

Les logiciels libres pour l'enseignement et la recherche

Jean M. Thiéry

<http://ModLibre.org>

Un bon enseignement doit transmettre des connaissances qui peuvent être expliquées et validées. L'enseignement de l'informatique ne peut donc se faire qu'avec des logiciels ouverts, dont on peut étudier tous les détails. Il en est de même pour tous les enseignements faisant appel à des bases de connaissances informatisées : celles-ci devraient être suffisamment ouvertes pour pouvoir être validées.

Pour qu'un nouveau résultat scientifique soit accepté, il faut qu'il puisse être reproduit par des laboratoires indépendants. Toutes les étapes de la recherche doivent donc être décrites avec le maximum de précision, en particulier au niveau des logiciels utilisés. Les logiciels ouverts sont alors particulièrement adaptés.

Les logiciels peuvent être diffusés avec des statuts très différents. Ils peuvent être *ouverts* ou *fermés*, *modifiables* ou *figés*, *gratuits* ou *payants*, *copiables* ou *non*, etc. La *Fondation pour le Logiciel Libre* (*Free Software Foundation* ou *FSF* : <http://www.fsf.org/>) a créé la licence *GNU GPL* (*GNU General Public License* : <http://www.fsf.org/licenses/licenses/>) qui est la licence libre la plus répandue. Cette licence a introduit 4 critères complémentaires pour vérifier si un logiciel est *libre* : liberté d'exécuter le programme pour tous les usages, liberté d'étudier le fonctionnement du programme, liberté de redistribuer des copies et liberté d'améliorer le programme et de publier les améliorations (http://www.aldil.org/projets/fiches_libres/). Un logiciel *ouvert* n'est pas toujours *libre* selon ces 4 critères.

- La liberté d'utilisation d'un programme *libre* ne peut être limitée que par la loi, qui doit décrire précisément les usages illicites.
- La liberté d'étudier le fonctionnement des programmes est indispensable pour l'enseignement de l'informatique et de toutes les technologies qui utilisent l'informatique. Et a fortiori pour les recherches dans ces disciplines.
- La liberté de distribution permet aux enseignants de transmettre à leurs élèves les logiciels indispensables pour leurs études.
- La liberté d'amélioration permet de créer des versions du logiciel adaptées à chaque application, en utilisant, par exemple, la langue ou les termes techniques appropriés.

De nombreux sites ont été créés pour faire connaître les logiciels libres ou ouverts disponibles. Le plus connu est probablement celui qui est géré par la *FSF* pour l'*UNESCO* (<http://directory.fsf.org/>). Les rubriques "*Education*", "*Internet applications*", "*Mathematics*", "*Science*" et "*Web authoring*" concernent directement les *JED'07*. On peut aussi trouver de nombreux liens similaires sur le site francophone de *Framasoft* (<http://www.framasoft.net/>) (en particulier dans les rubriques "*Education*", "*Internet*", "*Sciences*" et "*Site Web*").

Nous allons explorer différents domaines scientifiques en indiquant les logiciels libres les plus souvent utilisés, ainsi que quelques logiciels ouverts ou propriétaires qui n'ont pas encore d'équivalents libres. Cette démarche reste subjective dans la mesure où il y a peu de statistiques sur l'utilisation réelle des logiciels.

Il existe des outils libres à tous les niveaux des **recherches en informatique**.

- Les algorithmes fondamentaux sont libres, tout au moins en droit européen. L'ambiguïté existe pour certains algorithmes intégrés dans des composants matériels qui peuvent être brevetés. Il y a une forte pression juridique pour étendre les domaines d'application de cette brevetabilité (http://fr.wikipedia.org/wiki/Brevet_logiciel).
- Les langages de programmation sont libres pour la plupart. D'autres sont contrôlés par des sociétés commerciales. *SUN* va libérer *Java* en le distribuant sous la licence *GPL* (<http://www.april.org/association/positions/java-gpl.html>).
- L'un des premiers efforts de la *FSF* a été de développer le compilateur *GCC* (*GNU Compiler Collection*) entièrement libre pour permettre la maîtrise complète des développements informatiques. Ce compilateur possède des compléments pour les langages *C*, *C++*, *Objective-C*, *Fortran*, *Java* et *Ada*, avec toutes les bibliothèques nécessaires (<http://gcc.gnu.org/>).
- Tous ces outils ont permis de développer des systèmes d'exploitation complets comme *GNU/Linux* (<http://fr.wikipedia.org/wiki/Linux>).

De nombreux logiciels libres ont été développés pour les **mathématiques appliquées**.

- Les premières recherches ont porté sur les bibliothèques de calculs matriciels qui ont convergé vers la bibliothèque *LAPACK* (<http://en.wikipedia.org/wiki/LAPACK>, <http://www.cs.colorado.edu/~jessup/lapack/>).
- La bibliothèque *GSL* (*GNU Scientific Library*) fournit de nombreux sous-programmes pour la génération de nombres aléatoires, le calcul des fonctions spéciales et les ajustements par les moindres-carrés (<http://www.gnu.org/software/gsl/>).
- Des langages spécifiques ont été développés pour faciliter les calculs numériques : les calculs matriciels intensifs sont pilotés par un langage interprété de haut niveau souvent complété par des outils graphiques de simulation. On peut citer :
 - le système propriétaire *Matlab - Simulink* (<http://www.mathworks.com/>),
 - le logiciel libre *Octave* (<http://www.gnu.org/software/octave/>), équivalent de *Matlab*,
 - et le système ouvert *Scilab - Scicos* (<http://www.scilab.org/>) analogue à *Matlab - Simulink*.

On peut aussi noter les logiciels propriétaires *Maple* (<http://www.maplesoft.com/>), *Mathcad* (<http://www.mathcad.com/library/>) et *Mathematica* (<http://www.wolfram.com/>) qui permettent en plus des calculs symboliques.

- Malgré leurs différences de conception, d'ergonomie et d'efficacité, ces outils ont des domaines d'application qui se chevauchent largement. En général, un chercheur (ou un laboratoire) a tendance à tout calculer avec le même outil pour des raisons de formation et de compatibilité des fichiers de données. Plusieurs développeurs de logiciels de calculs numériques ont créé un groupe de travail commun (*the Numerical Mathematics Consortium* : <http://www.nmconsortium.org/>) pour standardiser les spécifications de ces logiciels.

Les calculs numériques sont surtout utilisés en ingénierie ou en mathématiques appliquées à d'autres sciences. Les **mathématiques fondamentales** font plus appel à des calculs symboliques.

- Les recherches en calcul symbolique ont commencé avec des langages comme *Lisp* dont une version libre est incorporée dans l'éditeur *Emacs* (<http://www.gnu.org/software/emacs/>).
- Les recherches faisant appel aux calculs symboliques sont très variées. Il n'y a donc pas eu, au niveau de ces logiciels, la convergence que l'on peut noter pour le calcul numérique. Parmi les logiciels libres de calculs symboliques on peut citer :
 - *Axiom*, un système polyvalent pour les calculs symboliques et les recherches mathématiques ... (<http://wiki.axiom-developer.org/FrontPage>).
 - *GAP*, un système de calcul pour les mathématiques discrètes, avec un accent sur la théorie des groupes, ... (<http://www.gap-system.org/>).
 - *Maxima*, un système de manipulation d'expressions symboliques et numériques (<http://maxima.sourceforge.net/>).
 - *PARI/GP*, un système de calcul formel très répandu, conçu pour des calculs rapides en arithmétique ... (<http://pari.math.u-bordeaux.fr/>).
 - *Sage* (*Software for Algebra and Geometry Experimentation*), un système écrit en *Python*, permettant l'intégration de nombreux logiciels mathématiques libres (<http://sage.sourceforge.net/>, voir le lien vers Wikipédia).
- On peut aussi noter le logiciel *MuPAD* (<http://www.mupad.de/>) qui est partiellement ouvert.

Dans le domaine des **statistiques**, la plupart des logiciels sont des logiciels propriétaires. Il existe cependant deux exceptions remarquables :

- Le langage *R* est une version libre du langage *S*. Il bénéficie de très nombreuses bibliothèques statistiques et permet de tracer d'excellents graphiques directement publiables (<http://www.r-project.org/>).
- *GRASS* est un Système d'Information Géographique pour l'analyse et la gestion des données spatiales. Il permet l'analyse d'images, la production de cartes et de graphiques, la modélisation spatiale et la visualisation (<http://grass.itc.it/>).

Dès leur apparition, les ordinateurs ont été utilisés pour des calculs de **physique**. Depuis des décennies, de nombreux logiciels ont été développés dans des centres internationaux comme le *CERN* (Organisation européenne pour la recherche nucléaire) où est né le web (<http://public.web.cern.ch/public/Welcome-fr.html>). Plusieurs logiciels pour la physique sont décrits sur ce site (<http://wwwasd.web.cern.ch/wwwasd/cernsoftware.html>).

Dans les autres domaines de la physique, les logiciels disponibles peuvent être retrouvés avec les moteurs de recherche usuels et des mots-clé appropriés : voir, par exemple, le site (http://www.mol.biol.ethz.ch/groups/wuthrich_group/software), spécialisé en résonance magnétique nucléaire.

L'importance des logiciels pour la **chimie théorique** a été reconnue dès 1962, date de création du *QCPE* (*Quantum Chemistry Program Exchange* : <http://qcpe.chem.indiana.edu/>). Cet organisme permet, depuis des décennies, la diffusion de logiciels ouverts pour un coût raisonnable couvrant les frais de gestion (envoi de bandes magnétiques ou de disquettes avant l'arrivée d'Internet !).

Avant les années 1980, la plupart des logiciels de calculs moléculaires étaient ouverts. Les meilleurs d'entre eux ont été intégrés ultérieurement dans des produits propriétaires. De nouveaux logiciels libres ont dû être développés pour la recherche. L'un des plus connus est *PyMOL*, un système de graphiques moléculaires intégrant un interpréteur *Python* (<http://pymol.sourceforge.net/>).

En **bioinformatique** (au niveau de la biologie moléculaire et des génomes) on trouve un site spécialisé pour les logiciels ouverts : l'*Open Bioinformatics Foundation* (<http://www.open-bio.org/>).

Les logiciels libres ne sont pas réservés aux sciences ! Par exemple les linguistes utilisent souvent le système de composition de documents *LaTeX* (<http://fr.wikipedia.org/wiki/LaTeX>) quand ils doivent combiner des langues écrites dans des alphabets différents.

Dans tous les cas, il ne faut pas hésiter à utiliser les moteurs de recherches usuels ! Les sites de synthèse sont généralement bien référencés.

Le mouvement des logiciels libres se transpose dans tous les domaines culturels avec le développement d'une *culture libre* accessible à tous, en particulier grâce à Internet.

L'exemple le plus souvent cité est l'encyclopédie *Wikipédia* (<http://fr.wikipedia.org/wiki/Accueil>) et ses nombreux projets comme *Wikibooks* et *Wikiversity*. Cette encyclopédie multilingue peut être utilisée pour l'enseignement de toutes les disciplines. En général, il n'y a pas de problème de crédibilité pour les articles purement techniques. Pour des sujets plus sensibles (culture, histoire, etc.), il est important d'exercer son esprit critique en utilisant d'autres sources d'information, éventuellement dans d'autres langues.

De même, de nouvelles licences permettent aux auteurs, qui le souhaitent, de diffuser leurs oeuvres librement sur Internet tout en gardant le contrôle de cette diffusion. Par exemple, les licences *creative commons* (<http://fr.creativecommons.org/>) offrent plusieurs contrats types (commercialisation *autorisée* ou *non*, modifications *autorisées*, *interdites* ou *acceptées sous la même licence*). Certaines revues scientifiques comme *BMC Evolutionary Biology* ont déjà choisi ce type de licence. Cette évolution est conforme à la Déclaration de Berlin sur le Libre Accès à la Connaissance en Sciences exactes, Sciences de la vie, Sciences humaines et sociales (http://www.inist.fr/openaccess/imprimersans.php3?id_article=38).

Le site francophone de *Framasoft* (<http://www.framasoft.net/>) permet de retrouver rapidement la plupart des ressources libres pour l'enseignement. D'autres sites de référence sont indiqués sur <http://ModLibre.org>. Là aussi, les moteurs de recherche usuels peuvent être utilisés avec succès.