

Chapitre I : Brève histoire d'Internet

Internet est la mise en réseau mondiale des ordinateurs, ce qui permet aux utilisateurs de communiquer (courrier électronique), de publier des informations (Web), de transférer des données (FTP), de travailler à distance (SSH), de discuter (messagerie instantanée comme ICQ ou MSN), etc.

1. Au commencement...

En 1957, l'URSS est la première des deux super-puissances à envoyer un satellite artificiel dans l'espace : c'est le fameux Spoutnik. Traumatisés, les États-Unis forment au sein du Département de la défense un groupe appelé ARPA ("*Advanced Research Projects Agency*"), constitué de scientifiques, chargé de concevoir des innovations technologiques appliquées à l'armée.

En 1962, l'US Air Force demande à un groupe de chercheurs de RAND (de "*Research ANd Development*", association non lucrative visant à développer les sciences et l'éducation aux États-Unis) de concevoir un réseau capable résister à une frappe nucléaire massive, afin de pouvoir riposter à son tour.

La solution est un système décentralisé, qui permet au réseau de continuer à fonctionner même si une ou plusieurs machines est touchée. *A contrario*, un système centralisé, lui, meurt dès que le centre est touché. L'idée de décentralisation est due à Paul Baran. Plus précisément, c'est lui qui pensa à un système où chaque machine, maillon d'un réseau en toile d'araignée, chercherait, à l'aide de paquets de données dynamiques, la route la plus courte possible d'elle-même à une autre machine, et où elle patienterait en cas de « bouchons. »

Le projet de Paul Baran est refusé par les militaires et ce n'est que 6 ans plus tard qu'il se concrétise.

2. Premières briques

1969 : l'Arpanet

En 1969, un réseau décentralisé se met en place sur commande de l'ARPA à BBN (Bolt Beranek and Newman Inc., une SSII de Cambridge, Mass.). Il comprend quatre grands centres universitaires américains :

- UCLA (Université de Californie à Los Angeles)
- SRI (Institut de recherche de Stanford)
- UCSB (Université de Californie à Santa Barbara)
- l'Université de l'Utah

Ces quatre centres étaient reliés par des câbles 50Kbps, et utilisaient le NCP ("*Network Control Protocol*"). C'est ce qu'on appelle l'*Arpanet*.

La date conventionnelle pour la « naissance d'Internet », c'est la date de publication de la Première RFC (« Host Software », par S. Crocker), le 7 avril 1969. C'est ce jour qui a été choisi pour célébrer l'anniversaire d'Internet.

Dr. LEILA BOUTAMINE MODULE : TICE 2ème année LMD Français groupes 6/8

Définition : qu'est-ce qu'une RFC ? RFC veut dire « *Request For Comment* ». Ce sont des documents de l'IETF (*Internet Engineering Task Force*) qui ont vocation à être les standards d'Internet. Que doivent contenir les en-têtes d'un courrier électronique, comment présenter une URL, tout cela est défini par des RFC. Le site RFC Editor contient une base de données de toutes les RFC, avec un moteur de recherche.

1971 : le courrier électronique

C'est 1971 est inventé ce qu'on appellera plus tard une *killer application* (une application qui tue des ours™), le courrier électronique. L'*e-mail* a donc un peu plus de 30 ans ! C'est l'une des utilisations les plus populaires d'Internet : chaque année, des milliards de courriers électroniques sont échangés de par le monde, et plus de 100 millions de gens possèdent une adresse électronique.

C'est Ray Tomlinson, de BBN, qui en est l'inventeur. À l'époque, Tomlinson travaille sur un système permettant à un utilisateur d'une machine de laisser un message à un autre utilisateur de la même machine (équivalent électronique d'un Post-It sur l'écran). En même temps, il teste un logiciel de transfert de fichiers via l'Arpanet. C'est en réunissant les deux concepts qu'il invente le courrier électronique. C'est également lui qui choisit l'arobase, ou arrobe (le fameux glyphe « @ » dont l'origine reste mystérieuse) comme séparateur pour les adresses électroniques.

Anecdote : que contenait le premier courrier électronique jamais envoyé ? Le premier message télégraphique de Samuel Morse était « *What hath God wrought?* », le premier message téléphonique d'Alexander Bell, « Mr. Watson, come here; I want you ». Ray Tomlinson ne s'en souvient plus bien, mais il pense que c'était : « QWERTYUIOP », la première rangée de lettres d'un clavier qwerty...

1973 : TCP/IP

En 1973 se développe ce que l'on appellera plus tard le protocole TCP/IP, l'une des pierres d'angle de l'Internet actuel, sous la houlette de Vinton Cerf, de Stanford, et de Bob Kahn, de la DARPA (nouveau nom de l'ARPA). Vous pouvez avoir quelques explications sur ce protocole dans notre cours sur les réseaux.

Ce sont ces deux hommes qui, en 1974, parlèrent pour la première fois d'« Internet ». Le protocole TCP/IP sera adopté par le Département de la défense pour l'Arpanet en 1976.

1983 : le DNS

Au début de l'Arpanet, les informations nécessaires à la connection des machines entre elles (conversion nom <-> adresse) sont contenues dans un fichier nommé hosts.txt. Ce fichier est maintenu par le *Network Information Center* (NIC en abrégé) de l'Institut de recherche de Stanford. Chaque administrateur d'une machine reliée à l'Arpanet doit envoyer ses modifications au NIC qui les centralise et redistribue périodiquement le hosts.txt mis à jour. Au fur et à mesure que l'Arpanet se développe, le système devient trop lourd à gérer : le NIC ne peut plus faire face à la charge réseau, et il y a des problèmes de collision (deux machines qui ont le même nom) qui peuvent mettre en danger le bon fonctionnement de l'Arpanet.

En 1983, pour résoudre ce problème, un groupe constitué de Jon Postel, Paul Mockapetris et Craig Partridge rédige les RFC 882 et 883 : le DNS ("*Domain Name System*") est inventé. C'est une base de données distribuée qui permet une gestion locale des noms de domaine, tout en rendant l'information disponible à tous. La base de

Dr. LEILA BOUTAMINE MODULE : TICE 2ème année LMD Français groupes 6/8

données est divisée en zones. Pour chaque zone, un ou plusieurs *serveurs de noms* (*name servers* en anglais) répond aux requêtes des *résolveurs*. Les résolveurs sont des programmes qui communiquent entre les programmes utilisés par une machine et les serveurs de noms, et permettent de faire le lien entre nom d'une machine (clipper.ens.fr par exemple) et adresse IP.

En 1984 se mettent en place les « *top level domains* », c'est-à-dire les suffixes comme .com, .gov, .net ou encore .org.

Le boom

1989 : le World Wide Web

C'est avec la naissance du Web qu'Internet s'étend au grand public. Soyons puristes : on fait généralement remonter la date de naissance du World Wide Web au texte « Information Management: A Proposal » de T. Berners-Lee : en mars 1989. Tim Berners-Lee était alors chercheur au CERN de Genève, le laboratoire européen de physique des particules. Il souhaitait ainsi fournir au plus grand nombre de chercheurs possibles un système d'information global, fondé sur le système de l'hypertexte.

Berners-Lee est l'inventeur du premier serveur Web, qu'il appelle « httpd », et du premier client Web, qu'il appelle « WWW », pour *World Wide Web*. Parmi les noms envisagés et rejetés, il y a avait « MOI » (pour *Mine of Information*, et « TIM » (pour *The Information Mine*)... À l'heure actuelle, Tim Berners-Lee travaille au MIT et dirige le W3C (« World Wide Web Consortium »), un consortium chargé de mettre au point les standards du Web. Vous pouvez consulter sa page Web.

1993 : Mosaic

C'est en mars 1993 qu'est inventé Mosaic, le premier des navigateurs grand public, doté d'une interface graphique. Son auteur est Marc Andreessen, étudiant à l'Université de l'Illinois, et assistant au NCSA (*National Center for Supercomputing Applications*). La première version de Mosaic est pour Unix, mais rapidement sortent des versions pour Windows et Mac OS. Mosaic connaît un succès et immédiat. Autre élément de l'importance de Mosaic dans l'histoire d'Internet : c'est le premier navigateur à avoir reconnu la balise IMG, autorisant ainsi l'emploi d'images sur des pages Web...

1995-2003 : la guerre des navigateurs

Marc Andreessen a continué de faire parler de lui en créant, en 1994, Netscape, un navigateur qui supplante rapidement Mosaic et règne en maître sur le marché des navigateurs, jusqu'en 1995, quand Microsoft lance Windows 95 et son propre navigateur, Internet Explorer. Menacé par ce puissant rival, Netscape, qui en est à Netscape 4, lance en 1998 le groupe Mozilla (d'après le nom de code de Netscape Communicator). Le groupe Mozilla est chargé de produire un navigateur libre et gratuit, que Netscape pourrait récupérer ensuite. Mais le groupe Mozilla décide de tout réécrire de fond en comble et tarde énormément à rendre un produit fini. Impatient, Netscape lance alors Netscape 6, un navigateur qui ne satisfait personne et sera un échec complet.

Quand le groupe Mozilla sort enfin Mozilla 1.0 en 2002, Netscape (racheté entre temps par AOL) peut lancer Netscape 7, mais c'est déjà trop tard, Internet Explorer détient plus de 90% du marché. En 2003, AOL prend acte de son échec et cesse de développer Netscape. Il aide à mettre sur pied la Mozilla Foundation, chargée de poursuivre l'écriture de Mozilla, permettant ainsi à une offre alternative de survivre.

Dr. LEILA BOUTAMINE MODULE : TICE 2ème année LMD Français
groupes 6/8

Aujourd'hui, Mozilla et ses dérivés sont des navigateurs de choix pour tous ceux qui n'utilisent pas Windows. Quant aux autres, il ne tient qu'à eux d'essayer.

Et maintenant...

L'histoire récente, vous la connaissez, vous êtes dedans ! On peut citer en 1996 le lancement d'ICQ, pionnier de la messagerie instantanée ; ou encore l'essor considérable des chats et autres salons de discussion ; Internet via un téléphone portable ou un assistant personnel, ou encore l'essor du sans-fil...

Chapitre II : Quelques concepts

QUELLE EST LA DIFFERENCE ENTRE...

Internet et le Web ?

Internet est le réseau des réseaux. Il est fait d'ordinateurs et de câbles et permet d'envoyer des blocs d'information où dans le monde. Internet propose différents types de services comme le courrier électronique ou encore le Web (la toile). Le Web est un espace virtuel d'information contenant des documents reliés entre eux par des liens hypertextes.

Http et html ?

HTTP (Hyper Text Transfert Protocol) est le protocole de communication entre les navigateurs et les serveurs. Il s'appuie sur le principe des hypertextes. HTML (Hyper Text Markup Language) est le langage du Web. Il permet de doter certains mots ou images d'une propriété d'hyperlien : en cliquant on accède à un autre document.

Un serveur et un ordinateur ?

Un serveur est aussi un ordinateur mais il contient des informations consultables par d'autres.

Un navigateur, un browser, un fureteur, un butineur ?

Aucune ! Il s'agit d'un logiciel qui permet de lire les documents sur le web : par exemple : Explorer, Fire fox...

Un navigateur et un moteur de recherche ?

Le navigateur permet de lire un document sur le Web. Le moteur est accessible grâce au navigateur et propose un service précis : trouver pour nous une adresse de site web à partir de mots-clés. Par exemple : Google.

Un site Web et une page Web ?

Un site est l'ensemble des informations mises sur Internet par une société ou par une personne. Une page web est l'élément de base d'un site.

Le titre d'une page ou d'un site et son adresse ?

Le titre d'une page est comme le titre d'un livre, c'est son nom. Le titre d'un site se trouve dans la barre d'adresse.

Doit-on taper l'adresse ou le titre sur un moteur de recherche ? Si vous avez l'adresse vous n'avez pas besoin de moteur de recherche car il a pour fonction de trouver des adresses pour vous. Il faudra donc taper l'adresse dans le navigateur. Si vous avez le titre seulement, tapez le dans le moteur entre « guillemets ».

Chapitre III : Moteur(s) de recherche

Un **moteur de recherche** est un logiciel permettant de retrouver des ressources (pages Web, forums Usenet, images, vidéo, etc.) associées à des mots quelconques. Certains sites Web offrent un moteur de recherche comme principale fonctionnalité ; on appelle alors *moteur de recherche* le site lui-même.

Moteurs et annuaires généralistes

Par ordre alphabétique.

Nom	Commentaires
Altavista.fr	Moteur généraliste, désormais lié à Yahoo!
Ask.com	Moteur généraliste, anciennement Ask Jeeves
Baidu.com	Surnommé le « Google chinois »
Google	Moteur généraliste avec services associés
Exalead.com	Moteur généraliste
Gigablast	Moteur généraliste
HotBot	Moteur généraliste du groupe Lycos
Lycos	Annuaire
Mozbot	Moteur généraliste ré-utilisant Google
MozDex	Moteur Open Source
MSN	Moteur généraliste de Microsoft
ardolab	Annuaire généraliste
Ujiko	Moteur généraliste, (c)

Voilà	Annuaire et moteur de recherche de Wanadoo
Yahoo	Annuaire

Définition

Outil de recherche sur le web constitué de « robots », encore appelés *spiders*, *crawlers* ou agents qui parcourent les sites à intervalles réguliers et de façon automatique (sans intervention humaine, ce qui les distingue des annuaires) pour découvrir de nouvelles adresses (URL). Ils suivent les liens hypertextes (qui relient les pages les unes aux autres) rencontrés sur chaque page atteinte. Chaque page identifiée est alors indexée dans une base de données, accessible ensuite par les internautes à partir de mots-clés.

Par abus de langage, on appelle également *moteurs de recherche* :

- des sites web proposant des annuaires de sites web : dans ce cas, ce sont des ressources humaines qui répertorient et classifient des sites web jugés dignes d'intérêt et non des robots d'indexation - on peut citer par exemple Voilà et Yahoo!, etc.
- des logiciels installés sur un ordinateur personnel : ce sont des moteurs dits *desktop* qui combinent la recherche parmi les fichiers stockés sur le PC et la recherche parmi les sites web - on peut citer par exemple Google Desktop et Copernic Desktop Search, etc.

On trouve également des métamoteurs, c'est-à-dire des sites web où une même recherche est lancée simultanément sur plusieurs moteurs de recherche (les résultats étant ensuite fusionnés pour être présentés à l'internaute) — on peut citer Mamma, Kartoo, Seek.fr, etc.

Fonctionnement

Le fonctionnement d'un moteur de recherche se décompose en trois étapes principales.

1. Le Web est systématiquement exploré par un robot d'indexation suivant récursivement tous les hyperliens qu'il trouve et récupérant les ressources jugées intéressantes. L'exploration est lancée depuis une ressource pivot, comme une page d'annuaire Web.
2. L'indexation des ressources récupérées consiste à extraire les mots considérés comme significatifs (pratiquement tous) correspondant à chaque ressource. Les mots extraits sont enregistrés dans une base de données organisée comme un gigantesque dictionnaire inverse. La partie *requêtes* du moteur de recherche peut ainsi rapidement retrouver les correspondances. Un algorithme gardé secret est généralement appliqué pour donner un poids variable aux correspondances, afin de pouvoir présenter les résultats des recherches par ordre de pertinence supposé. L'algorithme tient généralement compte du contexte du mot clé (titre, paragraphe, hyperlien...) et de la ressource (ressources liées, popularité du site...)

3. La ressource indexée peut ensuite être retournée dans les résultats d'une recherche d'un visiteur contenant un mot clé correspondant.

Chapitre IV : Blog

Introduction

Un **blog** ou **blogue** est un site Web constitué par la réunion d'un ensemble de *billets* triés par ordre chronologique. Chaque billet (appelé aussi *note* ou *article*) est, à l'image d'un journal de bord ou d'un journal intime, un ajout au blog ; le *blogueur* (tenant du blog) y porte un texte, souvent enrichi d'hyperliens et d'éléments multimédias et sur lequel chaque lecteur peut généralement apporter des commentaires.

1. Présentation et variété de la blogosphère

Que prétend regrouper le terme *blog* ? Un journal intime anonyme, les œuvres d'un dessinateur, l'opinion d'une journaliste, le carnet de bord d'un photographe, des satires d'hommes politiques, les vidéos d'une classe de collège, un roman en construction, les anecdotes quotidiennes d'une mère de famille, etc.

Le concept *blog* est assez vague pour autoriser toutes ces utilisations. Le phénomène blog connaît son succès grâce à une grande facilité de publication, une grande liberté éditoriale et une grande capacité d'interaction en temps réel avec le lectorat.

Même si beaucoup de blogs sont privés de visiteurs par la force des choses (abandon de milliers de blogs par mois), un blog est généralement libre d'accès, à l'image d'une vitrine survolée par de prétendus inconnus. Lecteurs et auteurs entreprennent *de facto* un contrat : tout est-il vrai ? Quels sujets aborder ? Sous quel angle ? On retrouve ces questions au sein des études de l'autobiographie. Cependant, sur un blog, la possibilité de commenter chaque billet modifie considérablement ce qu'on appelle le pacte autobiographique.

Par ailleurs, le nombre de lecteurs d'un blog varie bien plus rapidement que le nombre de lecteurs d'un livre. En effet, de par la force et la rapidité des échanges sur le Web, un blog peut devenir célèbre au sein de la communauté à laquelle il appartient en très peu de temps. Cette variabilité influe parfois sur l'auteur en le poussant par exemple à écrire autour d'un buzz (une technique marketing consistant, comme son nom l'indique à faire du bruit autour d'un nouveau produit ou d'une offre).

Les plates-formes proposant des blogs totalement privés protégés par mot de passe trouvent nombre de clients, et certaines affirment que la majeure partie de la blogosphère consiste en des blogs privés.

2. Format

La publication est généralement facilitée par l'emploi d'un logiciel spécialisé (un Système de gestion de contenu) : il met en forme le texte et les illustrations sous la direction du blogueur, il construit des archives automatiquement, il offre des moyens de recherche au sein de l'ensemble des billets et enfin gère les commentaires des internautes lecteurs. Ces outils permettent ainsi à chacun de publier du contenu, sans connaissances techniques préalables relatives à la conception de sites Web. Cette facilité d'utilisation a contribué pour une grande partie à l'explosion de l'utilisation des blogs depuis quelques années.

3. Le format invariant de la plupart des blogs

Un blog étant perpétuellement en cours d'écriture, la page d'accueil doit refléter les changements les plus récents ; l'auteur suit l'idée d'un public qui prend connaissance *régulièrement* de l'évolution du site. Les billets étant présentés figés dans le temps, l'auteur ne les modifie généralement qu'en publiant un addendum intitulé «mise à jour».

Depuis les années 2000, le phénomène blog évolue grâce aux améliorations techniques et matérielles de l'informatique et propose de publier toutes sortes de documents multimedia : images, mais surtout son et video, notamment par l'attrait de services d'hébergement (en particulier YouTube) ce qui enrichit très sensiblement la lecture ou devient le format privilégié du blog : voir photoblog et videoblog.

4. Gestion des commentaires

Un blogueur autorise souvent ses utilisateurs à laisser des commentaires, mêmes les plus critiques, suivant un formulaire Web automatisé. Cependant, le blogueur est légalement responsable de ce qui est affiché sur son site et doit donc, dès qu'il a connaissance de commentaires non conformes à la législation, supprimer les commentaires en question.

Tout blogueur peut également choisir de censurer *a priori* les commentaires en publiant lui-même ceux qu'ils jugent valides, ou de censurer totalement tout commentaire (sur un billet ou sur le blog entier.) Ce type de solution radicale est essentiellement adoptée suite à de trop nombreux trolls ou à du pollupostage.

De plus en plus de blogs deviennent le centre d'échanges approfondis au sujet duquel se passionnent auteurs et lecteurs (notamment au sujet de l'actualité ou du Web). Certains blogueurs accordent une place prépondérante aux commentaires laissés sur leur blog, qu'ils considèrent comme la raison d'existence de leurs écrits. Chaque billet revient alors à un nouvel espace public : le blog se transforme en une sorte de forum électronique et peut même recueillir une suite de commentaires sans queue ni tête, l'anonymat aidant au *blogosquat*, c'est-à-dire à un bavardage qui s'entretient tout seul, sans rapport avec le billet d'origine.

Les blogs sont souvent **centrés sur la publication d'opinions**. Ce n'est pas pour rien que c'est l'instrument favori de beaucoup de journalistes ou d'hommes politiques. Sur ce plan, c'est un intermédiaire entre la page perso (un parle et personne ne répond), et le forum de discussion (tout le monde parle à égalité).

5. Marché de la gestion de blogs

Dr. LEILA BOUTAMINE MODULE : TICE 2ème année LMD Français
groupes 6/8

Le développement de ces sociétés (Skyblog, Over-Blog, Blogger totalisent des millions de blogs personnels en 2006) est en pleine croissance : le « marché » du blog a explosé au cours des années 2004 et 2005, et tenir un blog est désormais une pratique courante.

Il s'établit également une concurrence entre blogs et média/sondeurs. Des opinions s'expriment avec force et de manière nombreuse, à la manière d'un sondage. Des opinions fortes et charismatiques peuvent même finir par diriger une partie de l'opinion ; on l'a vu lors des dernières élections présidentielles américaines et du référendum sur la constitution européenne en France.

6. Vocabulaire

Blog est issu de l'aphérèse d'un mot composé né de la contraction de « *web log* » (c'est-à-dire *carnet de bord Web*). La francophonie tente de trouver des équivalences ou des alternatives à cet anglicisme, bien que le franglais soit fréquent sur le Web, notamment parmi ses techniciens, qui rendent souvent compte de la nouveauté par le biais d'anglicismes et de néologisme.

Un **blogueur** / une **blogueuse** (en anglais *blogger*) est l'individu qui a l'habitude de **bloguer** : il écrit et publie les billets, sans entrer dans la composition de tous les commentaires qui y sont associés. La **blogosphère** est l'ensemble des blogs. Parfois, par métonymie, on désigne l'ensemble des blogs d'une communauté précise.

Référence bibliographique

Benzeroual. T (2018) : cours TICE, Département de Français. Université Batna2