Le Hublot

Journal des tuteurs — Numéro 7 — Mai 2000

Voici le dernier numéro du Hublot *pour l'année 1999–2000 et l'épilogue de nos grandes sagas de l'année... On revient encore plus top mieux en octobre, avec les conscrits, les vendanges, les bonnes résolutions... D'ici là, on vous proposera au mois de juin les divers cours du* Hublot *en fascicules séparés : typographie, cours Unix, réseaux, Netiquette, Netscape.*

Nous sommes toujours à votre disposition pour toutes questions, et surtout pour toutes suggestions pour l'an prochain... `tuteurs@clipper`*. Vous pouvez aussi mettre un mot dans le casier d'Émilia Robin. Certaines pages de notre site sont indiquées dans les articles : on suppose que toutes les adresses indiquées commencent par* `http://www.eleves.ens.fr:8080/tuteurs/`*. Certains mots sont suivis d'un astérisque*; ils renvoient au lexique en fin de numéro. Vous pouvez retrouver les anciens numéros du* Hublot *sur notre site :* `http://www.eleves.ens.fr:8080/tuteurs/hublot.html`

Merci au SPI pour son soutien, à Jacques Beigbeder, Éric Guichard et Daniel Béguin pour leurs articles, aux divers rédacteurs : Pierre Corbineau, Nicolas George, David Madore, David Monniaux, Thomas Pornin, Gilles Radenne, et aussi à tous ceux qui ont contribué, de près ou de loin, par leurs relectures et/ou leurs conseils, à la rédaction du Hublot *: Hélène, Hélène encore, Marie-Lan, Étienne, Verzi, MXK, Éric, Éric encore, Zpeat, Max, Isil, P'titboul, Olivier, Yann, Alain, Damien, Seb, Edd, Tmoy, Sbi... Et merci à tous ceux qui ont pris la peine de nous écrire pour nous donner des conseils et des encouragements !*

Dossier : les virus

On parle beaucoup de virus en ce moment dans la presse, avec le récent virus I Love You*, qui s'attaquait aux machines fonctionnant sous Windows et utilisant Outlook, par l'intermédiaire d'un fichier attaché. Un courrier collectif avait été envoyé à ce sujet le 5 mai par Jacques Beigbeder. Un filtre a été mis en place à l'entrée de l'École pour rejeter les courriers électroniques ayant comme sujet ILOVEYOU; environ 35 courriers ont été rejetés (ce qui est peu, sachant que l'ENS reçoit en moyenne 30 000 courriers électroniques tous les jours...). Nous profitons de l'occasion pour faire le point sur cette question des virus, et sur ce que la presse en a dit, également...*

Un virus informatique est un programme, de même qu'un traitement de textes ou un logiciel de calcul scientifique. Sa particularité est qu'il opère en grande partie à l'insu de l'utilisateur de la machine, et qu'il essaye de se transmettre à d'autres machines ou d'autres utilisateurs. Il peut n'avoir aucun comportement à part la réplication, ou créer des troubles plus ou moins graves (animation non désirée sur l'écran, remplacement de certains mots dans le texte du document édité, voire destruction complète de certaines données).

Comment se transmettent les virus ?

Il y a quelques années, les virus se transmettaient exclusivement par l'utilisation de programmes déjà infectés, c'est à dire de programmes normaux (traitement de textes, jeux...) modifiés par un virus afin de contenir une copie de lui-même. Lors de l'utilisation du logiciel en question, le virus s'active et essaye de se recopier au sein des logiciels installés sur le disque dur de la machine, voire du système (p.ex. Windows) lui-même. Le logiciel est ainsi installé à demeure sur la machine de l'utilisateur, qui ne soupçonne rien. À chaque fois que l'utilisateur recopie un logiciel (ex : un jeu) installé chez lui et le donne à un autre utilisateur, il copie le virus par la même occasion[1].

Certains virus récents, dont les très médiatiques *Melissa* et *I Love You*, se reproduisent au travers du courrier électronique. Leur mode d'action est le suivant : on peut attacher des documents (images, sons, textes...) au courrier. Lorsque l'utilisateur demande la visualisation du document, certains visualisateurs (p.ex. Microsoft Word, Microsoft Excel) exécutent des commandes incluses dans le document. Le virus va ainsi se faire

[1] Rappelons que la copie de logiciels dont vous n'êtes pas l'auteur n'est autorisée que si l'auteur vous le permet.

passer pour un banal texte, et lorsque l'utilisateur va demander à voir ce texte, il va lancer le virus au passage. Le virus va alors récupérer des adresses de courrier électronique et se reexpédier et/ou s'installer au sein du système.

Quelles sont les actions d'un virus ?

La plupart des virus transmis par courrier électronique sont essentiellement ennuyeux à cause du flot de courrier qu'ils génèrent. Le récent *I Love You* efface certaines données des utilisateurs, dont les fichiers JPEG (images photographiques), ce qui explique l'émoi qu'il a suscité auprès des journalistes.

Il n'existe aucun obstacle technique à ce qu'un virus fasse des choses plus insidieuses, comme récupérer des données personnelles (numéros de cartes bancaires, adresses, textes contenant certains mots-clefs) sur le disque dur et les renvoyer à l'auteur du virus... d'une manière éventuellement intraçable !

Comment se fait-il que les virus existent ?

— **Virus « conventionnels »** — Les systèmes d'exploitation Microsoft Windows 95, 98 et Millenium et Apple MacOS permettent à (presque) n'importe quel programme fonctionnant sur la machine de modifier n'importe quelle partie du système. Au contraire, Windows NT et 2000 et Unix (dont Linux) ne permettent pas aux programmes lancés par de simples utilisateurs de changer des paramètres globaux au système. Il est ainsi impossible (sauf bogues dans le système) à un simple utilisateur des machines Unix de l'ENS de lancer un virus qui se mettrait en place dans un traitement de texte utilisé par les autres utilisateurs[2].

En résumé, sous Windows, lancer le moindre programme de source inconnue (ex : un petit gadget qui affiche une animation sur l'écran, un jeu...) c'est s'exposer à ce qu'il installe des choses non désirées dans le système. Il n'est donc pas étonnant que certaines entreprises interdisent formellement à leurs employés l'installation de logiciels non approuvés !

— **Virus par courrier électronique** — Les logiciels de la société Microsoft (Outlook, Word, Excel...) offrent souvent la possibilité d'adjoindre à un document (texte, feuille de tableur...) des petites commandes *(macros)* exécutées automatiquement à l'ouverture du document. Cela peut être pratique par exemple pour offrir à l'utilisateur des options de remplissage automatique dans un formulaire, des menus... Jusque là, rien que de très normal. Ce qui l'est moins, c'est qu'une palette très large de commandes, y compris l'écriture de fichiers sur le disque dur, est permise dans des documents affichés directement depuis le courrier électronique.

On ne peut qu'être atterré par un tel manque de sérieux de la part d'une société qui se veut aussi professionnelle que Microsoft. **En résumé, leur système de courrier électronique exécute n'importe quels ordres contenus dans un courrier électronique de quelque provenance que ce soit lorsque l'utilisateur demande à voir le contenu du courrier**. C'est un peu comme si dans une entreprise le service chargé du courrier payait n'importe quelle facture lui arrivant par la poste, sans prendre la peine de vérifier son bien-fondé !

Les virus « mutent »-ils ?

Les virus naturels ont la particularité de pouvoir évoluer au cours du temps, et la sélection naturelle favorise les souches résistantes aux « contre-mesures ». Dans le cas des virus informatiques, dans l'état actuel des choses, rien de tel : les « mutations » sont en fait des modifications apportées par des programmeurs au code du virus, afin de produire leur propre version. Le virus *I Love You* effaçait les images Jpeg et les sons MP3, mais des modifications minimes peuvent le transformer en virus à effacer les documents Word (attention à vos mémoires !). Comme de telles modifications sont à la portée de n'importe quel programmeur moyen, le nombre de versions dérivées d'un virus à succès peut être considérable, et il n'est pas sûr qu'un anti-virus capable de détecter l'original détecte la version modifiée.

Quelles parades ?

Comme nous l'avons vu plus haut, le problème des virus dans le courrier électronique découle d'une faute grossière de conception du système de courrier électronique de Microsoft. Une première parade est de ne pas utiliser ce système. Les utilisateurs de machines Unix à l'ENS n'ont donc rien à craindre.

Si on doit lire son courrier sous Windows avec Outlook, il faut désactiver l'exécution automatique des attachements, ce qui est l'option par défaut.

Les logiciels anti-virus sont une parade souvent insuffisante. Ils sont essentiellement faits pour découvrir des virus déjà connus, alors que des nouveaux virus sont bricolés chaque semaine. De plus, on peut parfois se demander (en étant vraiment mauvaise langue) s'il n'existe pas un lien entre les auteurs de virus et les éditeurs de logiciels anti-virus (comment disent les Dupon{d,t} : chercher à qui profite le crime !).

N'a-t-on vraiment rien à redouter si on n'utilise pas Windows ou Macintosh ?

Il serait faux de dire qu'il est impossible d'écrire un virus pour Unix. C'est simplement largement plus difficile (il faut trouver des problèmes de sécurité compliqués à exploiter) et inefficace, notamment en raison des dispositifs de sécurité empêchant à un utilisateur de modifier les données d'un autre utilisateur. De plus, la diversité des environnements (Suns, PC Linux, etc...) rendrait le virus encore plus compliqué à écrire.

Les avertissements contre les virus

Écrire un virus du type *Melissa* ou *I love you* ne nécessite pas d'être un génie de l'informatique, loin de là. Mais il y a encore plus simple : écrire un avertissement

[2] NDLR : cela s'explique par le système des droits. Voir le numéro 5 du *Hublot*.

contre un virus... qui n'existe pas, lui adjoindre un petit message du style « C'est vraiment très important, avertissez tous vos amis », et attendre que des utilisateurs pleins de bonnes intentions propagent le message. Il s'agit pourtant d'une sorte de virus, qui se dissémine à l'aide de la naïveté de certains, et qui fait partie d'une famille fort vaste de « lettres en chaînes ». Outre les avertissements de virus, citons aussi les rumeurs du style « des trafiquants de drogues distribuent des bonbons au LSD aux enfants... avertissez tous les parents que vous connaissez ».

Inutile de dire qu'il convient de ne pas disséminer ce genre d'« informations ». L'école dispose du Service de Prestations Informatiques, dont les personnels se tiennent constamment au courant des problèmes de sécurité informatique. C'est à eux, et non pas aux élèves, de diffuser des avertissements concernant les virus. C'est eux (`spi@ens.fr`) qu'il faut contacter en cas de doute.

▷ David Monniaux, Laboratoire d'Informatique

Sur un article du *Monde*

Dans *Le Monde interactif*, supplément informatique du journal *Le Monde*, daté du mercredi 10 mai 2000, on lit en première page un grand titre qui proclame : « Le virus I Love You révèle les faiblesses du Web ». En un mot, c'est du n'importe quoi, mais du n'importe quoi en première page d'un journal réputé...

C'est quoi le Web ? Le Web, c'est un réseau d'informations, reliées entre elles par des liens (liens hypertexte si on veut faire branché). Quand on suit un lien en cliquant dessus, on est amené sur une autre page. Il se peut que la nouvelle page soit physiquement sur le même ordinateur, ou au contraire à 12 000 kilomètres de là, ça n'a aucune importance.

Le Web est l'un des aspects d'Internet. Internet, c'est la mise en réseau mondiale d'ordinateurs. Ça implique une infrastructure matérielle (câbles, routeurs...), et une série de protocoles pour faire transiter les informations. Par exemple, le Web utilise le protocole HTTP. Mais Internet c'est aussi le courrier électronique, le transfert de fichiers, la connexion à distance (telnet), les news... Pour chacune de ces fonctions, on utilise un logiciel : navigateur pour le Web, logiciel de courrier, etc. Si vous voulez, Internet c'est comme une route que les gens utilisent pour différentes fonctions (transport de marchandises, vacances, trajet boulot-dodo, balade à vélo...), et ces fonctions c'est le courrier, le Web, etc.

Tout ça pour dire quoi ? I Love You est un virus qui utilise un trou de sécurité d'*un* logiciel de courrier électronique, Outlook. Pour continuer la comparaison routière, imaginons qu'un fabricant de pneus (appelons-le Pierralu) a commercialisé une série de pneus défectueux, qui explosent. Les journaux parleraient de la mauvaise qualité des pneus Pierralu, mais ne titreraient pas « Les pneus qui explosent révèlent les faiblesses de la route »...

Il faudrait cesser de confondre Internet et le Web, et le réseau avec les logiciels qui l'utilisent...

Pratique

Typographie : les bibliographies

Le mois dernier, on a expliqué l'emploi des diverses fontes. Pour clore ce cours de typographie, on va voir quelles sont les principales règles françaises de présentation d'une bibliographie. Ces informations proviennent d'une part du *Lexique des règles typographiques en usage à l'Imprimerie nationale*, Imprimerie nationale, 1990, d'autre part d'une mise au point très claire de J.-M. Burlat et B. Prudhomme[3].

— **Normes internationales** — Il existe deux normes internationales de présentation des références bibliographiques. L'ISO 690 (1987) décrit le « contenu, la forme et la structure » des références bibliographiques, et l'ISO 690-2, datant de 1995, traite du référencement des documents électroniques[4].

Ces deux normes de l'International Standard Organisation (ISO) définissent la nature des informations à présenter et l'ordre dans lequel on les présente. C'est une description *logique*. La typographie proprement dite, la mise en page *graphique* n'est pas prescrite par la norme, et varie selon les pays. Il n'existe d'ailleurs pas de norme en la matière, mais des usages, que voici.

— **Où ?** — La bibliographie est placée le plus souvent en fin d'ouvrage, avant le ou les index, et les tables (des matières, des figures, etc). Dans les ouvrages d'érudition, on peut la trouver en début d'ouvrage, après les introductions.

— **Comment ?** — Une bibliographie se présente comme une liste numérotée, avec les numéros entre crochets. Le classement interne dépend du contexte : ordre alphabétique (bibliographie courte), classement thématique, classement selon la nature de l'ouvrage...

— **Les auteurs** — On indique le nom de famille en petites capitales, suivi d'une virgule et du prénom. Quand on connaît le prénom complet, on l'indique, sinon on se contente des initiales. N'oubliez pas le tiret des prénoms composés !

Quand il y a deux auteurs, on met « et » entre les deux noms ; quand ils sont trois, on n'écrit « et » qu'avant le dernier, comme dans une énumération classique. Si l'ouvrage compte plus de trois auteurs, on n'indique habituellement que les trois premiers, suivis de « *et. al.* ».

[3] Burlat J.-M., Prudhomme B., *Références bibliographiques — Rédaction et lecture.* 5ᵉéd. [On-line]. Villeurbanne (Fr) : Inst. Nat. Sci. Appl., Doc'INSA, Sep. 1997. Available from internet : <URL : `http://csidoc.insa-lyon.fr/docs/refbibli.html`>.

[4] Il existe aussi des normes françaises, éditées par l'AFNOR, principalement les normes Z 46-001, Z 46-002 et Z 46-003, qui reprennent le contenu des normes ISO.

La norme dit qu'il faut séparer les divers champs de la référence (auteurs, titre, édition...) par des points, ou des points-virgules, en réservant la virgule pour les distinctions à l'intérieur d'un champ. C'est une question d'usages. Si on utilise le point comme séparateur, il va de soi que, le cas échéant, on supprime le point d'abréviation (voir le dernier exemple ci-dessous) :

[1] Bitsch, M.-T., ...

[2] Knuth, Donald E., ...

[3] Poidevin, R. et Bariéty, J. *Titre...*

— **Le titre** — Le titre est écrit en italique. Dans le cas d'un article ou d'une contribution, le titre du passage est indiqué entre guillemets, puis on écrit en italique le nom de l'ouvrage ou de la revue dont le texte est extrait. Faut-il mettre « in » ou « dans » entre le titre de l'article et l'ouvrage ? Ce n'est pas clair... L'important est d'être cohérent, et de s'en tenir à la méthode choisie (pas comme dans nos exemples, donc).

[4] Rolland, Christian, *LaTeX par la pratique*, Paris, O'Reilly, 1999.

[5] Pornin, T., « Des pingouins partout », *Le Hublot*, mars 2000, n^{o} 5, p. 7–8

[6] Ménudier, Henri, « Kurt Schumacher et la nation allemande », in *l'Allemagne 1945–1955. De la capitulation à la division*, Krebs, G. et Schneilin, G., dir., Paris, Université de la Sorbonne Nouvelle, col. « Publications de l'Institut d'Allemand », n^{o} 22, 1996, p. 255–284.

Pour citer un ouvrage écrit en caractères non latins (grec, russe, arabe, chinois...), on a le choix entre utiliser une police exotique, ou recourir à la translittération en usage. Dans tous les cas, on indique ensuite la traduction du titre, entre crochets :

Quand on cite juste un chapitre ou un tome, on commence par donner les références de l'ouvrage, puis on indique la sous-partie.

[7] Molchanov, Nicolas Nicolaiévitch, *Vnechnaya politika Francii, 1944–1954* [*La politique extérieure de la France, 1944–1954*], Moscou, 1959, 404 p.

[8] Fontaine, André, *Histoire de la guerre froide*, Paris, Le Seuil, col. « Points histoire », n^{o} 65, 1983, 580 p. Volume 2, *De la guerre de Corée à la crise des alliances, 1950–1971.*

— **L'édition** — Après le titre, on indique le lieu d'édition, l'éditeur, l'année. On peut mentionner la collection, mais ce n'est pas obligatoire. Pour le lieu d'édition, il peut arriver que plusieurs villes soient indiquées ; vous pouvez ne mentionner que la première.

Dans le cas d'un article, on indique dans l'ordre, après le titre du périodique, l'année, le volume et le numéro.

[9] Guillen, Pierre, « La France et l'intégration de la RFA dans l'OTAN », *Guerres mondiales et conflits contemporains*, juillet 1990, vol. 40, n^{o} 159, p. 73–91.

— **Les pages** — Pour un ouvrage, on donne le nombre de pages : « 579 p. ». Pour un article, une contribution, on indique la pagination sous cette forme : « p. 123–148 » ; on met un tiret et non un trait d'union entre les chiffres. Attention : l'abréviation pour « page » est « p. » et non « pp. » ; n'oubliez pas de mettre un espace insécable entre l'abréviation et le numéro.

Commande du mois : `ispell` et `gospell`

Le correcteur orthographique le plus répandu dans le monde Unix s'appelle `ispell`. Il se manipule de manière très simple. Dans un terminal, il suffit de taper `ispell fichier`, et la correction commence... En anglais, qui est la langue par défaut. Pour obtenir une correction en français, il faut préciser sur la ligne de commande :

```
bireme-~ $ ispell -d francais orthographe.txt
```

On a alors droit aux commandes de correction en plein écran :

```
Une grose faute d'orthographe.

 0: glose      4: grouse
 1: grise      5: prose
 2: gros       6: rose
 3: grosse

[SP] <number> R)epl A)ccept I)nsert L)ookup U)ncap
Q)uit e(X)it or ? for help
```

Le programme nous montre les mots mal orthographiés, et propose des orthographes approchantes. Il va aussi souvent demander confirmation pour la décomposition d'un mot, en particulier quand il y a des apostrophes. Les options offertes sont alors :

* `espace` : ignorer la faute ;
* *`numéro`* : remplacer par un des mots proposés ;
* `R` : remplacer manuellement ; le(s) mot(s) alors saisis seront corrigés à leur tour ;
* `A` : accepter ce mot pour le reste du texte ;
* `I` : insérer le mot dans le dictionnaire personnel ; il sera alors accepté pour tous les autres textes ;
* `L` : chercher un mot dans un dictionnaire (hélas, toujours dans un dictionnaire anglais) ;
* `U` : comme `I`, mais en considérant le mot en minuscules ;
* `Q` : abandonner la correction ;
* `X` : arrêter là et sauver les corrections faites.

— `gospell` — Comme les accents posent des problèmes, on utilise un script* local à l'École, `gospell`, qui lance `ispell` avec les bonnes options. Il suffit de taper `gospell -f` pour le français, ou `gospell -e` pour l'anglais.

Pour finir avec les remarques, il faut signaler que le dictionnaire français n'est pas installé sur les machines Linux (`chaland` en salle S, `tremble` en salle T, `buyse` en Infi, `gondole` à Jourdan, `polacre` et `ponton` à Montrouge).

— **Sous *Emacs*** — *Emacs* fait tout, même le café. Il sait notamment utiliser `ispell` pour vous permettre de corriger confortablement le texte que vous êtes en train de taper. Pour cela, vous devrez d'abord probablement changer le dictionnaire à utiliser. Pour ce faire, tapez `Méta-X`, puis `ispell-change-dictionary`, et indiquez quel dictionnaire utiliser. Ensuite, vous pouvez lancer la correction par `Méta-X`, puis `ispell-buffer`. Avec *XEmacs*, vous pouvez également utiliser le bouton dans la barre d'outils.

Une petite fenêtre s'ouvre alors au-dessus de votre document, dans laquelle apparaissent les propositions d'`ispell`, qui sont les mêmes que quand on l'utilise tout seul.

▹ Nicolas George

Questions-Réponses

À quoi servent les sauvegardes ? En cette période de rapports, et bientôt de vacances, voici quelques conseils de bon sens.

Les plantages des ordinateurs fonctionnant sous Windows ou MacOS peuvent entraîner des pertes irréparables de données. Sachez que vous ne pourrez pas vous retourner contre le fabricant du logiciel, car la licence vous avez acceptée possède une clause qui exclut toute responsabilité pour les dégâts que pourrait causer le produit. D'autre part, quel que soit le système, un disque dur est un dispositif électromécanique pouvant être sujet à des pannes matérielles. Vous devez donc faire des sauvegardes régulières de vos données, au moins tous les mois[5]. Voici quelques conseils :

☞ Ne laissez jamais les sauvegardes dans la même pièce que l'ordinateur, en cas d'incendie vous perdiez tout (ne riez pas, c'est arrivé au département de physique). Sans parler des inondations venues du plafond, ou de la tasse de café...

☞ Rangez vos supports amovibles (disquettes, bandes, ZIP) loin de toute source de perturbations électromagnétiques, comme c'est d'ailleurs expliqué sur les boîtes de disquettes.

☞ Certains laissent avec confiance leurs données sur les Macs ou PC sous Windows en libre accès à l'École ; outre l'absence de confidentialité, n'importe quelle personne mal intentionnée peut les modifier, voire tout effacer (c'est impossible sur les Suns et les PC sous Linux, sauf si vous oubliez de vous déloguer). Même sans mauvaises intentions, il peut s'agir de quelqu'un qui a besoin de libérer de la place pour utiliser le scanner, par exemple. Ayez donc toujours une sauvegarde de votre travail !

☞ D'autre part, rappelez-vous bien qu'une disquette n'est pas un support fiable : faites toujours vos transferts en double exemplaire, sur deux disquettes ; à fortiori, **faites toujours au moins DEUX sauvegardes**. Ça nous fait froid dans le dos et mal au cœur de voir des affichettes « Perdu disquette orange très importante le 12 mars », et de savoir qu'il y a des gens qui confient à un seul support leur travail de l'année, ou même simplement de la semaine...

Enfin, sachez qu'à l'ENS, le Service de Prestations Informatiques effectue une sauvegarde des comptes des élèves (sur les Suns et les PC sous Linux) toutes les nuits, sur des bandes magnétiques.

Unix pas à pas

Ce mois-ci, on aborde la question des formats de fichiers. « Format » est un mot-valise, fourre-tout, bien pratique, et pas clair du tout... Dans la partie « Commande », on va voir la commande `file`, qui sert à identifier un fichier. Enfin, dans la partie « Usage », on passe en revue les principaux formats de fichiers, en indiquant quels logiciels utiliser.

Concept : « format de fichier »

Les fichiers stockés sur le disque dur d'un ordinateur contiennent des informations de nature diverses : textes, images, programmes... Toutes ces informations sont toutes écrites sous la forme d'une suite de nombres compris entre 0 et 255. Un format est une façon de décrire un objet informatique de façon à conserver toutes les informations utiles. Par exemple, pour une image, on veut la couleur de chaque point, mais aussi la taille de l'image, le nombre de couleurs, etc. On peut en outre vouloir ajouter une brève description, ou une notice de copyright. Bref, tout n'est pas si simple.

Au début, quand on écrivait un nouveau logiciel, on inventait une nouvelle manière de représenter l'information, c'est à dire un nouveau « format de fichiers ». Mais quand il existe déjà un logiciel similaire, pour peu que le format de fichiers qu'il utilise soit suffisamment général, et qu'on en connaisse les spécifications, on a tout intérêt a utiliser le même format pour garder la compatibilité. C'est pourquoi il existe aujourd'hui un certain nombre de formats de fichiers très répandus, et compris par la plupart des logiciels concernés. Il existe aussi un grand nombre de formats utilisés spécifiquement par un logiciel.

Un fichier ayant un format donné est généralement découpé en deux parties : d'une part l'entête, et d'autre

[5] Une sauvegarde tous les mois, ça veut dire que vous vous autorisez à perdre, éventuellement, un mois de travail.

part, les données elle-mêmes. L'entête contient (souvent) un code identifiant le format, et permettant au logiciel de vérifier qu'il est bien en train de travailler avec un de ses documents, des indications sur la variante du format utilisée, des paramètres numériques (taille par exemple), etc... Ensuite viennent les données, écrites selon le code du format.

L'usage veut que l'on ajoute au nom des fichiers un suffixe qui identifie son format. Par exemple, un fichier de son au format MP3 a un nom qui se termine en « `.mp3` »

— **Format texte et format binaire** — Un fichier est dans un format texte si les données qu'il renferme sont exprimées sous forme de texte à peu près lisible par une personne avertie, codé en utilisant le standard ASCII ou une de ses extensions internationales, à l'exclusion notamment des caractères dits « de contrôle ». En conséquence, un format texte est utilisable dans un éditeur de texte basique et modifiable à la main tandis qu'un format binaire ne l'est que par des outils spécifiques.

Un fichier est dans un format binaire si au contraire il utilise avant tout le fait que les fichiers sont des séries de nombres. Il est illisible pour un œil humain, et ne peut être lu qu'avec un logiciel spécifique.

Commande : `file`

La commande `file` permet de connaître le format d'un fichier. C'est très utile lorsqu'un fichier est renommé par accident, ou quand on tombe sur un fichier inconnu dont on ne connaît pas l'extension. Il suffit de taper `file` suivi du nom du fichier.

```
bireme-~ $ file hublot6.*
hublot6.aux:        ASCII text
hublot6.dvi:        TeX DVI file (TeX output 2000.04.21:1635 )
hublot6.html:       HTML document text
hublot6.image.tex:  ASCII text
hublot6.log:        TeX transcript text
hublot6.ps:         PostScript document text conforming at level 2.0
hublot6.tex:        data
```

Une remarque sur le format « data » : c'est ce que `file` répond quand il ne sait rien dire sur le fichier...

Usage : formats courants

— **Textes ASCII** — Il y a d'abord le format texte le plus élémentaire : juste des mots et rien d'autre... Le fichier peut avoir un suffixe « `.txt` ». Dans cette rubrique « textes ASCII » on inclut les fichiers correspondant à un langage spécifique, comme :

Fichiers LaTeX	: suffixe « `.tex` »
Fichiers HTML	: suffixe « `.html` »
Langage C	: suffixe « `.c` »
Langage C++	: suffixe « `.C` » ou « `.cc` »
Langage Perl	: suffixe « `.pl` »
Langage Fortran	: suffixe « `.f` »

Pour un fichier de texte, vous avez deux solutions : soit ouvrir le fichier avec un éditeur de textes, et vous pourrez alors modifier le fichier ; soit vous contenter de le regarder avec le programme `less` (que l'on quitte en tapant `q`) :

```
bireme-~ $ less remarques
```

Les fichiers HTML sont un cas particulier : à moins de vouloir voir le code, c'est le résultat qui vous intéresse. Dans ce cas, vous ouvrez le fichier avec Netscape ou un autre navigateur.

— **Visualisateurs** — Le format « `.dvi` » (binaire) est le résultat de la compilation LaTeX.

Le PostScript (« `.ps` ») est un langage de programmation utilisé pour décrire les documents imprimés. Grâce à son caractère généraliste, il s'est répandu très vite depuis sa création par Adobe Systems jusqu'à devenir un standard de fait.

Le PostSript encapsulé (« `.eps` ») est un fichier Postscript comportant en outre des indications, sous la forme de commentaires dans le langage Postscript, permettant d'obtenir des informations sur le fichier (telles que sa taille) sans avoir recours à la complexité d'un interpréteur Postscript. Il sert pour les fichiers PostScript destinés à être inclus dans un document.

Le format **PDF** (*Portable Document Format*), de la société Adobe, sert à diffuser des textes sans qu'il soit possible de les modifier (à moins d'acheter le logiciel idoine). En revanche, le visionneur est gratuit.

« `.dvi` »	: `xdvi`
« `.ps` »	: `gv`
« `.eps` »	: `gv`
« `.pdf` »	: `acroread`, `xpdf`

— **Traitements de textes** — Les traitements de textes produisent des formats généralement binaires :

Fichiers Word	: suffixe « `.doc` »
Fichiers ApplixWare	: suffixe « `.aw` »
Langage StarOffice	: suffixe « `.sdw` »

Pour lancer ApplixWare, tapez `asterix`, pour lancer StarOffice, tapez `soffice`.

Les traitements de textes créent un fichier dans leur propre format, qui n'est pas lisible par les autres traitements de textes, à moins que ceux-ci ne disposent d'un filtre d'importation. Pour que celui-ci fonctionne, il faut que le fichier importé soit dans un format « propre », c'est-à-dire non corrompu par des informations annexes. Désactivez donc la « sauvegarde rapide » de Word, qui ne sauvegarde pas le fichier dans un format Word correct, et insère au passage des extraits d'autres fichiers sans rapport.

Il existe un format particulier, appelé **RTF** (*Rich Text Format*)[6]. C'est un format créé par Microsoft pour servir de pivot entre différents formats de traitements de textes. Il a en particulier l'avantage de conserver les accents, d'être compris par tous les traitements de textes. En contrepartie, il ne conserve pas forcément toutes les informations de mise en page, et n'est pas toujours compatible avec lui-même. Cependant, nous ne saurions trop vous conseiller d'envoyer vos documents

[6] Remarque : le RTF est un format texte et pas binaire.

au format RTF, au lieu d'imposer à votre correspondant d'avoir telle version de tel traitement de textes.

— **Archives** — Il existe un plusieurs formats de fichiers compressés, dont :

Compresser	Suffixe	Décompresser
`gzip`	`.gz`	`gunzip`
`compress`	`.Z`	`uncompress`
`bzip2`	`.bz2`	`bunzip2`

Les archives sont des « paquets » de fichiers rassemblés en un seul fichier :

« `.tar` » : Archive simple. Désarchiver avec `tar xvf`.
« `.tgz` » : (ou « `.tar.gz` »). Archive compressée. Désarchiver avec `tar zxvf`
« `.zip` » : Archive compressée. Se désarchive avec `unzip`.

— **Images** — Il existe un grand nombre de formats d'images, texte ou binaires : XPM, XBM, JPG, GIF, PNG, TGA, BMP, TIFF... Pour visionner une image, utilisez le programme `xv`. Vous pouvez aussi utiliser le logiciel de dessin `gimp`, dont le format natif est « `.xcf` ». Les images vectorielles[7] ont un suffixe « `.fig` ». On les regarde avec le logiciel de dessin `xfig`.

— **Sons et animations** — Il existe différents formats : « `.au` », « `.wav` », « `.mp3` », fichiers pour séquenceur (MIDI), fichier modules « `.mod` », etc[8].

Les formats vidéos sont le MPEG (divers suffixes : « `.mpe` », « `.mpg` », « `.mpeg` »), le format AVI (« `.avi` »), et le Quicktime (« `.mov` », « `.qt` ».)

Fichiers MP3	: `mpg123`
Autres fichiers son	: `xplay`
Format MPEG	: `xplay`
Format AVI	: `xanim`

Internet

Dans cette section, vous trouverez la fin du cours de Thomas sur les réseaux; ce mois-ci, il présente différents problèmes qui peuvent se produire sur un réseau informatique. On termine le cours sur Netscape en expliquant comment télécharger des documents.

Cours : qu'est-ce qu'un réseau ?

On va clore ce cours avec un exposé des problèmes les plus fréquents sur un réseau. Lisez au moins le « plantage du serveur NFS », c'est un problème assez courant chez nous.

7- Intervention divine

Bizarrement, tout cet édifice aux proportions bibliques fonctionne la plupart du temps. Néanmoins, quand une rupture générale de la Force arrive, on peut se retrouver avec beaucoup de problèmes, tous différents. Voici quelques spécimens :

— **Ping-pong** — Deux routeurs se renvoient des paquets, chacun ayant décidé que Le Bon Chemin passe par l'autre. Cela se diagnostique avec la commande Unix `traceroute`. Solution : maudire Renater (le fournisseur d'accès Internet de l'ÉNS, grand spécialiste du tennis de table) et attendre que les routeurs oublient d'être stupides.

— **Panne des DNS** — Les DNS forment une structure hiérarchique qui a tendance à s'écrouler d'un coup. Dans ces conditions, tout semble figé, sauf si on utilise directement les adresses numériques ; mais il est difficile de retenir par cœur 4 milliards d'adresses... Le diagnostic : la commande `host` sur un hôte connu (`www.microsoft.com` par exemple) bloque et ne répond pas.

— **Plantage du serveur NFS** — Votre station est bloquée, des messages « `NFS server not responding, still trying` » apparaissent. C'est typique du serveur NFS (celui qui possède, physiquement, les disques partagés) qui a pris des vacances. Il faut attendre le reboot du serveur. NFS étant bien fait, cela débloquera les machines sans conséquence funeste pour les opérations en cours.

— **Plantage du serveur NIS** — NIS, *alias* les Pages Jaunes, est le protocole permettant de partager les mots de passe entre les stations. Pour l'utilisateur final, les conséquences sont semblables à un plantage NFS : il faut attendre le retour du serveur NIS. Une fois qu'il est revenu, tout remarche, rien n'a été perdu.

— **Embouteillage** — Une partie du réseau est à genoux. La plupart des paquets s'égarent, tout va lentement, c'est horrible. Pas de solution, sinon espérer que de nouvelles lignes seront bientôt mises en place, pester contre ces milliers de crétins qui ne pensent qu'à récupérer les dernières photos de Pamela Anderson, et revenir se connecter à une heure plus adaptée (pour les États-Unis, la bonne heure est 6 heures du matin : les Américains sont couchés, les Européens pas encore levés).

Il existe beaucoup d'autres pannes possibles, plus ésotériques. Elles sont la cause de certaines de ces discussions d'informaticiens, que personne ne comprend mais qui semblent les amuser follement.

[7] Pour en savoir plus sur les images bitmap ou vectorielles : `http://www.eleves.ens.fr:8080/home/radenne/Docs/images/Images.ps`.
[8] `http://www.eleves.ens.fr:8080/home/corbinea/modfiles.html`.

En savoir plus

Le cours d'Architecture et Systèmes comprend une partie animée par Jacques Beigbeder, qui décrit, entre autres, le fonctionnement du réseau.

Les gens motivés se référeront aux RFC (Request For Comments). Ce sont les documents de référence ; on les trouve là : `ftp://ftp.inria.fr/rfc/`

Le protocole IP est décrit dans la RFC 791.

▷ Thomas Pornin

Utiliser Netscape

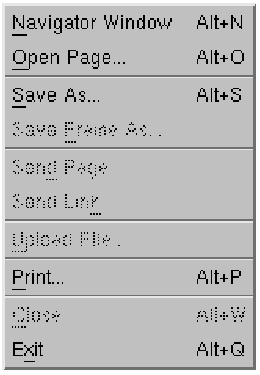
Le mois dernier, on a vu comment créer des marques-pages. Ce mois-ci, pour clore cette présentation des fonctionnalités principales de Netscape, on va voir comment télécharger des documents. Jusqu'ici, j'ai utilisé la version 3 de Netscape, mais par défaut nous sommes passés à la version 4. J'indique donc en parallèle les deux versions.

— **Principe** — Télécharger se dit *to download* en anglais, et quand on veut faire pro, on dit « downloader un machin ». Ça veut dire récupérer sur sa machine le fichier que l'on est en train de consulter, et qui se trouve physiquement sur un disque dur à l'autre bout du monde ou presque.

La page que vous consultez correspond en effet à un ou plusieurs fichiers (texte, images...). Quand vous vous connectez sur un site, vous demandez au serveur de vous envoyer telle page, c'est-à-dire tels ou tels fichiers. L'ordinateur s'exécute et envoie les données qui sont stockées dans un cache* sur votre machine. Quand on demande à sauvegarder les données, elles sont écrites sur le disque dur, chez vous.

— **Commment faire** — Dans le menu `File` de Netscape, vous trouvez une entrée `Save As` (ou `Alt-S` au clavier). Il suffit donc de sélectionner `Save As` pour récupérer le fichier.

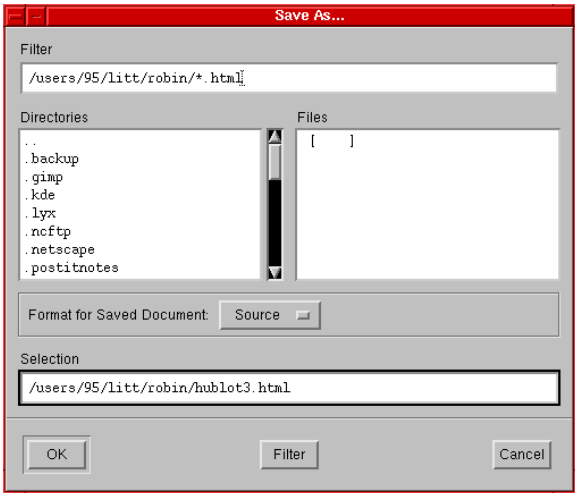
Netscape 3



Netscape 4

Netscape 4 fournit une entrée supplémentaire pour récupérer le contenu d'un cadre quand on se trouve sur une page avec *frames*. Pour cela, cliquez sur le cadre à télécharger, puis sur `Save Frame As`.

Ensuite, une fenêtre apparaît, qui ressemble à celle-ci :



Il y a deux choses importantes dans cette fenêtre. La première, c'est le format sous lequel on veut sauvegarder le document. Vous voyez ici « `Source` ». Ça veut dire qu'on va récupérer le fichier source, au format HTML, c'est-à-dire encodé pour être une page Web. Ce fichier au suffixe `.html` pourra être regardé avec un navigateur comme Netscape ou Lynx.

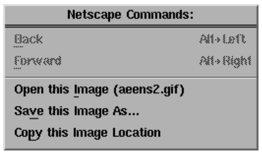
Vous pouvez aussi choisir « `Text` » ou « `PostScript` ». Au format texte, vous ne récupérez que le texte, formaté en paragraphes, sans encodage, comme un courrier électronique. Ce sera un fichier au suffixe `.txt`. Vous pouvez modifier à votre guise le HTML et le format texte.

Le format PostScript est le format de fichier spécifique aux imprimantes. Vous obtenez un fichier `.ps`, que vous pouvez regarder avec `gv` et imprimer avec `lpr`. Vous ne pouvez pas le modifier (enfin si, mais il faut le vouloir pour modifier du PostScript à la main !).

Dans le cadre `Selection` s'affiche le chemin d'accès du fichier que vous allez télécharger. Par défaut, le fichier est placé dans votre répertoire d'accueil. Vous pouvez modifier cette ligne pour mettre le fichier dans un sous-répertoire ; par exemple, pour mettre ce fichier dans mon sous-répertoire `tuteurs/`, j'écrirais dans le cadre : `/users/95/litt/robin/tuteurs/hublot3.html`

— **Télécharger des images** — Quand une page Web s'affiche, elle est en fait constituée de plusieurs fichiers : le fichier HTML qui la code, et les fichiers des images qui ont éventuellement été insérées. À moins de télécharger le fichier au format PostScript, les images ne sont pas récupérées avec la page quand vous la sauvegardez chez vous.

Il est cependant possible de récupérer des images. Quand vous cliquez avec le bouton droit sur une image, ces menus apparaissent (ils peuvent être plus étoffés si l'image supporte un lien) :



Netscape 3

Netscape 4

Il suffit de sélectionner « `Save (this) Image (as)` » pour récupérer l'image elle-même, sans le texte.

— **Déontologie** — C'est tout facile de récupérer plein de choses... Et le principe est le même pour les sons... Donc la correction et la modération s'imposent, comme le rappelle la charte signée en début d'année :

> Tout utilisateur s'engage à utiliser correctement les ressources mises à sa disposition : mémoire à ne pas saturer, espace disque, bande passante des réseaux, imprimantes, etc. (Règlement d'utilisation des moyens informatiques de l'École normale supérieure, 1er septembre 1999, article 5.1)

Comme on l'a vu dans le dernier numéro du *Hublot*, la place disque est partagée entre les utilisateurs. Si vous tenez à posséder des séries de photos intitulées « Pamela » ou « Samantha », ça vous regarde, mais ça n'a rien à faire sur votre compte.

Il est aussi plus simple de créer un marque-page vers une page intéressante, plutôt que de récupérer son contenu chez vous, dans la mesure où à l'École vous disposez d'une connexion permanente.

D'autre part, il y a des questions de copyright. Avant de télécharger quelque chose, vérifiez la nature du document. Certains cours ou articles sur le site d'universités sont soumis à copyright, l'image des personnages de bandes dessinées aussi. Le cas échéant, demandez l'autorisation à l'auteur du site avant de récupérer le résultat de son travail, même pour un usage privé. Nous ne parlerons pas des fichiers de musique piratée... Relisez les articles 6.1 à 6.5 de la Charte.

Lexique

— **Cache** — Le cache est un espace de stockage où le navigateur garde une copie des pages récemment visitees, afin d'accélérer leur chargement lors de la consultation suivante (il suffit alors de vérifier que la page n'a pas changé plutôt que de la télécharger de nouveau).

Il y a en fait deux types de cache : d'une part un cache en mémoire vive, qui ne conserve les données que durant une session de navigation, et d'autre part un cache disque (parfois désactivé pour économiser de la place) qui est conservé d'une session à l'autre.

— **Exécutable** — Il existe deux sortes de langages de programmation : les langages compilés et les langages interprétés. Les langages compilés fonctionnent selon le principe suivant : le programmeur écrit le code du programme, puis on passe par une étape de *compilation*, c'est-à-dire de traduction du code en langage machine. On dispose alors d'une part du code source du programme, que l'on peut modifier et recompiler, d'autre part d'un exécutable, illisible par un humain, mais compréhensible pour l'ordinateur.

Quand vous achetez un logiciel dans le commerce, vous achetez un exécutable produit à partir d'un code source, dont vous ne disposez pas. Vous pouvez utiliser le logiciel, le recopier, mais pas le modifier. Ce n'est pas vrai pour les logiciels libres.

— **Libre** — Un logiciel libre est un logiciel que l'on peut copier, modifier et distribuer librement. Le modifier implique de pouvoir disposer du code source du logiciel, et pas seulement de l'exécutable*. On peut revendre commercialement un logiciel libre ; dans ce cas, c'est un service que l'on vend (pressage du CD-ROM, documentation, service après-vente...), et non pas le logiciel lui-même. Un logiciel libre n'est donc pas obligatoirement gratuit.

Là-dessus se greffe la notion de *copyleft* (par opposition à copyright) : en plus d'être libre, le logiciel est sous une licence (telle que la GPL) qui garantit que les versions dérivées seront également libres. Des exemples de logiciels libres : pine, mutt, LaTeX, netscape, linux, gimp, xkobo...

— **Script** — Programme servant à automatiser des tâches répétitives, généralement simple. Principales caractéristiques : le langage utilisé interprète directement le script, de manière transparente, sans avoir besoin de passer par un traducteur externe ; les constructions utilisées sont généralement de haut niveau par rapport à un langage de programmation traditionnel, et collent plus avec la notion de « faire telle tâche, puis telle autre ».

— **Texte (mode texte)** — Un programme lancé dans un terminal peut être en mode texte, auquel cas il s'affiche dans la fenêtre où il a été lancé. Pour sa mise en page, il utilise le jeu de caractères disponibles, c'est-à-dire les caractères alphanumériques et les ponctuations. Il peut être en couleurs ou en noir et blanc. Certains logiciels en mode texte peuvent s'utiliser avec la souris, possèdent des menus déroulants, etc. Ils s'utilisent cependant le plus souvent avec le clavier. Pine est un logiciel de courrier en mode texte.

Par opposition, un logiciel en mode graphique se lance dans une nouvelle fenêtre, s'utilise avec la souris, et propose divers moyens graphiques (fontes, couleurs, icônes) pour son utilisation, comme Netscape.

L'Atelier Internet : bilan de l'année 1999–2000

Cette année, les membres de l'Atelier Internet se sont avant tout consacrés à l'organisation du colloque « comprendre les usages d'Internet », qui s'est tenu à l'École les 3 et 4 décembre 1999. Cette manifestation a réuni quelques trois cents experts des sciences humaines et exactes, ainsi que des responsables d'entreprises, du gouvernement, pour faire le point sur les « usages ». Il en est ressorti, outre un grand nombre de collaborations et de programmes de recherches, le désir de prolonger nos travaux suivant trois axes :

* L'incidence de l'informatique et d'Internet sur nos pratiques intellectuelles ; nous intégrons les deux aspects, matériel et conceptuel de cet environnement (le support comme la pratique d'écriture) ;
* La compréhension des logiques économiques : la « bulle spéculative » n'est peut-être pas si creuse qu'on le prétend, et l'opposition logiciel libre*/commerciel (comme le disent les francophones d'Amérique du Nord) est moins franche qu'on ne l'imagine. En revanche, la place croissante des savoirs dans l'espace marchand pose la question des modalités de l'enseignement dans les années à venir (donné ou vendu, public ou privé, à distance et/ou en « présentiel ») ;
* Développer une batterie d'outils adaptés à l'écriture, la navigation sur le Web et à l'analyse des pratiques. Tout reste à faire dans ce domaine, le seul frein à la production est l'absence d'imagination.

La convergence pressentie il y a deux ans entre sciences humaines et sciences exactes nous apparaît aujourd'hui évidente après avoir rencontré à Toronto des personnes comme :

— Mark Olsen, fondateur de l'ARTFL Project à Chicago, grand passionné de Linux et responsable de la publication en ligne des 5 premiers dictionnaires de l'académie ;

— Marie-Luce Demonet, éditrice du corpus de Rabelais, dont les outils permettent de reconnaître le sens d'un mot, quelle que soit sa graphie (pouvant distinguer la préposition « à » de la forme conjuguée du verbe avoir, quand l'accent de la préposition est absent) et qui risque de devenir « leader » de l'OCR[9] si elle continue à vouloir lire des ouvrages trop fragiles pour être posés à plat sur un scan...

– ou Barry Russel, ancien lecteur à l'École, expert ès-théâtre français des XVII^e^ et XVIII^e^ siècles, qui jongle avec PHP3 et JavaScript, et qui a de fabuleuses idées pour suivre (et accompagner) la démarche intellectuelle d'un lecteur de pages Web.

Tous sont persuadés que les sciences exactes ont développé des méthodes et des outils, aujourd'hui suffisamment affinés pour permettre une étude intelligente de ce qui était encore hier exclu de leur champ d'investigation : le langage, l'écriture, la pensée.

Bon, nous n'allons pas racheter M$ demain ! Mais dans quelques mois, nous lancerons des recherches communes avec des élèves de plusieurs Grandes Écoles (les normaliens sont bien sûr toujours bienvenus) pour développer tout d'abord des systèmes d'aide à la navigation sur Internet (visualisation de l'architecture de sites Web visités), des (petits) moteurs de recherche, des mesures de la proximité sémantique entre serveurs, tout en prolongeant la cartographie automatique du Web (lectorat de serveurs, localisation de tous les serveurs d'une zone donnée, comme le RIPE), et des outils d'analyse syntaxique (questions posées à un indexeur, analyse de pages Web).

En effet, nous voyons Internet comme un territoire, dont nous ne pouvons prendre la mesure que si nous l'investissons pleinement : tout comme l'idée d'un anthropologue qui ne fait pas de terrain se conçoit difficilement (ainsi que celle de sa neutralité en milieu autochtone), nous ne concevons pas une analyse des pratiques sociales et intellectuelles d'Internet sans nous investir en tant qu'auteurs, éditeurs et lecteurs, bref en tant qu'acteurs à part entière, dans ce nouveau territoire.

En attendant, il nous reste à publier les actes du colloque de décembre, ce qui n'est pas une mince affaire.

Mais nous sommes impressionnés par la façon dont notre réseau national et international se renforce cette année 2000, par le biais d'échanges solides avec des acteurs majeurs de l'Internet ; ces personnalités hors du commun naviguent sans encombre entre diverses cultures : économistes, linguistes ou philosophes, ils manipulent avec aisance les giga-octets et ont une activité éditoriale de premier plan. Ils redonnent aux termes « information » et « base de données » leurs lettres de noblesse, puisque l'une et l'autre sont des recueils de Racine, des cahiers de Valéry, voire des bibliothèques entières, et non pas des théories « sémiologiquement correctes » qui souvent élèvent au rang de connaissance des séries télévisuelles ou des publicités crétinisantes.

Que dire de plus ? Que les amoureux et amoureuses des langues vivantes, soucieux d'intégrer dans nos dictionnaires des mots comme php3, Java, perl et Kmel, les passionnés qui voudraient construire eux-mêmes un Littré en ligne, les aventuriers à la recherche de terrains intellectuels quasi-jamais foulés, tous peuvent venir nous rejoindre à l'Atelier Internet dès la rentrée prochaine.

▷ Éric Guichard

[9] OCR : *Optical Character Recognition.* NDLR.

Des pingouins partout

Le pingouin contre-attaque

Entre un NASDAQ hoquetant et une presse spécialisée dithyrambique à défaut d'être sérieuse, le monde du pingouin est en ébullition. En effet, l'Empire Maléfique de Redmond tremble sur ses bases, le terrible Bill Gates est poussé dans ses derniers retranchements par les chevaliers anti-trust, la bataille finale approche, et elle décidera du sort du monde libre. On n'aurait pas cru, vu d'ici, que les quelques démêlés d'un délinquant fortuné eussent pu impliquer autant. Présentons ici les quelques factions qui se disputent le fameux marché des systèmes d'exploitation, cela permettra de bien illustrer ce que c'est au juste que ce fameux logiciel libre.

— **Aux origines de Microsoft** — Rendons à César ce qu'il a dûment OPAté, commençons par Microsoft. Cette société, ancienne startup des années 70, alors nommée Traf-o-Data et spécialisée dans la numération des véhicules individuels, prend son patronyme actuel en 1978. Parmi la dizaine d'associés chevelus qui composent l'intégralité du personnel de cette société naissante, un jeune homme probablement brillant, Bill Gates. En ces temps-là, IBM est bien embêtée : Apple commence à bouffer du marché avec ses nouveaux ordinateurs (Apple I et Apple II, pas encore le MacIntosh, faut pas rêver, j'évoque des âges reculés). Les finauds de chez IBM conçoivent un plan machiavélique : construire une machine pourrie, la lancer sur le marché pour bien jeter la confusion, et le client hésitant se tournera tout naturellement vers la tranquille pérennité des ordinateurs chers mais éprouvés d'International Business Machines. D'où le premier PC. IBM refusera le 68 000, processeur avancé de Motorola, et choisira Intel, qui fournit le 8086, chimère monstrueuse, union contre-nature du passé et du futur, un peu comme le serait un TGV à vapeur. Pour le système d'exploitation, pas question de choisir un dérivé d'Unix (trop professionnel) ou de CP/M (trop évolutif) ; Billou arrive alors avec sous le bras un code d'un pote à lui. Bill a acheté ce code 1 000 dollars, il le revend pour un million (déjà on voit qu'il est doué). Et surtout, il ne vend qu'une licence non exclusive, et il se garde le droit de faire de nouvelles versions et de les vendre lui-même, comme un grand. Ce code, c'est Ms-Dos.

C'est qu'il est visionnaire, l'ami Gates ! Contrairement aux demeurés d'IBM, il a bien vu ce qui allait se passer : le PC, boosté par son indigence totale en matière de brevets, est copié, concurrentialisé, guerre-desprisé, sur-diffusé, et enfin conquiert le marché. Et on en est là : 95% des ordinateurs sont des PC. Et IBM s'est fait enfoncer par Compaq sur le marché des PC, par Sun sur le marché des stations de travail. Il y a au moins un peu de justice dans cette affaire. Pendant ce temps, William Gates, III^{e} du nom, a vendu diverses versions de Ms-Dos, et du complément graphique de ce dernier, le bien connu Windows, qui a connu moult opus, les derniers étant Windows 98 et Windows 2000.

Alors là, maintenant, suivez bien : Bill, c'est le méchant. Il profite odieusement de sa position monopolistique pour vendre plein de copies de son Windows, qu'il a truffé de bugs volontairement pour vendre plein de nouvelles versions, il a fait pression sur les revendeurs[10] pour les plier à ses buts diaboliques de contrôle du monde, il s'est allié avec la mafia, il soutient le terrorisme mondial, il a vendu son âme au Malin et il mange les petits enfants. Bon, ok, il est possible que je sacrifie un peu de ma légendaire objectivité afin de donner un peu d'emphase lyrique au personnage. Mais quand on s'aventure chez le pingouin, il faut s'y faire, c'est le style de la maison. Quand on arrive chez la Comtesse, on lui présente ses hommages ; quand on est invité dans l'antre du volatile, on crache sur Billou. C'est l'usage.

— **Aux origines de Linux** — Voyons donc son challenger actuel. En 1991, un étudiant finlandais, Linus Torvalds, s'achète un PC. Il en est très content. Mais il aimerait bien savoir comment ça marche. Comme c'est un étudiant, il a du temps libre ; alors il se met à programmer un nouveau système d'exploitation, juste pour voir. Magie d'un Internet balbutiant, d'autres oisifs se joignent à ses efforts, et, inconséquents qu'ils sont, décident de se servir du nouveau joujou pour faire du vrai travail. En quelques mois, Linux était né (même s'il a failli s'appeler Phreax, en fait). Le pingouin lui-même sera désigné comme mascotte en 1996, sous l'impulsion du nordique créateur sus-désigné, qui n'a jamais fourni de raisons sérieuses pour ce choix pour le moins étrange. Probablement un traumatisme dans sa petite enfance.

Où en est ce Linux, maintenant ? Quelque part entre 2 et 10 millions d'exemplaires en usage dans la nature. À la grosse louche, 3% du marché mondial des systèmes d'exploitation sur micro-ordinateurs. C'est peu, face aux fenêtres bi-millénaires. Mais suffisant pour remuer des capitaux dans un marché un peu enthousiaste, voire même franchement téméraire, pour tout dire absolument cinglé et digne de la camisole. Il faut dire que Linux, c'est de l'informatique, donc c'est aussi une religion, et donc il y a des fanatiques. Ce sont les Linuxiens. Le Linuxien, il vit pour et par Linux. Il installe Linux partout où il peut, non pas parce que ça pourrait aider quelqu'un à bosser (et pourtant, le logiciel du pingouin, il marche plutôt bien, souvent mieux que les productions de Bill), mais parce que c'est Linux et que Linux doit être partout, c'est ainsi. Le vrai Linuxien est à la fois attendrissant et donne une envie irrésistible de taper dessus, c'est un Bisounours hi-tech. Le Linuxien hait par dessus tout toute forme de production venant

[10] Tout ceci est vrai, et justifie le procès actuel fait à Microsoft. NDLR.

de chez Microsoft, et également tout quidam qui pourrait évoquer, ne serait-ce qu'un instant, l'idée d'utiliser Windows. Comme quoi rien ne change au cours des siècles, on est toujours le païen de quelqu'un.

Quand pingouin fâché, lui toujours faire ainsi

— **GNU** — La secte Linux reconnaît une autorité philosophique supérieure : il s'agit de GNU. GNU, c'est le mouvement de libération du logiciel. Attention, c'est très sérieux : le vrai GNUiste peut mordre s'il estime qu'on attaque son dogme. Le gourou de cette église, c'est Richard M. Stallman ; il ressemble à un hippie qui aurait découvert les bienfaits du courant électrique. C'est lui qui a inventé le logiciel libre*, dont Linux est un parfait exemple et Windows pas du tout (parce que Billou fait du logiciel propriétaire* ; bouh le méchant Billou). L'idée est simplissime, lumineuse même : un logiciel, ça coûte très cher à produire et rien du tout à reproduire, alors on va le donner, ce logiciel, et se faire du fric sur le support technique. Mais, puisque RMS était investi d'une mission divine, il ne suffisait pas de donner le logiciel, il fallait encore convertir les égarés. Tel le virus de la variole, la *General Public License* arrive à la rescousse : remarquable mécano de juriste, cette licence d'utilisation d'un logiciel libre stipule qu'une fois qu'on a mis le doigt dedans, on se fait arracher le bras avec. En gros, si on intègre un morceau de logiciel GPL dans un autre logiciel, le tout devient GPL.

Les cieux s'ouvrirent, la terre trembla, le logiciel libre était né. C'est dingue comme une idée peut être à la fois aussi américaine (distribution gratuite, pour bien s'emparer du marché) et aussi socialiste (les moyens de production doivent être gratuits). RMS créa son temple, la *Free Software Fundation*, afin de promouvoir le but grandiose de la libération planétaire, que dis-je, galactique, des hackers.

Le projet de RMS, à l'origine (en 1983), c'était de refaire le monde. Pour un informatheux, le monde, c'est l'informatique, bien sûr. Donc, RMS avait constaté qu'il n'y avait pas de système d'exploitation à sa convenance. Alors il fallait qu'il en fasse un qui serait mieux que tous les autres. Il a donc commencé à écrire plein d'outils pratiques, notamment ce qu'il faut pour développer un système d'exploitation[11]. Ces outils ont été promptement adoptés et absorbés par le pingouin, et maintenant, le prototype du système GNU (dénommé Hurd) est en train de rejoindre le placard, comme tant d'autres projets géniaux avant lui.

Un hérétique, Eric S. Raymond, a fondé son propre culte réformiste. Lui, il utilise le terme « Open-Source », qui veut dire à peu près autant de chose que « Free Software », c'est-à-dire tout et rien. ESR est plus fourbe que RMS, et, d'ailleurs, ses idées plaisent plus que celles de RMS dans le milieu des entreprises (bien que ce soient les mêmes idées, au fond ; comme quoi, tout est affaire de marketing).

— **BSD** — Pour parfaire ce tableau somptueux, évoquons BSD : dérivé du travail d'une équipe de l'université de Berkeley, le code BSD se pose comme une sorte de concurrent à la fois de GNU et de Linux, même si des échanges ont lieu dans les deux sens. Bien évidemment, les BSDistes ne se privent pas pour dire que leur code à eux il est bien plus libre que le code GNU. Les GNU-men rétorquent qu'eux défendent bien plus la Cause que BSD. Querelle de clocher, finalement. Le code BSD n'a pas le côté viral de la GPL, sinon c'est pareil.

Les systèmes de BSD (car il y en a plusieurs, il ne faudrait quand même pas imaginer que les schismes n'existent pas) sont souvent décrits comme étant beaucoup plus professionnels et performants que ceux du pingouin finlandais. Ce qui est objectivement avéré, c'est qu'ils sont beaucoup plus austères. C'est vrai que ça fait vraiment professionnel, parce que ça n'a vraiment aucune chance de conquérir le grand public, un système d'exploitation qui glorifie le mode texte*.

Pétrole ! Pétrole !

Quelle est la finalité de tout cela ? La justice ? L'équité ? Le bonheur des utilisateurs ? Non, bien sûr, rien de tout cela. La charité, ça ne remue plus les foules. Le Bien, le Mal, c'est très suranné. La vraie raison de toute cette histoire, le moteur, le nerf, c'est le pognon. Le marché du libre est juteux, dégoulinant même, c'est une terre de miel et de lait, et est arrosé de fleuves dans lesquels un quelconque Midas a dû plonger les mains.

Et pourquoi cet engouement ? L'humanité n'a-t-elle pas vécu plusieurs millénaires sans logiciel libre (car sans logiciel du tout), et pas si mal que ça quand on regarde de près ? Et bien, ce n'est pas important. Que l'ordinateur soit indispensable ou superflu, là n'est pas la question. La seule Vérité, c'est que tout le monde en veut un, et puis c'est tout. Robert Gropagu, agriculteur dans la Creuse, sait désormais que sans informatique ses navets ne pousseront jamais, aussi il s'y met, et il consomme. Et comme il se pique d'être un grand défenseur des libertés individuelles, il fait dans le logiciel libre, parce que même quand on est paysan on a le droit de faire smart.

— **Valeur, valeurs et valeur** — Le logiciel libre est le symbole de l'esprit révolutionnaire retrouvé, la Liberté montant à l'assaut des remparts obscurantistes du logiciel propriétaire, tendance bourgeois. Alors ça fait bien, de défendre de grands idéaux de formats libres, de logiciels de qualité développés pour le simple plaisir de les voir tourner, et tout ça dans une grand messe financière. Car *la Liberté guidant le Peuple*, c'était il y a peu sur tous les bons billets de 100 francs. On réconcilie l'humanitaire et le mercantile, c'est presque l'incarnation du bonheur prolétaire communiste. Lénine en serait tout ému.

[11] Exemples d'outils, au sens large... : `emacs` (l'éditeur de textes), `gnuchess` (le jeu d'échecs), `gimp` (logiciel de dessin et de retouche d'images)...

On aurait tort de prendre le marché des « nouvelles technologies » comme preuve de l'ampleur du phénomène, parce que là-bas ils sont tous fous. Mais il est vrai que tout ce qui est estampillé « logiciel libre » connaît une inflation big-banguesque, tellement que la start-up moyenne a pour unique destin et but d'être introduite en bourse, moment magique où elle rembourse moult fois l'investissement en quelques jours. En fait, ladite start-up est souvent vide de sens, et la vente d'actions est en soi une arnaque ; citons par exemple VA-Linux, produit de spéculation pure puisque la société fondée pour l'occasion n'a pas l'ombre d'un début de commencement d'activité. Il n'empêche qu'elle a multiplié par huit son cours le premier jour (mais le dit cours est retombé plus bas que terre depuis, une sorte d'intrusion de la réalité dans cette belle histoire).

Certes, une légère « correction », comme ils disent, a eu lieu, et un léger coup de balai subséquent. Et Microsoft a des soucis et risque le démantèlement. Mais il ne faut pas se leurrer : le fantasme de la conquête de l'Ouest, la ruée vers l'or, existe encore, et on va continuer à voir fleurir des affiches dans le métro, vantant des sites web avec beaucoup trop de k et de oo dans leur nom. D'ailleurs, pour en revenir sur le cas de Microsoft : c'est le délire chez les Linuxiens, en ce moment. Ça y est, la bête est morte ! Bien évidemment, ils oublient que ça ne va rien changer, Windows continuera d'être partout, le format des documents Word ne sera pas plus spécifié, et puis, même si Word est finalement porté sous Linux (hypothèse très probable, en fait : Linux, c'est un marché aussi important que MacIntosh, maintenant), ça n'empêchera pas Word de planter. Il est très significatif, dans ces conditions, que le démantèlement soit prononcé non pas pour cause de fumisterie flagrante des programmeurs de chez Microsoft, mais en vertu de la loi anti-trust, qui, en l'occurrence, préfère avoir deux entreprises faisant du code propriétaire qu'une seule.

Alors, cette folie du logiciel psychotrope, à qui la faute ? J'en vois déjà qui cherchent des coupables. Cessez votre quête, la réponse est évidente : les coupables, c'est vous. Oui, vous, lecteurs qui pensez être aussi purs que l'agneau qui vient de naître et dont les défenses immunitaires ont déjà envoyé *ad Patrem* des milliers de microbes innocents. Car c'est vous qui vous émerveillez devant tous les gadgets qui clignotent, les révolutions des informations autoroutières, les arbres de Noël virtuels. *Panem and circenses*, comme toujours. Et c'est pour vous contenter, vous donner ce divertissement que vous réclamez, que les banquiers allongent 20 plaques au premier glandu qui arrive avec une idée en bois (mais c'est une idée « Internet »). Quand on voit un premier ministre faire des discours sur un thème, de toutes façons, il faut bien se dire que le thème en question est démagogiquement porteur

— **Alors, quel avenir ?** — C'est là qu'on rigole : les utilisateurs vont souffrir. Et pourquoi ? Mais c'est qu'ils n'intéressent personne ! Oui, je sais, toutes les pubs vous disent que l'informatique, c'est pensé pour l'utilisateur, pour qu'il puisse s'instruire, s'informer et bosser. Et bien c'est du pipeau ! Et des deux côtés encore bien.

D'abord, du côté des vendeurs : ce qu'ils cherchent, c'est de la chair à canon. Le grand public, ah, la belle affaire. Non, tous les linuxiens vous le diront (sans honte et sans fard, c'est assez étonnant pour que ce soit noté), la cible, c'est le décideur. Le décideur, c'est le gros bourgeois cigarifiant qui, dans la boîte, décide ce qui sera acheté et ce qui ne le sera pas. Et lui, le brave homme, il n'utilise pas d'ordinateur (il n'est pas fou). Le logiciel libre, peu lui chaut. Pour ravir son cœur, il faut lui en mettre plein la vue ; pour cela, des gadgets (lui aussi, il aime ça), et surtout des troupes.

C'est l'effet pervers du pingouin : se répandre tel le virus, demander à être installé partout, recruter, entraîner, former des hordes de linuxiens à envoyer dans la bataille. Qu'importe si Linux n'est pas plus adapté aux besoins (même fictifs) des utilisateurs que Windows. Non, la Cause est tout ce qui compte. Le-logiciel-libre, à-mort-le-code-propriétaire. Un mantra vidé de son sens et scandé selon la grande tradition du martellement idéologique. D'aucuns auraient invoqué Krishna ; mais Linux est un meilleur dieu de nos jours. Chez Microsoft, ils font pareil, de toutes façons, alors les Linuxiens (et assimilés) se sentent pardonnés par avance.

Maintenant, les clients : certes, ils se font estamper, mais ils l'ont bien cherché. Après tout, les pubs leur disent bien qu'un gros ordinateur, ça sert à s'instruire, à créer et à travailler. Personne ne prétend jamais que ça sert à jouer ! Et, franchement, à quoi ça rime de pleurer sang et eau à cause du prix d'une licence Windows, quand on allonge sans rechigner 400 balles pour chaque version de *Tomb Raider* ? Hypocrisie générale. Il n'y a pas d'autre mot.

Les pigeons de l'affaire, ce sont les pauvres gars qui veulent simplement travailler sur ce damné ordinateur qui n'est même pas le leur. Ils regrettent la simplicité du temps où on utilisait des machines à écrire, et où on ne multipliait pas les impressions, faute de photocopieuse (qui restait encore à inventer).

Mais ceux là sont condamnés, parce que l'époque ne leur est pas favorable. Faut dire, ah là là, y'a plus de saison, ma bonne dame.

▷ Thomas Pornin