

Technologie XML

Par :
Pr. Imad Zeroual

Faculté des Sciences et Techniques – Errachidia / 2022-2023



Introduction

Informations, Notions.

Langage XML

Introduction, Arbre XML, Règles de syntaxe.

Définition de Type de Document (DTD)

Définition, Éléments, Attributs, Entités.

Schémas XSD

Structure de base, Éléments, Attributs, Indicateurs.

XPATH & XSLT

Langage de requêtage XPath, Feuilles de style XSL.



Introduction

Informations, Notions.

01



Introduction Informations



‘Web,

en français, la toile

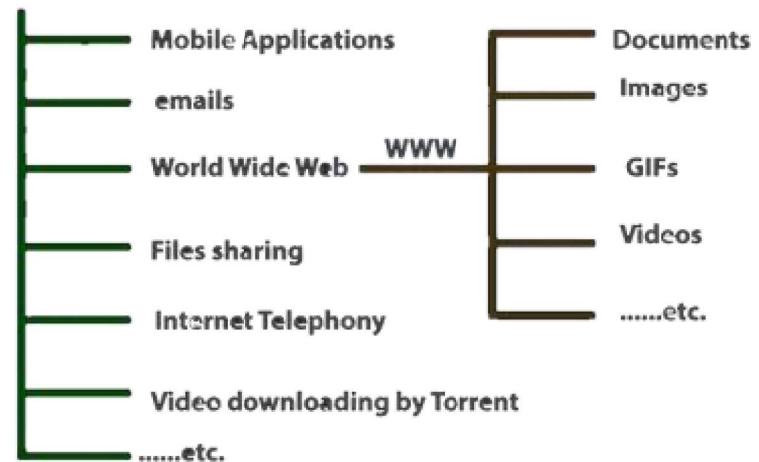
Le **World Wide Web** (abrégé www ou le Web) est un système hypertexte public fonctionnant sur Internet.

Le Web n'est qu'une des applications d'Internet !!

L'**Internet** est le réseau informatique mondial accessible au public.

Un **réseau informatique** est un ensemble d'équipements reliés entre eux pour échanger des informations.

Internet





Introduction Informations



‘Web,

en français, la toile

Le **World Wide Web** (abrégé www ou le Web) est un système hypertexte public fonctionnant sur Internet.

Le Web n'est qu'une des applications d'Internet !!

L'**Internet** est le réseau informatique mondial accessible au public.

Un **réseau informatique** est un ensemble d'équipements reliés entre eux pour échanger des informations.

20 Mrd < d'appareils connectés dans le monde !!





Introduction Informations



‘Web,

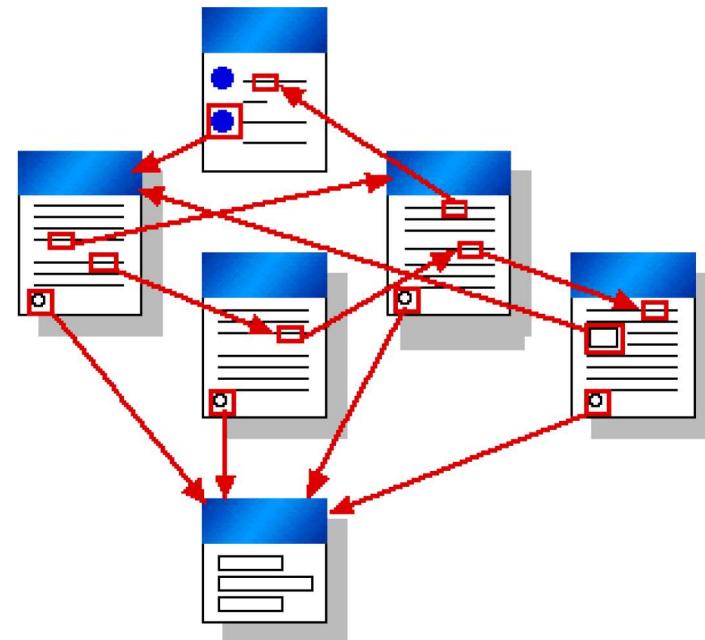
en français, la toile

Le **World Wide Web** (abrégé www ou le Web) est un système **hypertexte** public fonctionnant sur Internet.

Le Web permet de consulter, avec un navigateur, des pages accessibles sur des sites.

Un **hypertexte** est un document contenant des unités d'information liées entre elles par des **hyperliens**.

Un **hyperlien** est une référence dans un système hypertexte permettant de passer automatiquement d'une page consultée à un autre endroit de même page, une autre page ou un autre site.





Introduction Informations



Un **IRI** est un identificateur de ressource **internationalisé** qui prit en compte les divers alphabets utilisés dans les différentes langues du monde.

Une **URL** (Localisateur Uniforme de Ressource) est une chaîne de caractères uniforme qui permet d'**identifier l'emplacement d'une ressource** (un document, une image, un enregistrement sonore, etc.) du **World Wide Web** et de préciser le **protocole** internet pour la récupérer (par exemple **http** ou **https**).

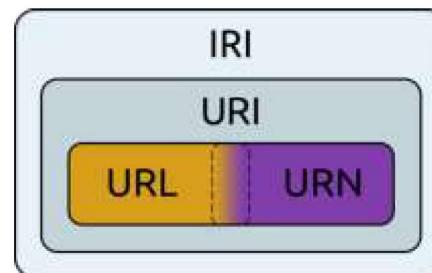
'Web,' en français, la toile

موقع.السعودية/
موقع.الجزيرة/

http://

موقع.الجزيرة/

http://



Un **URI** identifie une ressource de manière permanente sur un réseau.

Par opposition à l'**URL**, une **URN** (Nom Uniforme de Ressource) identifie la ressource **elle-même**, indépendamment de son **emplacement**.

Syntaxe de URN : urn:<nid>:<nss>

Exemples :

urn:isbn:0451450523

urn:ietf:rfc:2648

https://fr.wikipedia.org/wiki/Web_semantique#/media/Fichier:Semantic_Web_Stack.png

URL

URN

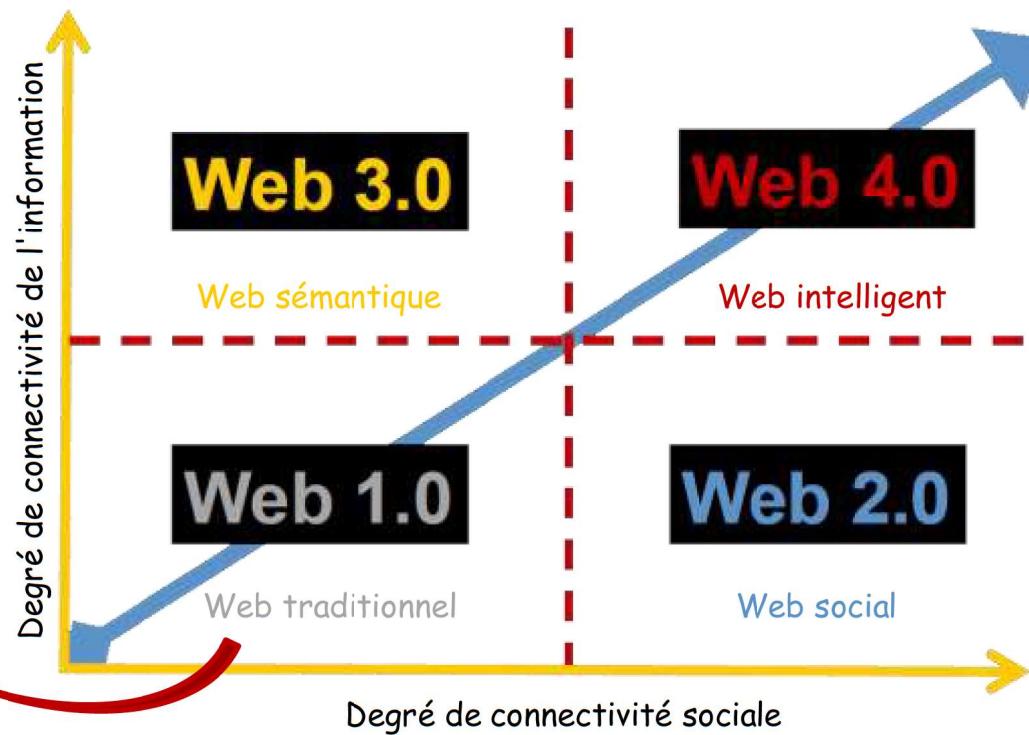


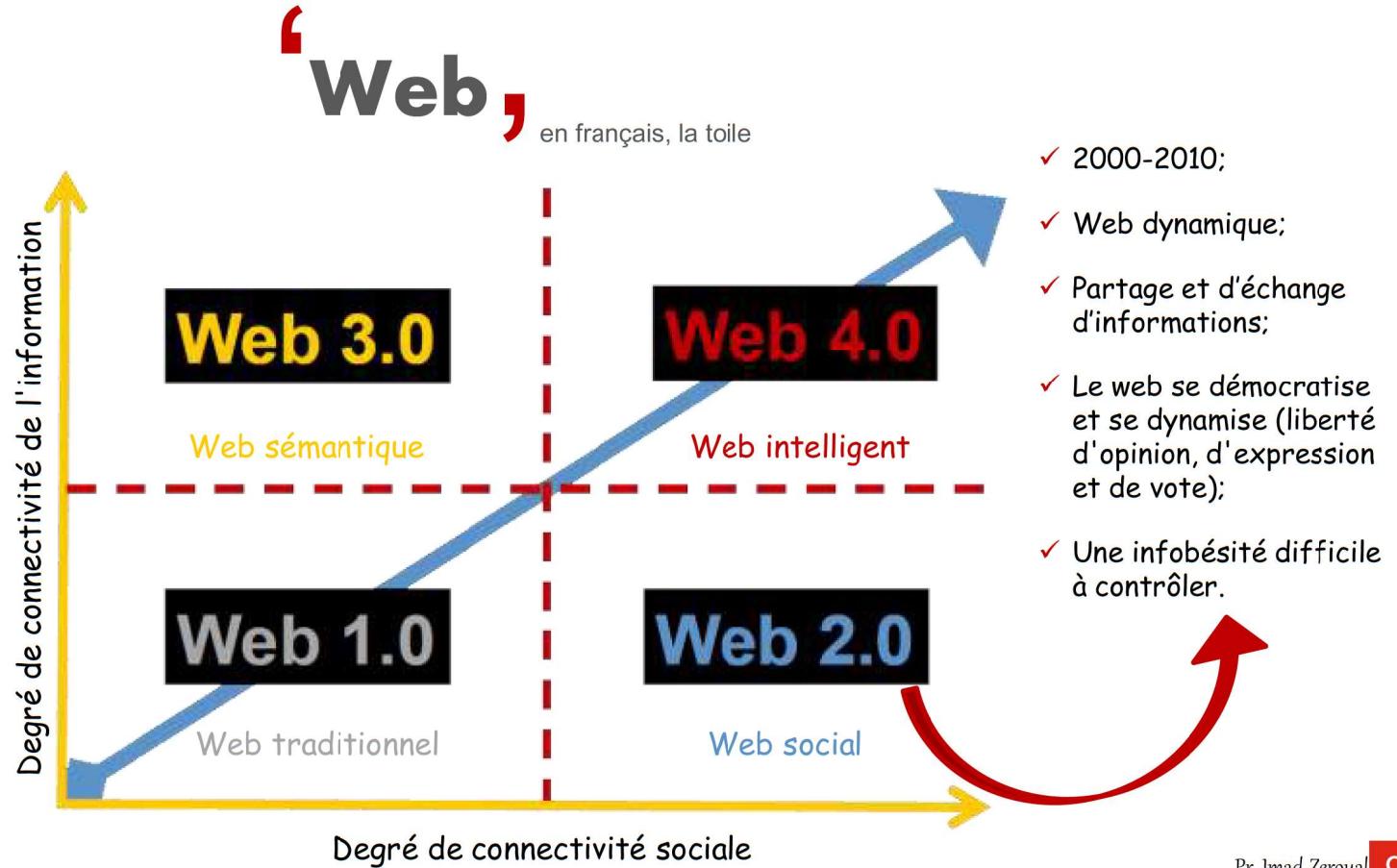
Introduction Informations



‘Web,’ en français, la toile

- ✓ 1990-2000;
- ✓ Web statique;
- ✓ Centré sur la distribution d'infos;
- ✓ Sites en lecture seule





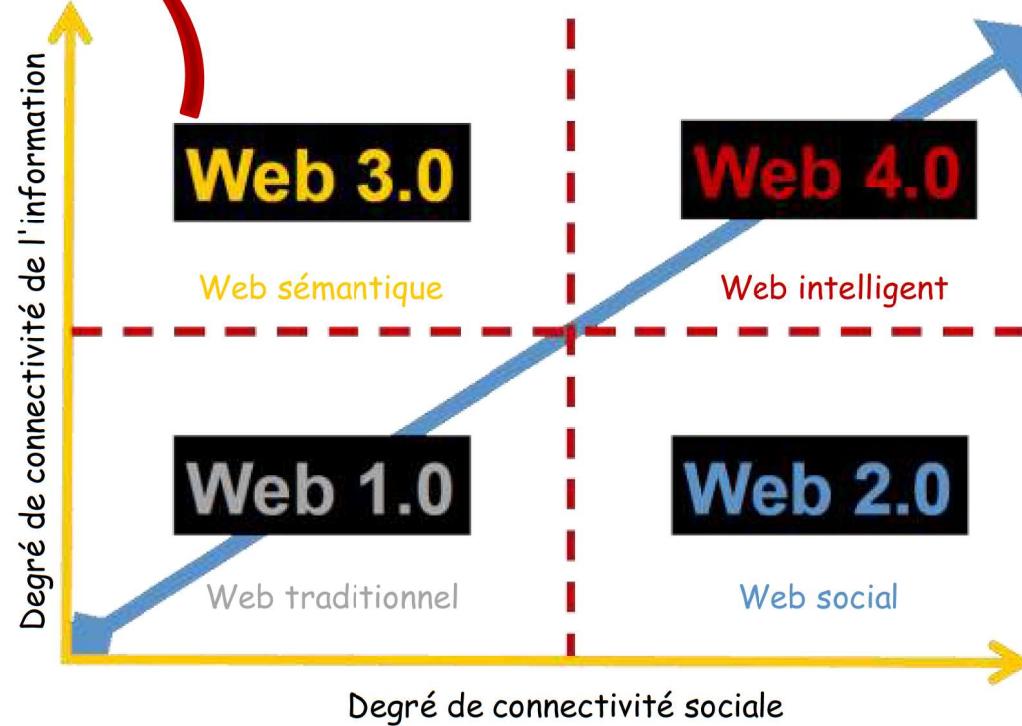


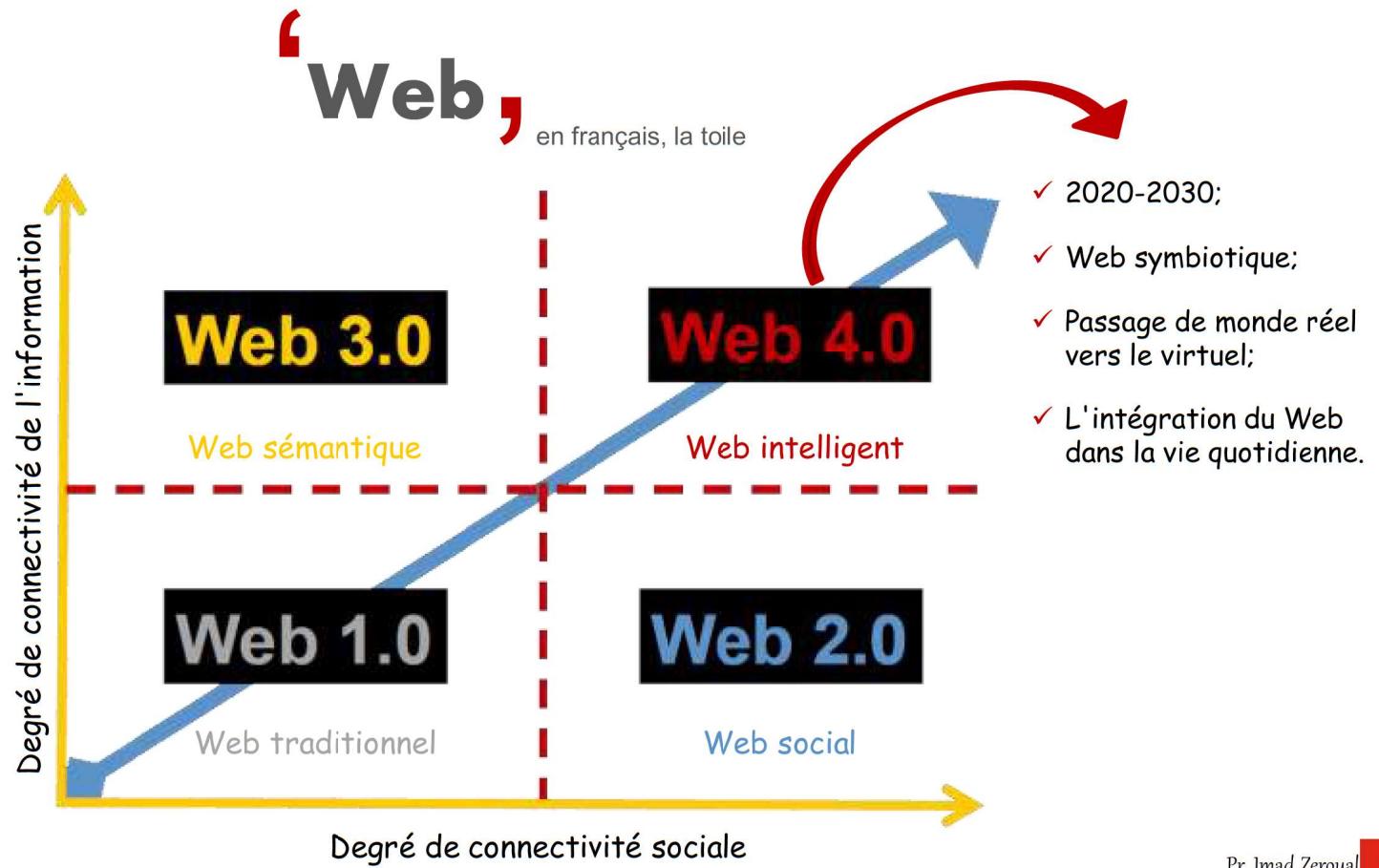
Introduction Informations



Web, en français, la toile

- ✓ 2010-2020;
- ✓ Web de connaissances;
- ✓ La mise à niveau du back-end du web;
- ✓ Consiste à lier et structurer l'information sur Internet.
- ✓ Enseigner à l'ordinateur ce que signifient les données
- ✓ RDF, OWL, SWRL, SPARQL,







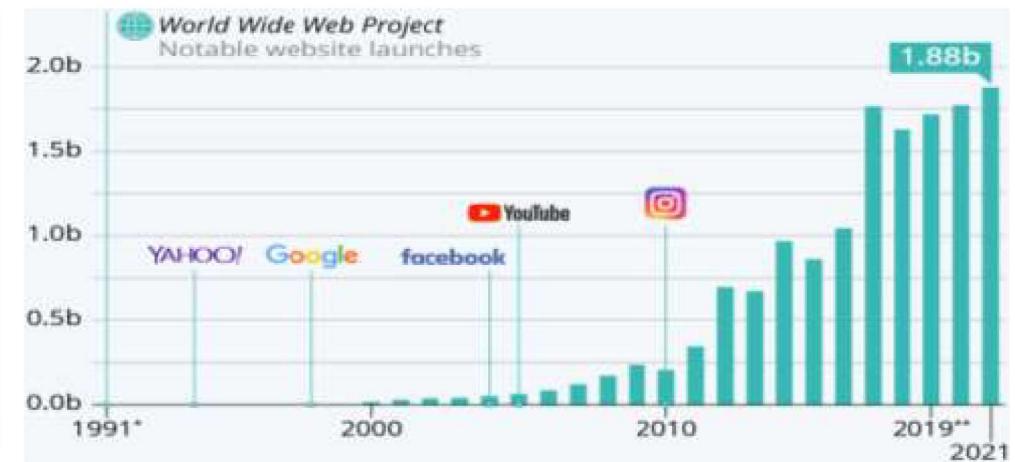
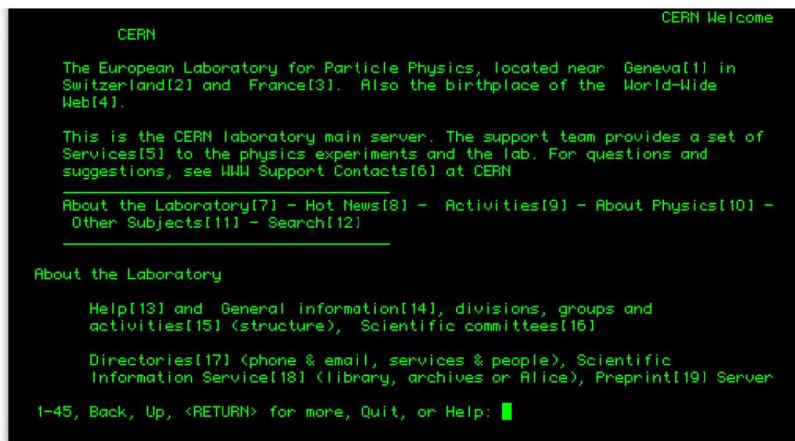
Introduction Informations



‘Web, en français, la toile

- De 1 site en 1991 à ≈ 2 milliard aujourd'hui (Décembre 2021).

<http://info.cern.ch>



- 130 mile milliards est le nombre de pages indexées sur Google en 2016.



Introduction Informations



‘ Web,

en français, la toile

- En 2009, le contenu du Web a dépassé ce qui est dans tous les **livres du monde** par **5 milliards de fois**.

5.000.000.000

X





Introduction Informations



SURFACE WEB

Bing

Google

4%

(not accessible to Surface Web crawlers)

DEEP WEB

Medical Records

Academic Information

Legal Documents

Multilingual Databases

Scientific Reports

Financial Records

Subscription Information

Government Resources

Competitor Websites

Organisation-specific Repositories

90%

DARK WEB

(only accessible through certain browsers such as TOR. Deep web technologies has zero involvement with the Dark Web)

TOR Encrypted sites

Drug Trafficking

6%

Private Communications

Political Protests

Illegal Information

'Web,

en français, la toile

- Pages protégées par mot de passe.
- Dossiers de présence, certificats médicaux, etc.
- Registres bancaires.
- Messages de chat cryptés (WhatsApp, Facebook Messenger, etc.)
- E-mails envoyés entre utilisateurs.
- Contenu nécessitant un achat lorsqu'il n'est pas indexé sur les moteurs de recherche (chansons / vidéos, cours, etc.).



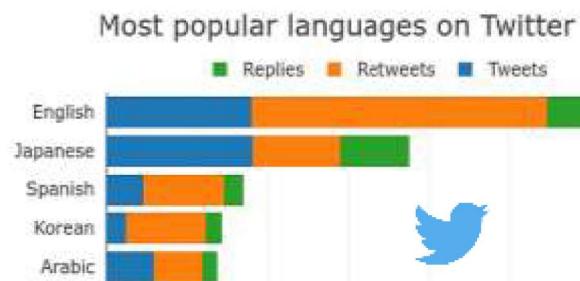
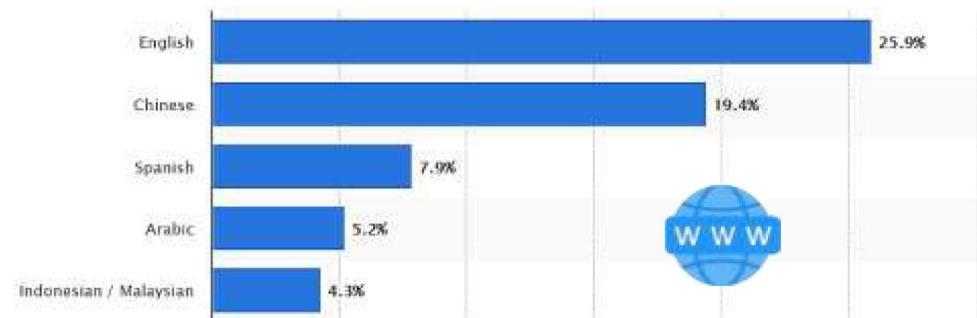
Introduction Informations



‘Web,’ en français, la toile

Les **TOP 5 Langues** les plus couramment utilisées sur Internet **2022***

L'anglais était la langue la plus populaire en ligne, représentant **25,9%** des internautes dans le monde. Les dix principales langues représentaient **76,9%** des internautes mondiaux.



A screenshot of the Google Translate interface. It shows a translation from Swahili to English. The text reads: 'Gutuma buri wese, aho ari ho hose amenyia isi aikanavuga ibyo atekereza mu ndimi zitandukanye.' Below it, the English translation is: 'Let everyone, anywhere, know the world and express their thoughts in different languages.'

*<https://www.statista.com/statistics/262946/share-of-the-most-common-languages-on-the-internet/>



Introduction

Informations

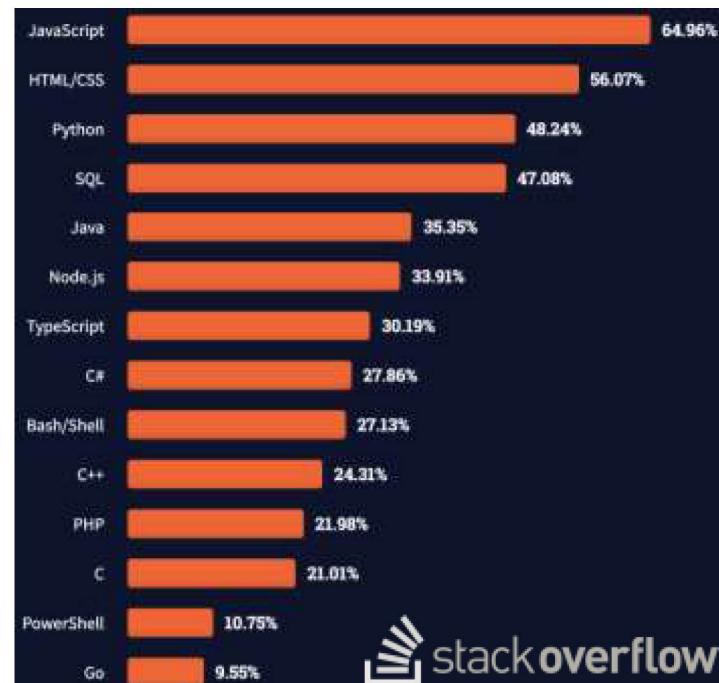
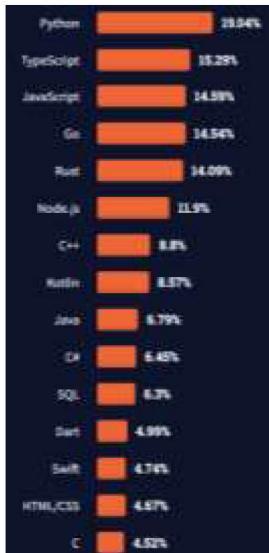


‘Web,

en français, la toile

Les **TOP Langages** de programmation à apprendre en **2022***

Les langages les plus **demandées**



*<https://medium.com/dare-to-be-better/top-programming-languages-and-web-frameworks-to-learn-in-2022-776113bfefeb/>