SNT · LE WEB · SÉANCE 4

Chercher, publier, se protéger

Trois problèmes du quotidien, des règles à construire

Recherche · droit d'auteur · hameçonnage · navigateur

Problème n° 5 : une page sans adresse

Comment trouver une page dont on ignore l'URL ?



Problème n° 5 : une page sans adresse

Comment trouver une page dont on ignore l'URL ?

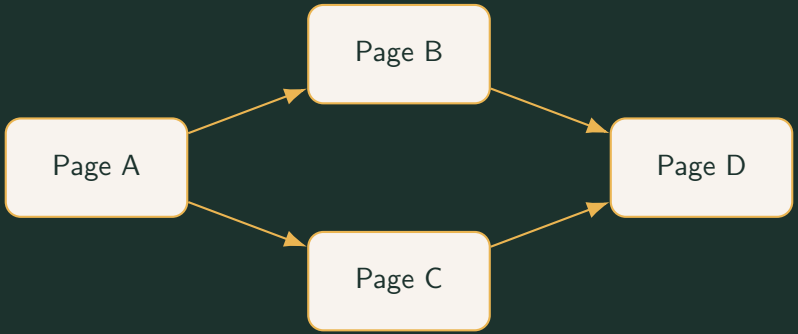


L'obstacle

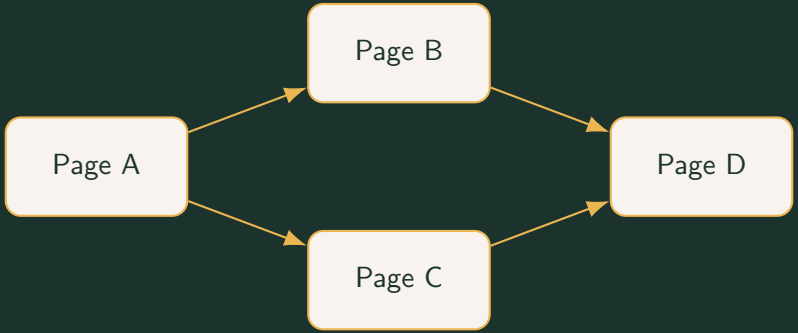
Il serait beaucoup trop long de parcourir tout le Web au moment où l'on tape chaque recherche.

À faire formuler : Comment se préparer à répondre très vite ?

Qui découvre les pages ?



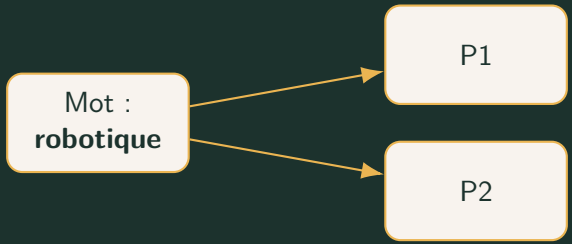
Qui découvre les pages ?



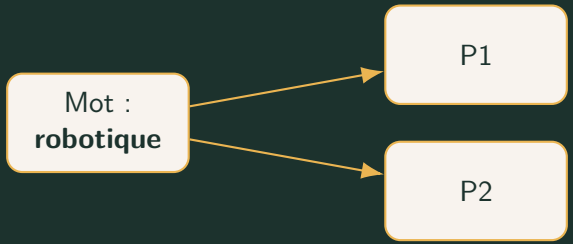
Exploration

Un **robot** suit les liens et visite des pages pour recueillir leur contenu. Ce travail se

Comment retrouver un mot sans relire tous les documents ?



Comment retrouver un mot sans relire tous les documents ?



Index

Comme à la fin d'un livre, un **index** associe les mots aux pages qui les contiennent.

À faire formuler : Dans quelles pages du livret apparaît le mot « tournoi » ?

Activité 4.1 : indexer à la main

Et si nous étions un moteur de recherche ?

1. Lire trois mini-pages.
2. Construire la liste des pages contenant chaque mot.
3. Chercher « tournoi robotique ».
4. Comparer avec « inscriptions mars ».

Livret 4 · Activité 4.1

L'index ne suffit pas

Deux pages correspondent : laquelle afficher en premier ?

P1
« inscriptions mars »

P3
« inscriptions mars »

L'index ne suffit pas

Deux pages correspondent : laquelle afficher en premier ?

P1
« inscriptions mars »

P3
« inscriptions mars »

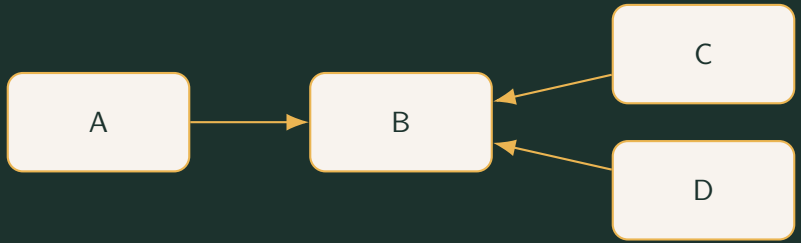
Le nouveau problème

Plusieurs pages peuvent contenir les mêmes mots. Il faut les **classer** en tenant compte de critères supplémentaires.

À faire formuler : Une page citée par beaucoup d'autres serait-elle plus utile ?

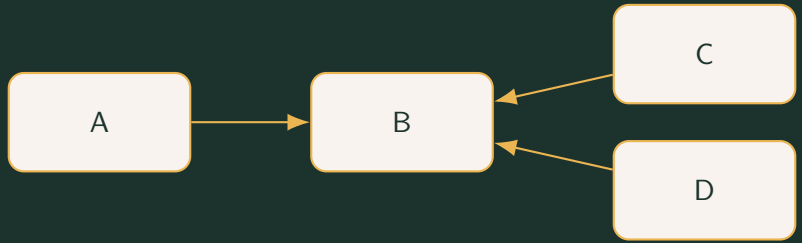
Les liens comme votes

Une page reçoit trois liens. Une autre n'en reçoit qu'un.



Les liens comme votes

Une page reçoit trois liens. Une autre n'en reçoit qu'un.



Première intuition

Les **liens entrants** peuvent indiquer qu'une page est citée par d'autres. Mais tous les liens n'ont pas la même valeur.

Pourquoi compter les liens ne suffit pas

Et si je créais 1 000 fausses pages ?

- Le simple nombre de liens peut être manipulé.

Pourquoi compter les liens ne suffit pas

Et si je créais 1 000 fausses pages ?

- Le simple nombre de liens peut être manipulé.
- Un lien venant d'une page reconnue peut avoir davantage de poids.

Pourquoi compter les liens ne suffit pas

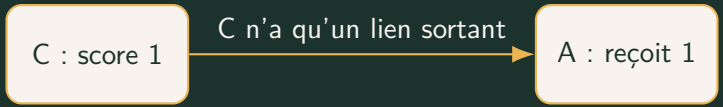
Et si je créais 1 000 fausses pages ?

- Le simple nombre de liens peut être manipulé.
- Un lien venant d'une page reconnue peut avoir davantage de poids.
- Une page qui crée de nombreux liens partage son poids entre eux.

Modèle simplifié de PageRank

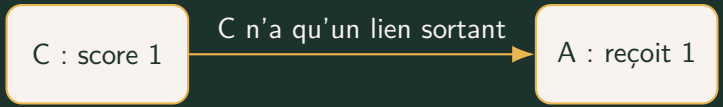
Chaque page commence avec un score de 1, puis le partage entre ses liens sortants.
Nous allons effectuer **un seul tour** de calcul.

Qui donne du score à qui ?



Activité 4.2 : calcul guidé

Qui donne du score à qui ?



Méthode

Contribution d'un lien = score de départ de la page / nombre de ses liens sortants.

Calculons A ensemble, puis B, C, D et E.

Le classement mesure-t-il la fiabilité ?

- Un lien très consulté peut contenir une erreur.

Un résultat en premier est-il vrai ?

Le classement mesure-t-il la fiabilité ?

- Un lien très consulté peut contenir une erreur.
- Un premier résultat peut être une annonce.

Le classement mesure-t-il la fiabilité ?

- Un lien très consulté peut contenir une erreur.
- Un premier résultat peut être une annonce.
- Pour un exposé, il faut regarder **qui parle**, **quand**, et **sur quelles preuves**.

Livret 4 · Activité 4.3

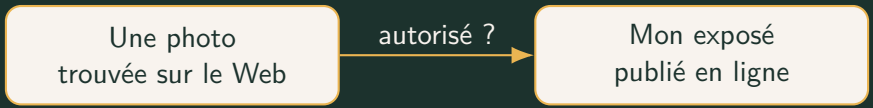
Une image trouvée en ligne

Puis-je la publier sur le site du lycée ?



Une image trouvée en ligne

Puis-je la publier sur le site du lycée ?



Principe

Une œuvre originale bénéficie du **droit d'auteur** dès sa création. Être visible en ligne ne signifie pas « libre à réutiliser ».

Comment partager avec des règles claires ?

L'auteur peut-il autoriser certaines réutilisations ?

L'auteur peut-il autoriser certaines réutilisations ?

La licence

Une **licence** précise les usages autorisés et les conditions à respecter.

- BY Citer l'auteur et fournir les informations de licence.
- NC Pas d'utilisation commerciale.
- ND Pas de diffusion d'adaptation.
- SA Diffuser une adaptation sous une licence compatible.

Un e-book acheté se revend-il comme un livre papier ?

- Pour un livre papier, je possède l'exemplaire matériel.

Acheter un fichier, quels droits ?

Un e-book acheté se revend-il comme un livre papier ?

- Pour un livre papier, je possède l'exemplaire matériel.
- Pour un téléchargement, la copie reste soumise aux droits et conditions d'utilisation applicables.

Un e-book acheté se revend-il comme un livre papier ?

- Pour un livre papier, je possède l'exemplaire matériel.
- Pour un téléchargement, la copie reste soumise aux droits et conditions d'utilisation applicables.
- Posséder une copie numérique ne signifie pas pouvoir la redistribuer librement.

Livret 4 · Activité 4.4

Le piège du faux site

Deux adresses se ressemblent. À laquelle faire confiance ?

Adresse A

`https://www.mabanque.example/compte`

Adresse B

`https://www.mabanque.example.connexion-securisee.example/compte`

Le piège du faux site

Deux adresses se ressemblent. À laquelle faire confiance ?

Adresse A

`https://www.mabanque.example/compte`

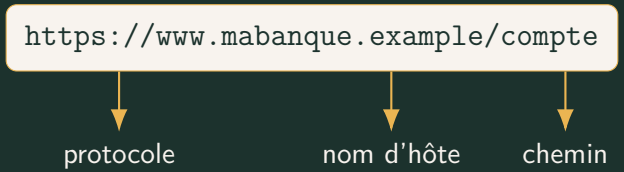
Adresse B

`https://www.mabanque.example.connexion-securisee.example/compte`

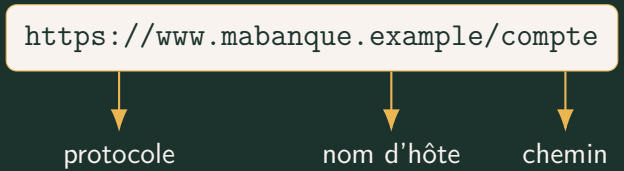
Le nom de la banque apparaît dans les deux !

À faire formuler : Quelle partie indique réellement le site contacté ?

Où se trouve le nom d'hôte ?



Où se trouve le nom d'hôte ?



Méthode

Après `https://`, lis jusqu'au prochain `/`. Identifie ensuite le **domaine principal** en partant de la droite.

HTTPS garantit-il l'honnêteté ?

Un faux site peut-il afficher le cadenas ?

HTTPS

protège la **communication** avec le site

Un faux site peut-il afficher le cadenas ?

HTTPS

protège la **communication** avec le site

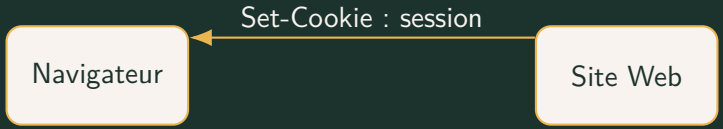
Mais attention

Un fraudeur peut lui aussi utiliser HTTPS. Il faut contrôler l'adresse et le contexte de la demande avant de saisir des informations sensibles.

Livret 4 · Activité 4.5

Pourquoi un site me reconnaît-il ?

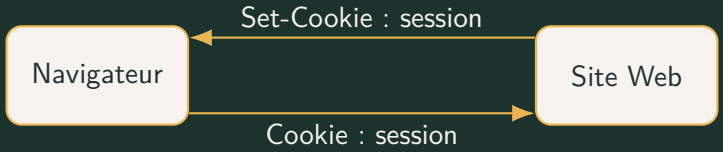
Comment rester connecté en changeant de page ?



Cookie

Une petite donnée conservée par le navigateur peut aider le site à retrouver la session de l'utilisateur.

Comment rester connecté en changeant de page ?

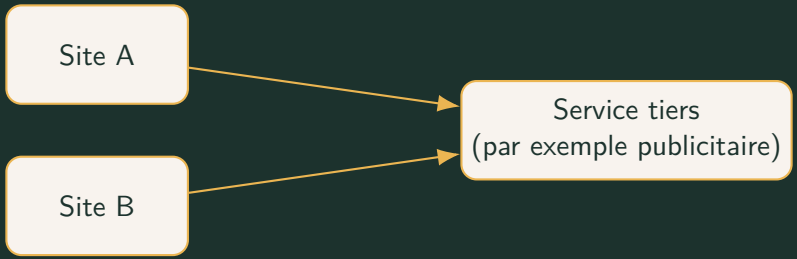


Cookie

Une petite donnée conservée par le navigateur peut aider le site à retrouver la session de l'utilisateur.

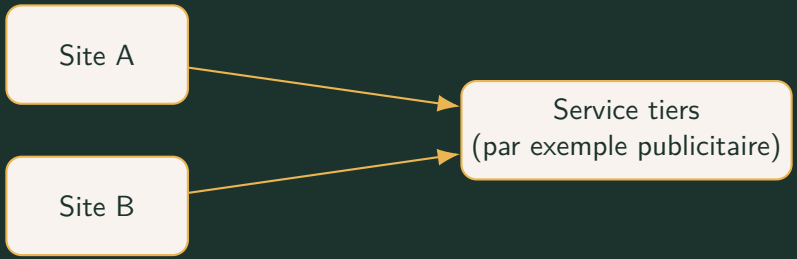
Quand le suivi dépasse un seul site

Pourquoi un même service apparaît-il sur plusieurs sites ?



Quand le suivi dépasse un seul site

Pourquoi un même service apparaît-il sur plusieurs sites ?



Pistage

Certains services tiers peuvent reconnaître un navigateur d'un site à l'autre à l'aide de traceurs. Les protections du navigateur cherchent à limiter ce suivi.

Qui ne voit plus mes visites ?

- La navigation privée limite les traces conservées **sur l'appareil après la session**.

Qui ne voit plus mes visites ?

- La navigation privée limite les traces conservées **sur l'appareil après la session**.
- Le site consulté reçoit toujours les requêtes.

Qui ne voit plus mes visites ?

- La navigation privée limite les traces conservées **sur l'appareil après la session**.
- Le site consulté reçoit toujours les requêtes.
- Le réseau utilisé peut encore observer des informations sur les connexions.

En classe

Dans Firefox : **Paramètres** puis **Vie privée et sécurité**. Observer sans modifier irréversiblement les réglages du poste.

Activité 4.6 : observer Firefox

Que nous montre réellement le navigateur ?

1. Trouver les options de protection et les autorisations.
2. Ouvrir **F12** puis **Stockage** pour observer les cookies visibles.
3. Cliquer sur le **bouclier** pour lire les protections signalées.

Activité 4.6 : observer Firefox

Que nous montre réellement le navigateur ?

1. Trouver les options de protection et les autorisations.
2. Ouvrir **F12** puis **Stockage** pour observer les cookies visibles.
3. Cliquer sur le **bouclier** pour lire les protections signalées.

Important

Cookies stockés, domaines contactés et traqueurs bloqués sont **trois observations différentes**.

Vérification finale

Une adresse commence par HTTPS. Est-elle forcément sûre ?

A. Oui, le site est nécessairement honnête.

B. Non, HTTPS chiffre la communication mais ne prouve pas l'honnêteté.

C. Oui, si le nom de la banque apparaît quelque part.

Une adresse commence par HTTPS. Est-elle forcément sûre ?

A. Oui, le site est nécessairement honnête.

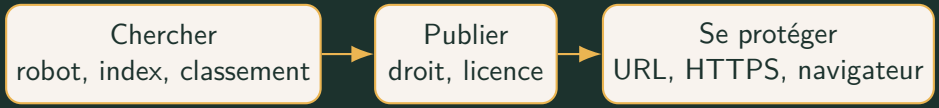
B. Non, HTTPS chiffre la communication mais ne prouve pas l'honnêteté.

C. Oui, si le nom de la banque apparaît quelque part.

Réponse B

Vérifier le véritable domaine principal et le contexte avant de saisir ses identifiants.

Bilan du Web



Trois questions à retenir

par Comment trouve-t-on une information ?

Quels droits a-t-on sur une œuvre ?

Comment vérifier une adresse et limiter ses traces ?