

Activité 4 le fonctionnement d'un moteur de recherche

vidéo : lienmini.fr/1046-307 (2mns) lienmini.fr/3389-309 (26s)

1. Expliquer ce que font en permanence les moteurs de recherche . (doc1)

Al'aide de robots, les moteurs de recherche explorent en permanence le web. A chaque fois qu'ils trouvent une nouvelle page, ils l'indexent et la classent.

2. Expliquer les principaux critères de classement des résultats d'un moteur de recherche . (doc1)

Les résultats d'un moteur de recherche sont classés en fonction de la pertinence des pages affichées, de leur popularité, de leur taux de visite, du nombre de liens entrants ou sortants et des mots-clés qu'elles contiennent.

3. A partir du document 2 , noter le type de résultat fourni en premier, par le moteur de recherche.

Le premier résultat fourni en premier est une annonce commerciale.

Doc 1 Le fonctionnement des moteurs de recherche

Exploration du Web

Des robots parcourent les pages Web. Indexation et classement des pages Web.

Recherche de l'internaute

Doc 2 Les premiers résultats d'un moteur de recherche

musique streaming

Environ 346 000 000 résultats (0,45 secondes)

Toute la musique en streaming | 3 mois offerts | deezer.com
www.deezer.com/

La musique en illimité, partout, tout le temps 3 mois offerts, profitez-en ! Profitez de plus de 56 millions de titres en illimité, partout, sans publicité. 56M de Titres Disponibles. Son Haute Qualité. Mobiles. PC/Mac. Tablettes. Types: Rap & Hip Hop, Pop, Rock, Soul.

Offre Deezer Famille
Profitez de 6 profils Premium pour seulement 14,99€/mois avec Deezer !

Découvrez Deezer
Qui sommes nous ?
Nos fonctionnalités

Streaming Musique | Qobuz | Streaming Illimité | qobuz.com
www.qobuz.com/

Ecoutez toutes les musiques que vous aimez en illimité. Sans engagement. Qobuz est la seule

Le référencement payant permet aux annonceurs commerciaux d'apparaître en tête de liste dans les résultats d'une requête, fournie par un moteur de recherche.

VIDÉO WEB
Les moteurs de recherche
lienmini.fr/1046-307

4. Regarder la vidéo <https://www.youtube.com/watch?v=DVG8jGFZJig> (9mns)

DOC 3 L'optimisation du référencement

Le **référencement naturel** ou SEO (*Search Engine Optimization*) permet d'améliorer le classement d'un site Web dans les résultats de recherche. Pour cela, les éditeurs de sites doivent notamment bien choisir les mots-clés dans leur contenu rédactionnel et les placer à des endroits stratégiques. Le référencement naturel est à distinguer du référencement payant des annonces et des liens sponsorisés.

Une page optimisée pour le référencement

The screenshot shows a website header with navigation tabs: 'compétences', 'éducation', 'numérique', 'évaluation', 'compas', and 'pix'. An orange arrow points to the 'pix' tab with the label 'Tags (liens)'. The main content area features the 'pix' logo and a paragraph of text. An orange box highlights the text '5 grands domaines' with the label 'Mots-clés de l'article'. Below the main text, there is a section titled 'L'objectif de PIX...' with another orange box highlighting '5 grands domaines' and '16 compétences', labeled 'Maillage interne, lien renvoyant à un autre article'.

Que permet de faire le référencement naturel ou SEO (Search Engine Optimization) ?

Le recensement SEO permet d'améliorer le classement d'un site Web dans les résultats de recherche.

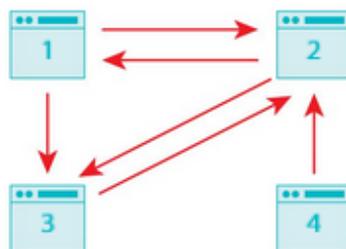
Comment peut-on optimiser le référencement naturel d'un site web ?

Il faut choisir les mots clés et les placer à des endroits stratégiques et mettre davantage de liens hypertextes.

25 Le surf au hasard

CAPACITÉ : Mener une analyse critique des résultats d'un moteur de recherche

Dans ce graphe de liens entre quatre pages Web, la flèche allant de 1 vers 3 signifie que la page 1 référence la page 3 et l'absence de flèche de 3 vers 4 signifie que la page 3 ne référence pas la page 4.



Si l'internaute part de n'importe quelle page :

1. Citer la page qui a le moins de chance de recevoir de visites et dites pourquoi.

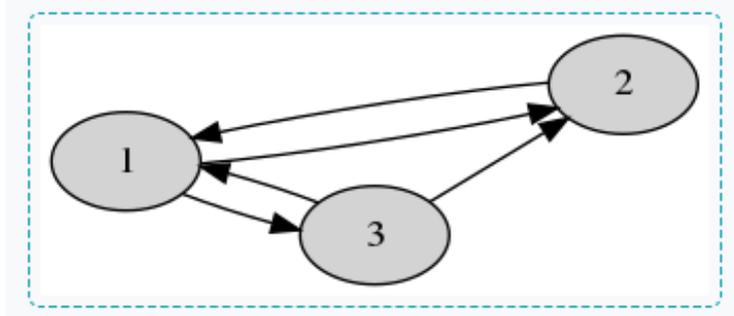
C'est la page 4 car seuls les internautes qui connaissent son adresse peuvent la trouver

2. Citer la page qui a le plus de chance de recevoir des visites et dites pourquoi.

C'est la page 2 car trois pages la référencent

Exercice : recherche d'un pagerank à l'aide d'edupython

On se donne 3 pages web reliées par des liens hypertextes dont on donne une modélisation sous forme de graphe.



1. Citer la page qui a le moins de chances de recevoir des visites.

La page 3 a le moins de chances de recevoir des visites

2. Citer la page qui a le plus de chances de recevoir des visites.

La page 1 a le plus de chances de recevoir des visites

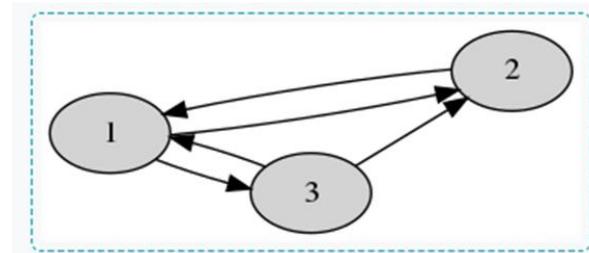
3. **On suppose qu'un robot est sur la page 1.** Il se déplace aléatoirement à l'aide des liens hypertextes.

Lorsque le robot est sur la page 1, il peut soit choisir la page 2 soit la page 3 de manière aléatoire.

Lorsque le robot est sur la page 3, il peut soit choisir la page 1 soit la page 2 de manière aléatoire.

Lorsque le robot est sur la page 2, il n'a pas le choix et doit se rendre sur la page 1.

Compléter le programme python ci-dessous afin qu'il affiche le compteur de chacune des pages lorsque l'on demande à un robot de se déplacer 10 000 fois aléatoirement dans notre mini-toile de trois pages.



```
from random import*
page = 1
compteurpage1 = 0
compteurpage2 = 0
compteurpage3 = 0
for k in range(10000):
    if page == 1:
        if randint(1, 2) == 1: #"Le robot est en 1 et va vers 2"
            page = 2
            compteurpage2 = compteurpage2+1
        else: #"Le robot est en 1 et va vers 3"
            page = 3
```

```

compteurpage3 = compteurpage3+1
elif page == 2:    #"Le robot est en 2 et va vers 1"
    page = 1
    compteurpage1 = compteurpage1+1
else:
    if randint(1, 2) == 1: #"Le robot est en 3 et va vers 1"
        page = 1
        compteurpage1= compteurpage1+1
    else:
        page = 2 #"Le robot est en 3 et va vers 2"
        compteurpage2 = compteurpage2 + 1

```

```
print(compteurpage1, compteurpage2 , compteurpage3)
```

```

1 from random import*
2 page = 1
3 compteurpage1 = 0
4 compteurpage2 = 0
5 compteurpage3 = 0
6 for k in range(10000):
7     if page == 1:
8         if randint(1, 2) == 1: #"Le robot est en 1 et va vers 2"
9             page = 2
10            compteurpage2 = compteurpage2+1
11        else: #"Le robot est en 1 et va vers 3"
12            page = 3
13            compteurpage3 = compteurpage3+1
14    elif page == 2: #"Le robot est en 2 et va vers 1"
15        page = 1
16        compteurpage1 = compteurpage1+1
17    else:
18        if randint(1, 2) == 1: #"Le robot est en 3 et va vers 1"
19            page = 1
20            compteurpage1= compteurpage1+1
21        else:
22            page = 2 #"Le robot est en 3 et va vers 2"
23            compteurpage2 = compteurpage2 + 1
24
25 print(compteurpage1, compteurpage2 , compteurpage3)
26
27

```

```

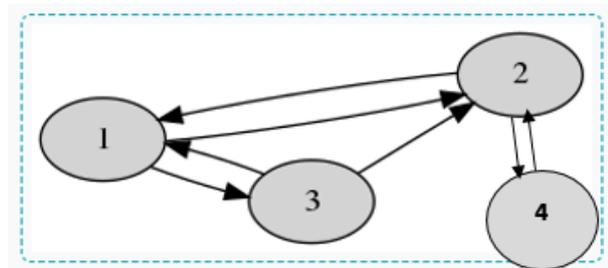
win32. ***
*** Distant Python engine is active ***
>>>
*** Console de processus distant Réinitialisée ***
4439 3322 2239
>>>
*** Console de processus distant Réinitialisée ***
4433 3311 2256
>>>

```

Ouvrir Edupython et copier coller le programme ci-dessus. Joindre une capture d'écran de la console. Sauvegarder votre travail dans SNT/Theme3 sous le nom pagerank

Lorsque le robot effectue 10 000 déplacements dans la mini-toile , il visite 4433 fois le site 1 ,3311 fois le site 2 et 2256fois le site 3.

4. On se donne 4 pages web reliées par des liens hypertextes dont on donne une modélisation sous forme de graphe.



En s'inspirant de la question précédente, écrire un programme Python permettant de simuler le déplacement du robot. Joindre une capture d'écran de son travail et sauvegarder sous pagerank2.

```

1 from random import*
2 page = 1
3 compteurpage1 = 0
4 compteurpage2 = 0
5 compteurpage3 = 0
6 compteurpage4 = 0
7 for k in range(10000):
8     if page == 1:
9         if randint(1, 2) == 1: # "Le robot est en 1 et va vers 2"
10            page = 2
11            compteurpage2 = compteurpage2+1
12        else: # "Le robot est en 1 et va vers 3"
13            page = 3
14            compteurpage3 = compteurpage3+1
15    elif page == 2:
16        if randint(1, 2) == 1: # "Le robot est en 2 et va vers 1"
17            page = 1
18            compteurpage1 = compteurpage1+1
19        else: # "Le robot est en 2 et va vers 4"
20            page = 4
21            compteurpage4 = compteurpage4+1
22    elif page == 3:
23        if randint(1, 2) == 1: # "Le robot est en 3 et va vers 1"
24            page = 1
25            compteurpage1 = compteurpage1+1
26        else: # "Le robot est en 3 et va vers 2"
27            page = 2
28            compteurpage2 = compteurpage2+1
29    else: # "Le robot est en 4 et va vers 2"
30        page = 2
31        compteurpage2 = compteurpage2+1
32
33
34 print(compteurpage1, compteurpage2 , compteurpage3,compteurpage4)

```

```

Console Python
*** Console de processus distant Réinitialisée ***
2633 4047 1294 2026
>>>

```

Activité 5 : sécurité et vie privée sur le web

vidéo : lienmini.fr/1046-306 (3mns)

1. Identifier les deux informations qui permettent de repérer que l'accès à une page web est sécurisé.

(doc1)

Lorsque l'adresse d'une page web est précédée d'un cadenas et le nom du protocole http est suivi d'un S alors l'accès est sécurisé

2. Préciser la raison pour laquelle le site d'achats en ligne est sécurisé . (doc1)

Le site d'achats en ligne est sécurisé pour éviter que les paiements soient détournés par des personnes malveillantes.

3. Interpréter les deux écrans ci-dessous (le 1^{er} écran on a juste passé la souris sur le lien , dans le 2^{ème} on vient juste de cliquer sur le lien)

Doc 1 Les accès à une page Web



<http://monsiteweb1.fr>

Page non sécurisée



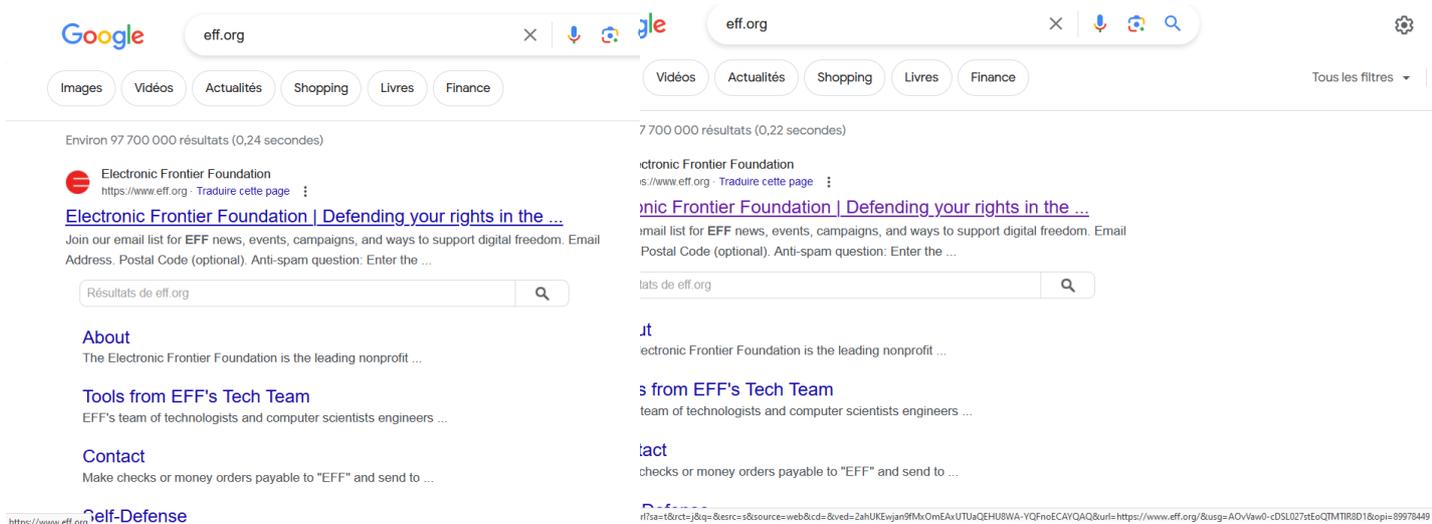
<https://achatsenligne.fr>

Page sécurisée
cadenas fermé 

Les pages Web qui utilisent le protocole **HTTPS** (*HyperText Transfer Protocol Secure*) sont sécurisées.

Un cadenas apparaît alors dans la barre d'adresse du navigateur pour indiquer que la communication entre le client et le serveur est sécurisée.

3. Interpréter les deux écrans ci-dessous (le 1^{er} écran on a juste passé la souris sur le lien , dans le 2^{ème} on vient juste de cliquer sur le lien)



Dès que vous cliquez sur le lien, la page envoie une requête à google.com, indiquant à l'entreprise d'où vous venez et où vous allez. De cette façon, Google peut suivre non seulement vos recherches, mais aussi les liens sur lesquels vous cliquez.

4. Les cookies

Rappel : Un cookie est un petit fichier stocké par un serveur dans le terminal (ordinateur, téléphone, etc.) d'un utilisateur et associé à un domaine web (c'est à dire dans la majorité des cas à l'ensemble des pages d'un même site web). Ce fichier est automatiquement renvoyé lors de contacts ultérieurs avec le même domaine.

Les cookies ont de multiples usages : ils peuvent servir à mémoriser votre identifiant client auprès d'un site marchand, le contenu courant de votre panier d'achat, la langue d'affichage de la page web, un identifiant permettant de tracer votre navigation à des fins statistiques ou publicitaires, etc. Certains de ces usages sont strictement nécessaires aux fonctionnalités expressément demandées par l'utilisateur ou bien à l'établissement de la communication et donc exemptés de consentement. D'autres, qui ne correspondent pas à ces critères, nécessitent un consentement de l'utilisateur avant lecture ou écriture.

Citer les différentes traces que l'on peut laisser lors de la navigation sur le Web . (doc2). Indiquer comment les contrôler.

Le navigateur mémorise les sites visités, recueille des données (nombre de pages visités, clics , recherches , saisies dans les formulaires, etc). Pour les contrôler, on peut agir sur les paramètres de son navigateur : historique, cookies , mots de passe

4.Regarder la vidéo : lienmini.fr/3389-307 (5mns24s)

Doc 2 Contrôler et sécuriser sa navigation

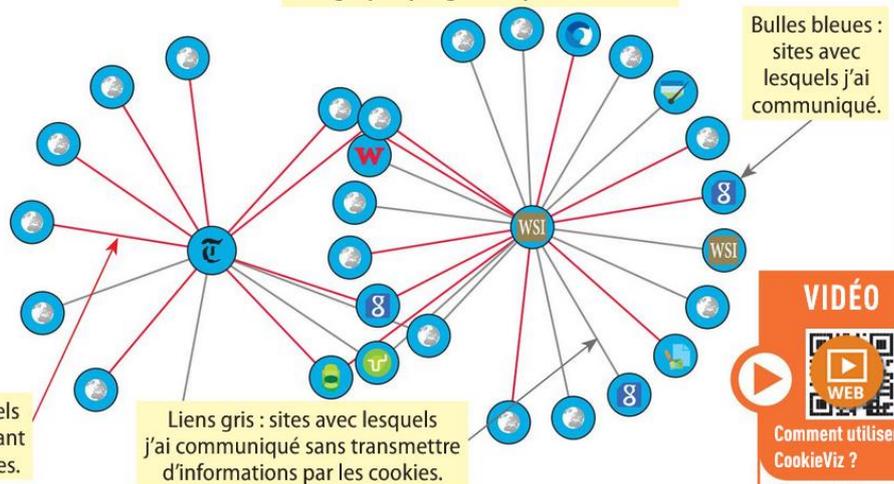
Quand on navigue sur le Web, on laisse des traces (identité, historique, choix des articles, produits, clics, etc.) qui permettent de vous connaître et de vous proposer des informations. Pour protéger sa vie privée et ainsi éviter de laisser des traces, on peut régler les paramètres de contrôle de son navigateur (cookies, historique, sécurité, confidentialité).



DOC 3 Les cookies

CookieViz est un logiciel créé par la CNIL (Commission nationale de l'informatique et des libertés) qui propose un outil de visualisation pour mesurer l'impact des cookies et autres traqueurs lors de sa navigation.

Un graphique généré par CookieViz



Quel est l'intérêt d'un logiciel comme CookieViz ? (firefox utilise l'outil Disconnect)

CookieViz permet de voir les types d'informations qui sont récoltés lors de sa navigation. C'est un outil de visualisation pour mesurer l'impact des cookies et autres traqueurs.

5.La navigation privée est t'elle un moyen efficace de protéger mes données personnelles ?

<https://support.mozilla.org/fr/kb/idees-recues-sur-navigation-privée>

Quelles sont les avantages de la navigation privée ? La navigation privée me permet elle d'être anonyme sur le web ?

La navigation privée permet d'éviter les cookies mais ne permet pas d'être totalement invisible sur le net.