

Formats ouverts, ça veut dire quoi ?

Stéphane Bortzmeyer

<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 22 Février 2006. Dernière mise à jour le 10 Juin 2009

<http://www.bortzmeyer.org/formats-ouverts.html>

Maintenant que le logiciel libre est entré dans les mœurs, il est temps de se demander quelle est sa relation avec les **formats ouverts**. Qu'est-ce qu'un format ouvert ?

S'il existe plusieurs définitions du logiciel libre (Debian <http://www.debian.org/social_contract#guidelines>, FSF <<http://www.gnu.org/philosophy/free-sw.html>> ou bien OSD <<http://www.opensource.org/docs/definition.php>>), toutes donnent, en pratique, le même résultat. Il est exceptionnel qu'une licence soit considérée comme libre par une de ces définitions et pas par les autres. Donc, on peut dire que « logiciel libre » est un concept clair aujourd'hui. Même ses adversaires les plus bouchés ne confondent plus, par exemple, libre et gratuit.

Il en est tout autrement pour **format ouvert** (le format d'un fichier est un cas particulier d'une norme ; je les utilise ici de manière interchangeable). Si tous les partisans du logiciel libre, et même d'autres personnes, approuvent l'idée de formats ouverts, ils n'en ont pas pour autant de définition claire. Par exemple, un excellent article dans Groklaw, "*Is ODF an Open Standard?*" <<http://www.groklaw.net/article.php?story=20060209093903413>> examine plusieurs définitions possibles, sans en adopter aucune. (L'article est surtout consacré à un long et intéressant exposé sur la création d'Open Document.)

Pourtant, c'est une question essentielle : beaucoup d'organismes publics et d'administrations, par exemple, stockent toutes leurs données exclusivement dans des formats fermés, au mépris de la pérennité de ces documents, et de leur accessibilité par les citoyens qui auraient choisi le logiciel libre.

Avant de regarder les définitions existantes, voyons quelques principes : pour moi, il est clair que XML est un format ouvert et que MS Word ne l'est pas. Mais en vertu de quoi ? Les critères qu'on peut employer sont :

- La possibilité d'accéder à la spécification du format.
- La possibilité légale (via la licence) et pratique (via la documentation du format) de mettre en œuvre ce format, surtout dans un logiciel libre.
- La possibilité de contribuer à son évolution. En informatique, une norme figée a moins d'intérêt.
- La disponibilité d'une mise en œuvre de ce format dans un logiciel libre.

Le premier critère, l'accession à la norme, n'est pas rempli par certaines organisations pourtant censées être sans but lucratif. C'est ainsi que l'ISO fait payer pour ses spécifications (qui ne peuvent donc certainement pas être considérées comme ouvertes) et interdit la redistribution (un enseignant en informatique est donc dans l'illégalité s'il photocopie une norme ISO pour demander à ses étudiants de l'implémenter).

Paradoxalement, il est plus facile d'obtenir la spécification du format Microsoft Word que d'accéder à une norme ISO. Mais, le deuxième critère, la possibilité d'implémentation, n'est évidemment pas rempli lorsque le format est secret ou lorsqu'il est plombé par des brevets, dont la nocivité, surtout pour le logiciel, n'est plus à démontrer. On voit souvent dans le jargon des organisations de normalisation les termes de RF (*"royalty-free"*) et de RAND (*"Reasonable And Non-Discriminatory"*). Le premier désigne un brevet dont la licence d'exploitation est gratuite (en argent : il peut avoir d'autres conditions). Le deuxième est moins net, c'est un brevet dont la licence d'exploitation est accordée à tous sur des critères « raisonnables ». Certaines organisations de normalisation n'intègrent de techniques brevetées que s'il y a une licence RAND, d'autres exigent une RF et certaines laissent le choix.

Le troisième critère, la possibilité de faire évoluer la norme est également très importante. Beaucoup de formats par ailleurs plutôt ouverts (comme le format de fichier PDF, le langage de programmation Java ou bien le format de syndication RSS) sont contrôlés par un acteur privé, qui ne laisse pas aux autres le soin de décider des évolutions futures. Celui qui contrôle la norme peut donc faire peser sur ses concurrents la menace d'une évolution subite, qui les rendrait incompatibles.

Le quatrième critère, une mise en œuvre en logiciel libre, est d'un maniement plus délicat. La plupart des organisations de normalisation ne cherchent pas à développer du logiciel. Une norme peut être très ouverte et ne jamais avoir suscité l'intérêt d'un développeur.

Voyons maintenant les propositions de définition existantes.

Le fameux *"Comparative Assessment of Open Documents Formats. Market Overview."* <<http://europa.eu/idabc/en/document/3439/5585#ODF>> de l'Union Européenne donne une définition minimale *"that the document format is completely described in publicly accessible documents, that this description may be distributed freely and that the document format may be implemented in programs without restrictions, royalty-free, and with no legal bindings."*, ce qui correspond à mes deux premiers critères. Une définition ultérieure, publiée dans *"European Interoperability Framework for pan-European eGovernment Services"* <<http://europa.eu/idabc/en/document/3761>> est plus précise en ajoutant *"The standard is adopted and will be maintained by a not-for-profit organisation, and its ongoing development occurs on the basis of an open decision-making procedure available to all interested parties (consensus or majority decision etc.)"*, mon troisième critère.

Le très connu Bruce Perens, auteur déjà d'une définition du logiciel libre utilisée par Debian, a écrit une *"Open Standards Principles and Practice"* <<http://www.perens.com/OpenStandards/Definition.html>>. Ce texte mélange des critères objectifs et mesurables (*"Open Standards are available for all to read and implement"*) et des souhaits qui ne peuvent pas se traduire de manière précise (*"Open Standards create a fair, competitive market for implementations of the standard"*). Le processus de développement du format n'est pas présenté comme un critère de choix. Par contre, Perens ajoute un critère que je n'ai pas cité : la nécessité de permettre des sur-ensembles ou des sous-ensembles de la norme, ce qui peut faire hésiter. (Un langage de programmation comme Ada prend bien soin de ne pas les autoriser, au nom de la portabilité des programmes.)

Ken Krechmer a publié *"The Meaning of Open Standards"* <<http://www.csrstds.com/openstds.html>>. Contrairement à Perens, il mentionne mon troisième critère, les conditions dans lesquelles le format est développé et maintenu.

Georg Greve a proposé cinq points <http://www.fsfe.org/en/fellows/greve/freedom_bits/an_emerging_understanding_of_open_standards>, qui ressemblent beaucoup aux miens mais il met très fortement l'accent sur l'importance d'avoir une implémentation en logiciel libre, considérant la norme comme surtout théorique, autrement. Ces cinq points sont discutés au sein du Forum pour la Gouvernance de l'Internet via le groupe DCOS <<http://igf-dcos.org/>>.

On voit déjà qu'il n'y a pas de consensus entre ces deux définitions. Mais il y en a d'autres :

- Celle du Wikipédia anglophone <http://en.wikipedia.org/wiki/Open_standard>. Elle note à juste titre qu'il n'existe pas de consensus. Elle ne mentionne pas les conditions dans lesquelles la norme est développée. Et elle cite des normes comme SQL qu'il est difficile et coûteux d'obtenir (elle n'est pas disponible en ligne).
- Celle du Wikipédia francophone <http://fr.wikipedia.org/wiki/Format_ouvert>. Elle ne mentionne pas non plus les conditions de développement de la norme.
- "*Open standards, open source, open minds, open opportunities*" <http://www.ibm.com/developerworks/blogs/page/BobSutor?entry=open_standards_vs_open_source1> est surtout un bon résumé de la notion de norme et explique les organismes de normalisation.
- <<http://www.openformats.org/>> se penche plutôt sur des questions pratiques (jusqu'à discuter de la qualité de la compression de JPEG par rapport à celle de PNG).
- La LCEN parle de formats ouverts : <<http://formats-ouverts.org/blog/2004/07/01/12-UnArticleDeLoiDefinitCeQueSontLesFormatsOuverts>>

À noter une des rares études comparatives sur les SDO, menée à la demande des autorités danoises : "*Evaluation of Ten Standard Setting Organizations with Regard to Open Standards*" <<http://www.itst.dk/arkitektur-og-standarder/Standardisering/Aabnestandarder/baggrundsrapporter/Evaluation%20of%20Ten%20Standard%20Setting%20Organizations.pdf>>.

Bref, encore du pain sur la planche pour débroussailler cette question. En attendant, lisez l'excellent blog dédié, <<http://formats-ouverts.org/>>.