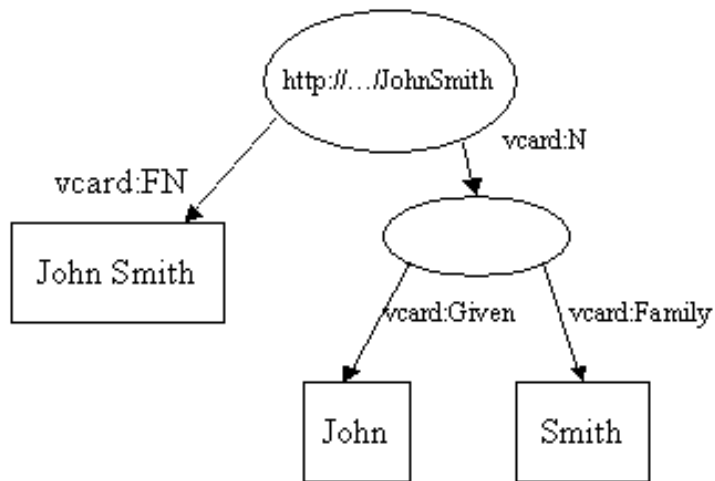
API RDF : Jena

- ▶ JENA est une API java pour manipuler des modèles de données RDF
- ▶ https://jena.apache.org/tutorials/rdf_api.html

```
Resource johnSmith =  
    model.createResource(personURI)  
        .addProperty(VCARD.FN, fullName);
```

```
// some definitions  
static String personURI = "http://somewhere/JohnSmith";  
static String fullName = "John Smith";  
  
// create an empty Model  
Model model = ModelFactory.createDefaultModel();  
  
// create the resource  
Resource johnSmith = model.createResource(personURI);  
  
// add the property  
johnSmith.addProperty(VCARD.FN, fullName);
```

API RDF : Jena



```
// some definitions
String personURI  = "http://somewhere/JohnSmith";
String givenName  = "John";
String familyName = "Smith";
String fullName   = givenName + " " + familyName;
// create an empty Model
Model model = ModelFactory.createDefaultModel();
// create the resource
// and add the properties cascading style
Resource johnSmith
    = model.createResource(personURI)
        .addProperty(VCARD.FN, fullName)
        .addProperty(VCARD.N,
            model.createResource()
                .addProperty(VCARD.Given, givenName)
                .addProperty(VCARD.Family, familyName));
```


API RDF : Jena

```
// list the statements in the Model
```

```
StmtIterator iter = model.listStatements();
```

```
// print out the predicate, subject and object of each statement
```

```
while (iter.hasNext()) {
```

```
    Statement stmt    = iter.nextStatement(); // get next statement
```

```
    Resource subject  = stmt.getSubject();    // get the subject
```

```
    Property predicate = stmt.getPredicate(); // get the predicate
```

```
    RDFNode object    = stmt.getObject();    // get the object
```

```
    System.out.print(subject.toString());
```

```
    System.out.print(" " + predicate.toString() + " ");
```

```
    if (object instanceof Resource) {
```

```
        System.out.print(object.toString());
```

```
    } else {
```

```
        // object is a literal
```

```
        System.out.print(" \"" + object.toString() + "\"");
```

```
    }
```

```
    System.out.println(" .");
```

```
}
```

API RDF en php

- ▶ <http://www.easyrdf.org/>
- ▶ <http://wifo5-03.informatik.uni-mannheim.de/bizer/rdfapi/tests.html>

```
<?php
```

```
require 'vendor/autoload.php';
```

```
$foaf = new  
EasyRdf_Graph("http://njh.me/foaf.rdf");
```

```
$foaf->load();
```

```
$me = $foaf->primaryTopic();
```

```
echo "My name is: ".$me->  
get('foaf:name')."\\n";
```

SPARQL

<https://web-semantic.developpez.com/tutoriels/>

- ▶ SPARQL est un langage de requête et un protocole pour accéder au RDF conçu par le groupe de travail du W3C RDF Data Access.
- ▶ SPARQL est « orienté données » en ce sens qu'il interroge uniquement les informations détenues dans des modèles ;
- ▶ Vidéos :
 - ▶ <https://www.youtube.com/watch?v=FvGndkpa4K0>
 - ▶ https://www.canal-u.tv/video/inria/demos_autour_de_sparql.20324

SPARQL

<https://web-semantic.developpez.com/tutoriels/>

- ▶ <https://web-semantic.developpez.com/tutoriels/jena/arq/introduction-sparql/>
- ▶ <http://jplu.developpez.com/tutoriels/web-semantic/nouveautes-sparql-1-1/>
- ▶ <https://www.w3.org/TR/rdf-sparql-query/>

SPARQL

<https://web-semantic.developpez.com/tutoriels/>

- SPARQL requête des graphes RDF

- Exemple :

```
SELECT ?x
```

```
WHERE
```

```
{ ?x <http://www.w3.org/2001/vcard-rdf/3.0#FN> "John  
Smith" }
```

```
-----  
| x |  
-----  
| <http://somewhere/JohnSmith/> |  
-----
```

@prefix vCard: <http://www.w3.org/2001/vcard-rdf/3.0#> .
 @prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
 @prefix : <#> .

```
<http://somewhere/MattJones/>
  vCard:FN "Matt Jones" ;
  vCard:N [ vCard:Family
            "Jones" ;
            vCard:Given
            "Matthew"
          ] .
```

```
<http://somewhere/RebeccaSmith/>
  vCard:FN "Becky Smith" ;
  vCard:N [ vCard:Family
            "Smith" ;
            vCard:Given
            "Rebecca"
          ] .
```

```
<http://somewhere/JohnSmith/>
  vCard:FN "John Smith" ;
  vCard:N [ vCard:Family
            "Smith" ;
            vCard:Given
            "John"
          ] .
```

```
<http://somewhere/SarahJones/>
  vCard:FN "Sarah Jones" ;
  vCard:N [ vCard:Family
            "Jones" ;
            vCard:Given
            "Sarah"
          ] .
```



SPARQL

- ▶ WHERE => recherche de correspondance
- ▶ ?x => variable
- ▶ <> => URI
 - ▶ " " littéral ordinaire
 - ▶ ^^ littéraux typés
 - ▶ @ étiquettes de langues

SPARQL

x	name
<http://somewhere/RebeccaSmith/>	"Becky Smith"
<http://somewhere/SarahJones/>	"Sarah Jones"
<http://somewhere/JohnSmith/>	"John Smith"
<http://somewhere/MattJones/>	"Matt Jones"

SELECT ?x ?fname

WHERE {?x <http://www.w3.org/2001/vcard-rdf/3.0#FN>
?fname}

SPARQL 1.1

jplu.developpez.com/tutoriels/web-semantic/nouveautes-sparql-1-1

- ▶ Les projections
- ▶ Les négations
- ▶ L'agrégation
- ▶ Les chemins
- ▶ Les requêtes imbriquées
- ▶ Inférence sur les données
- ▶ Les requêtes distribuées

<http://dbpedia.org/sparql>

```
SELECT ?name
WHERE {
    ?x foaf:givenName ?name .
    OPTIONAL { ?x foaf:birthday ?date } .
    FILTER (!bound(?date))
}
```

Support d'intégration avec RDF

- ▶ R2RML => RDB to RDF Mapping Language
- ▶ CSVW => <https://www.w3.org/ns/csvw>
- ▶ GRDDL => <https://www.w3.org/TR/grddl/>
- ▶ LDP => <https://www.w3.org/TR/ldp/>

Liens

- ▶ <https://www.youtube.com/watch?v=ZbtZqFG7LG0>
- ▶ <http://www.enssib.fr/le-dictionnaire/web-semantique>
- ▶ <http://www.futura-sciences.com/tech/definitions/internet-web-semantique-3993/>
- ▶ <https://web-semantique.developpez.com/tutoriels/>
- ▶ <http://www.cotechnoe.com/blog-web-semantique/2016/05/24/sept-raisons-pour-mettre-la-technologie-du-web-semantique-au-centre-dune-politique-du-numerique/>
- ▶ <http://www.agencereferencement-webmarketing.com/web-semantique-web-3/>

vidéo

- ▶ https://www.canal-u.tv/video/msh_val_de_loire/introduction_au_web_semantique.41925