

INF344 – Données du Web

Introduction : le Web, les navigateurs

Antoine Amarilli

24 avril 2017

- 1969** ARPANET, ancêtre d'Internet, est lancé
- 1974** Introduction de TCP
- 1990** Introduction du World Wide Web, HTTP, HTML
- 1994** Yahoo! est fondé
- 1995** Amazon.com, Ebay, AltaVista sont fondés
- 1998** Google est fondé
- 2001** Wikipédia est lancé

- Près de 300 millions de **domaines**, dont 120 millions dans **.com**¹
- **52%** du contenu est en **anglais**, **4%** en français²
- Près de **la moitié** de la population mondiale utilise Internet³
- Google connaît plus d'un **trillion** (10^{12}) d'URLs uniques⁴
 - Une large partie du Web n'est pas indexable : le **Web caché**

1. <https://investor.verisign.com/releasedetail.cfm?releaseid=947518>

2. https://w3techs.com/technologies/overview/content_language/all

3. <https://www.itu.int/en/ITU-D/Statistics/Pages/facts/default.aspx>

4. <https://googleblog.blogspot.fr/2008/07/we-knew-web-was-big.html>

Table des matières

Introduction

Navigateurs Web

Autres clients Web

URLs

Comment les gens vont sur le Web ?

- On utilise un **navigateur Web**
- Dépend du **système d'exploitation sous-jacent**
- Fonctionne sur un **téléphone** ou sur un **ordinateur**

Navigateurs Web historiques

Mosaic Premier navigateur graphique répandu, 1993-1997

- De 80% en 1994 à < 10% en 1996

Navigateurs Web historiques

Mosaic Premier navigateur graphique répandu, 1993-1997

- De 80% en 1994 à < 10% en 1996

Netscape Lancé en 1994, basé sur Mosaic

- Propriétaire, usage non-commercial gratuit
- 80% d'usage en 1996, < 10% en 2001
- N'est plus maintenu depuis 2008

Navigateurs Web historiques

Mosaic Premier navigateur graphique répandu, 1993-1997

- De 80% en 1994 à < 10% en 1996

Netscape Lancé en 1994, basé sur Mosaic

- Propriétaire, usage non-commercial gratuit
- 80% d'usage en 1996, < 10% en 2001
- N'est plus maintenu depuis 2008

Internet Explorer Lancé en 1995, basé sur Spyglass Mosaic

- Propriétaire, fourni avec Windows 95
- IE 6 lancé en 2001 atteint 80% du marché
- Mène au procès antitrust aux USA, 1998-2001

Navigateurs Web historiques

Mosaic Premier navigateur graphique répandu, **1993-1997**

- De **80%** en 1994 à **< 10%** en 1996

Netscape Lancé en **1994**, basé sur Mosaic

- Propriétaire, usage non-commercial **gratuit**
- **80%** d'usage en 1996, **< 10%** en 2001
- N'est **plus maintenu** depuis 2008

Internet Explorer Lancé en **1995**, basé sur **Spyglass Mosaic**

- Propriétaire, fourni avec **Windows 95**
- IE 6 lancé en 2001 atteint **80%** du marché
- Mène au procès **antitrust** aux USA, 1998-2001

Firefox Lancé en **2002** (Netscape libéré en 1998)

- **Libre** et gratuit
- Navigation à **onglets**
- Attaque le **monopole** de IE 6

Navigateurs Web actuels

IE Développement de **IE 7** annoncé en 2005, lancé en 2006.
Meilleure **conformité aux standards**

Firefox Toujours **dans la course**

Safari Lancé en 2003, navigateur par défaut de **Mac OS X**
(succédant à Netscape puis IE sous Mac OS)

Opera Lancé en **1996**, mais **payant** jusqu'en 2000 et financé par
la **publicité** jusqu'en 2005. Maintenant **gratuit** mais
propriétaire

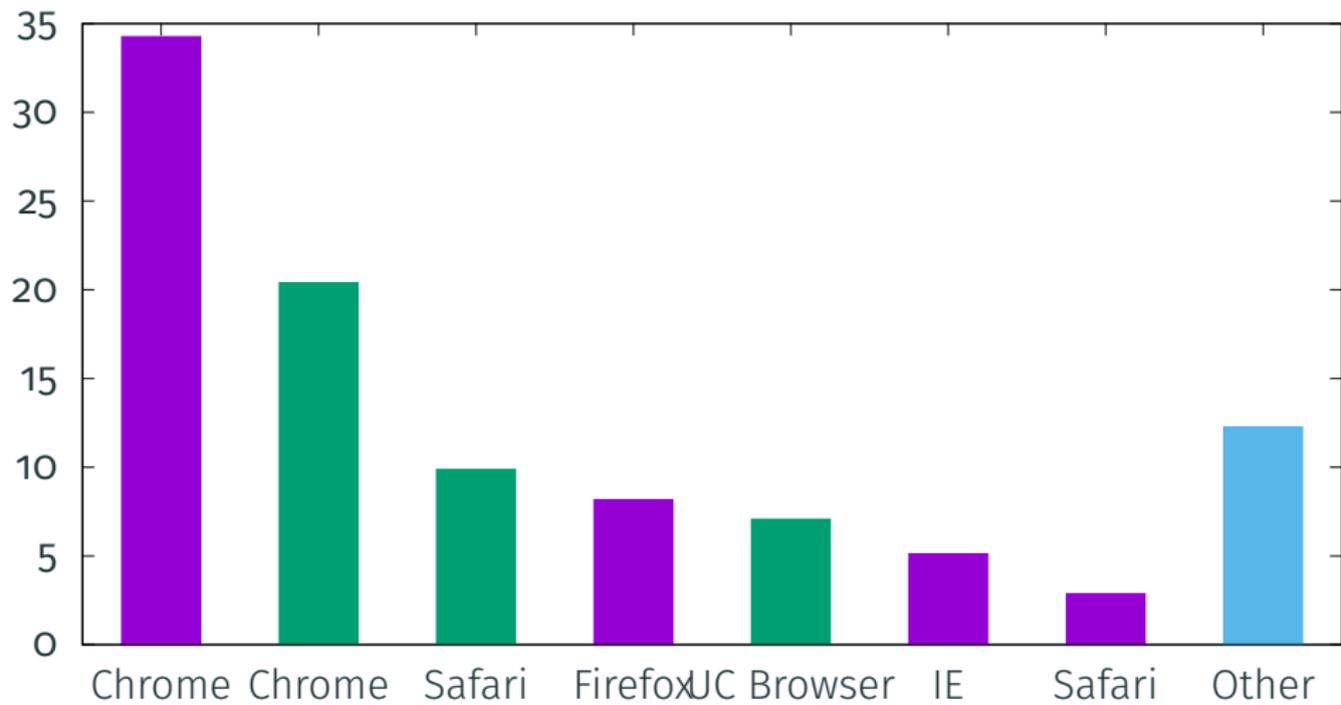
Chrome Lancé en **2008** par **Google**. Propriétaire mais version
libre également disponible

Mobile Important! **55%** d'usage d'après `gs.statcounter.com`

- **Safari** (iOS), **Android browser** ou **Chrome** (Android)
- Firefox Mobile, Blackberry, Opera, UC Browser...

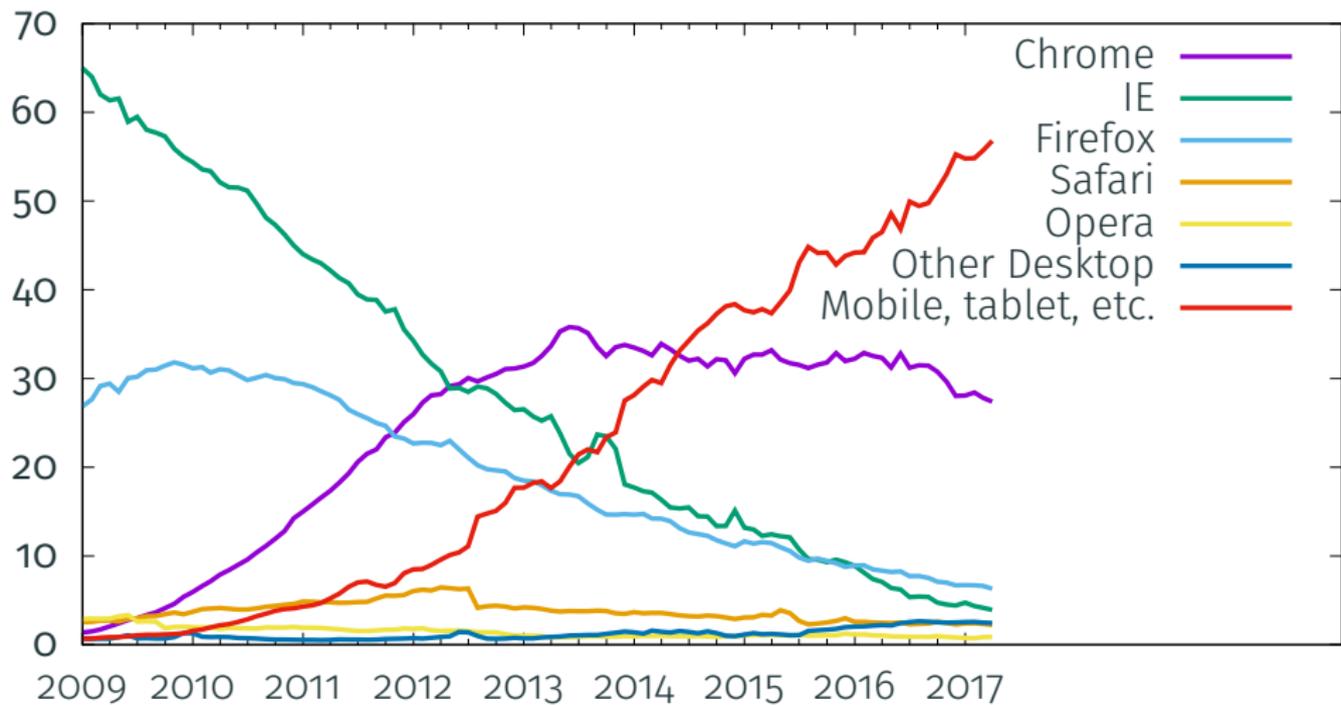
Différentes **versions**, cf `browsershots.org` pour la compatibilité

Répartition récente



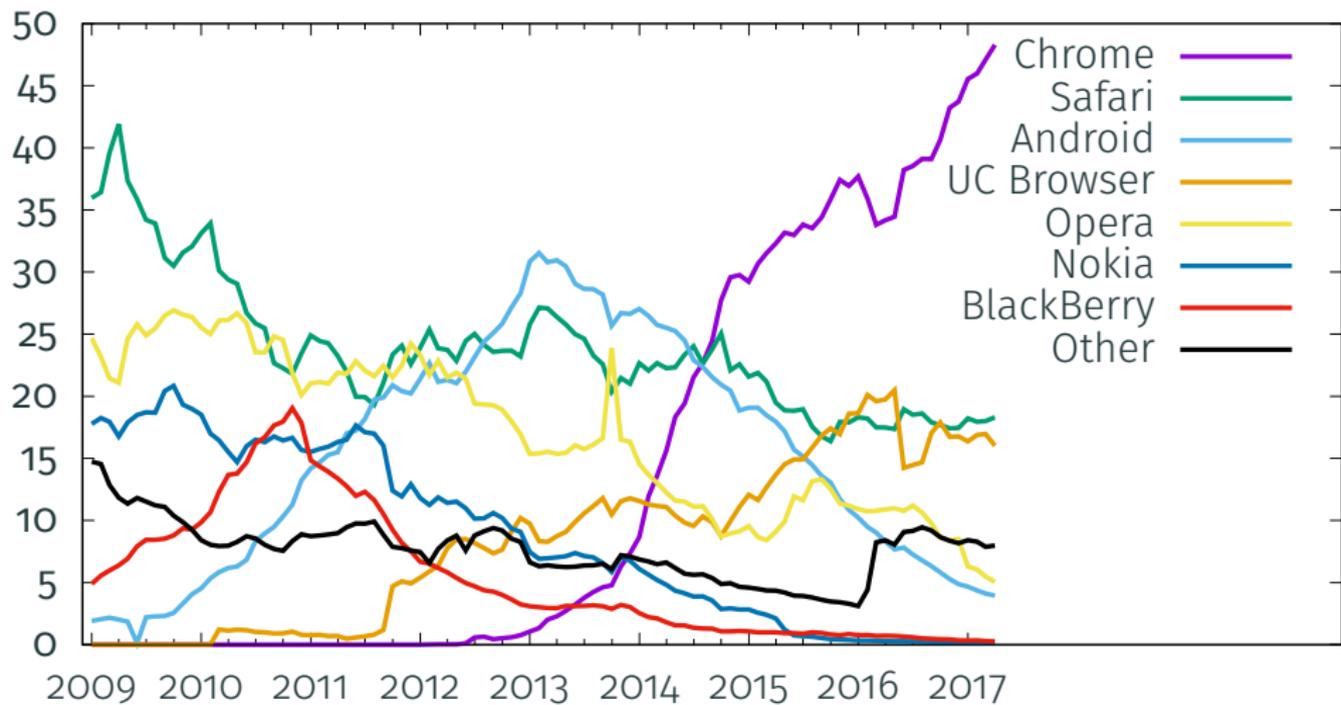
Source : gs.statcounter.com (mars 2017)

Évolution



Source : gs.statcounter.com

Évolution (mobile)



Source : gs.statcounter.com

Moteurs de rendu

Trident Internet Explorer seulement, propriétaire

Gecko Firefox et apparentés, libre et gratuit

KHTML Libre et gratuit, projet KDE

WebKit Fork de KHTML, Safari et (anciennement) Chrome, navigateurs Android et iOS, libre et gratuit

Blink Fork (avril 2013) de Webkit : Chrome, Opera

Presto Propriétaire, anciennes versions d'Opéra

Autres Dillo, autres moteurs exotiques...

Table des matières

Introduction

Navigateurs Web

Autres clients Web

URLs

Navigateurs textuels

```
sampi:~$ w3m 'http://en.wikipedia.org' (14:24:59)
```

```
Télécom ParisTech
```

```
From Wikipedia, the free encyclopedia
```

```
(Redirected from Telecom ParisTech)
```

```
Jump to: navigation, search
```

```
"ENST" redirects here. For the airport with this ICAO airport code, see Sandnessjøen Airport, Stokka.
```

```
Crystal Clear This article may need to be rewritten entirely to comply with Wikipedia's quality  
app kedit.svg standards. You can help. The discussion page may contain suggestions. \(May 2009\)
```

```
Coordinates: 48°49′35″N 2°20′47″E﻿ / ﻿48.82639°N 2.34639°E﻿ / 48.82639; 2.34639
```

```
Télécom ParisTech
```

```
Logo telecomparisTech.png
```

```
Motto L'École au coeur de la Société de l'Information
```

```
Established 1878
```

```
Type French Grande École
```

```
President Yves Poilane
```

```
Admin. staff 340 (2006)
```

```
Students 1249 (2006)
```

```
Location Paris, France
```

```
Campus Paris, Sophia Antipolis
```

```
⌂ ↑ ↓ Viewing <Télécom ParisTech - Wikipedia, the free encyclopedia>
```

Aussi : **navigateurs auditifs** pour les malvoyants...

Nombreux **programmes automatisés** sur le Web :

- **Crawlers** pour moteurs de recherche (indexation) :
Googlebot (Google), bingbot (Bing)...
- **Crawlers** plus obscurs : meanpathbot, ahrefsbot, ezoomsbot, Career-X bot, msnbot, news bot, seznambot, ChangeDetection bot? ...
- **Lecteurs RSS** et agrégateurs
- **Harvesters** d'adresses email
- Utilisateurs d'**API**

Table des matières

Introduction

Navigateurs Web

Autres clients Web

URLs

- Structure générale :

`http://en.wikipedia.org/wiki/Telecom_ParisTech#History`
protocole machine chemin fragment

- Uniform Resource Locator
- Identifie une **ressource** sur le Web
- Barre d'adresses

- Matériel de cours inspiré de notes par Pierre Senellart