

Nicolas Pettiaux
Prof De Sc
Association électronique libre
Avenue du pérou 29
B-1000 Bruxelles
Nicolas.pettiaux@ael.be
Portable : 0496 24 55 01

Monsieur Christophe Van Vaerenbergh
Président
NBN - Bureau de normalisation
Rue de la brabançonne 29
B-1000 Bruxelles

Bruxelles, le 25 juin 2007

Commentaires à NBN - Normalisatie Bureau – Bureau de Normalisation à propos de la demande de standardisation ISO du format ECMA 376 (JTC1-Sc34)

Par l'association électronique libre (AEL)

Introduction

La spécification « ECMA 376 Office Open XML », dénommée « ECMA 376 » dans la suite de ce message, a été soumise à l'ISO en vue de son approbation en tant que standard international. Le NBN, tout comme les instances officielles de normalisation de certains autres pays, est appelée à prendre avant le 2 septembre 2007 une position favorable ou non à la mise en oeuvre de cette approbation.

De très nombreux commentaires permettant de mettre en avant un doute sérieux sur le bien fondé de la reconnaissance par ISO du format ECMA 376 ont été rassemblés par les utilisateurs, tant privés que industriels. Plusieurs sont repris ci-dessous car nous paraissent significatifs d'un défaut qui empêchent de considérer que le format ECMA 376 peut aujourd'hui être considéré comme adéquat pour devenir un standard international.

Nous pensons cependant aussi que le premier exemple repris dans la section [Démonstration du point concernant l'impossibilité de définir des dates antérieures à 1900 dans le format ECMA 376 pour les tableurs](#) suffit à lui seul à mettre en évidence la pertinence de nos propos. Nous le détaillons donc.

Nous demandons aux membres du Comité national belge de voter « **NON** » à la résolution ISO DIS 29500 (format Open OfficeXML ou OOXML) notamment pour les raisons exposées ici.

L'AEL considère que la procédure actuelle, suite de la procédure « fast-track », n'est pas justifiée, compte tenu des risques et des incertitudes qui grèvent lourdement le standard proposé et qui sont incompatibles avec la mission des instances nationales et internationales de normalisation.

La première difficulté est évidente car elle découle du volume du dossier. Le standard ECMA 376, en effet, est décrit par un document de 6.000 pages environ. Or ce document n'a été finalisé et validé que le 7 décembre 2006 par l'ECMA. Les échéances qui nous était imposée du 5 février 2007 puis du 2 septembre 2007, pour une spécification de cette taille portant sur une matière aussi complexe, rendent matériellement impossible toute analyse sérieuse. Un tel rapport entre le nombre de pages de la spécification et le délai imparti est sans précédent connu dans l'histoire de la standardisation. L'AEL est d'autant plus préoccupée que, dès son origine, le projet Office Open XML se développe sous le signe d'une précipitation qui ne semble pas compatible avec la bonne tenue d'une démarche de standardisation. L'AEL a donc travaillé en collaboration avec des associations et autres structures pour établir ses remarques.

En effet, paradoxalement et bien que cette spécification représente, en volume, un record absolu, elle a été mise au point et validée au sein de l'ECMA plus rapidement que tout autre standard dans ce

domaine.

Dans les circonstances actuelles du marché, il convient de s'interroger sur les motifs d'une telle précipitation, et de veiller aux risques d'instrumentalisation des instances de normalisation à des fins non conformes aux objectifs réels de la standardisation. L'AEL, compte tenu du rôle essentiel qu'a joué la société Microsoft dans la genèse de ce nouveau standard, se doit de vous alerter sur la coïncidence entre l'agenda du projet ECMA 376 et celui du lancement commercial du logiciel bureautique Microsoft Office 2007.

Le lancement d'une procédure accélérée auprès de l'ISO est d'autant plus préoccupant que le projet ECMA 376 comporte des éléments inquiétants pour l'avenir des standards dans un secteur important des technologies de l'information, ainsi que pour le crédit des organismes de normalisation. Plusieurs caractéristiques essentielles du dossier (dont il n'est pas possible de dresser la liste exhaustive dans ce message) sont en effet contraires à l'esprit (et parfois aux règles) de la standardisation. Nous devons attirer votre attention sur les plus immédiatement visibles d'entre elles.

Non respect des critères pour l'évaluation des normes

Qu'est-ce qu'une norme ? Plusieurs définitions pertinentes sont disponibles. L'ISO dit :

"[Un] document, établi par consensus, et approuvé par une organisation reconnue, qui fournit, pour un usage courant et répété, des règles, des directives ou des caractéristiques pour les activités ou leurs résultats, visant à atteindre le meilleur degré d'ordre dans un contexte donné.

NOTE : les normes devraient appuyer sur les résultats éprouvés de la science, de la technologie et de l'expérience, et viser le bénéfice maximal pour la communauté."¹

Les standards britanniques de la BSI sont:

"... une norme est une façon adoptée et reproductible de faire quelque chose. C'est un document publié qui contient une spécification technique ou d'autres critères précis destinés à être utilisés de façon cohérente comme règle, directive ou définition. Les normes aident à se simplifier la vie et à augmenter la fiabilité et l'efficacité des biens et des services utilisés. Ils sont censés être *incitatifs* – un condensé des meilleures pratiques plutôt que de la pratique commune. Les normes sont créées en rassemblant les expériences et l'expertise de toutes les parties intéressées, telles que les fabricants, les revendeurs, les utilisateurs et les régulateurs d'un matériel, produit, procédé ou service particulier."²

Les directives ISO/IEC JTC1 mentionnent:

"L'un des buts de la normalisation en Technologies de l'Information est d'assurer que les produits disponibles sur le marché possèdent les caractéristiques d'interopérabilité, de portabilité et d'adaptabilité culturelle et linguistique. Pour cela, les normes développées devraient refléter les pré requis des Caractéristiques Stratégiques Communes suivantes :

- Interopérabilité ;
- Portabilité ;
- Adaptabilité culturelle et linguistique."³

¹ ISO/IEC Guide 2:2004, définition 3.2. Plusieurs organismes de normalisation nationaux ont aussi adopté cette définition ISO, p. ex la DIN allemande.

²<http://www.bsi-global.com/en/Standards-and-Publications/About-standards/What-is-a-standard/>

³JTC1 Directives, 5th Edition, Version 3.0, Section 1.2

D'après ce qui précède et d'après d'autres définitions nationales, quelques thèmes communs émergent pour la caractérisation des normes :

1. Elles définissent des critères communs précis pour faire quelque chose de façon reproductible.
2. Elles fournissent un niveau optimal d'ordre dans un contexte donné, censé être incitatif, donnant les résultats scientifiques, technologiques et expérimentaux éprouvés, un résumé des meilleures pratiques plutôt qu'un état des pratiques courantes.
3. Elles encouragent l'interopérabilité et la portabilité.
4. Elles s'adaptent aux différentes langues et cultures.

Ce document évalue la proposition DIS 29500 "Office Open XML" (OOXML) vis-à-vis de chacun de ces critères. Certains exemples spécifiques de problèmes avec la spécification OOXML sont mentionnés, mais ce ne sont qu'une poignée d'exemples extraits d'une liste de plusieurs centaines. Le nombre élevé de problèmes sérieux posés par OOXML démontre son immaturité comme spécification et sa non-pertinence dans le cadre d'une procédure accélérée d'adoption (*Fast Track approval*) comme norme ISO.

Redondance globale dans le domaine des documents bureautiques

Le périmètre du projet ECMA 376 recoupe très précisément celui du standard ISO/IEC 26300:2006, dit OpenDocument (ODF). Selon certaines affirmations (non étayées par des éléments techniques précis), ECMA 376 aurait pour vocation de combler certaines lacunes du standard ODF. Cette argument est dénué de fondement sérieux, dans la mesure où ECMA 376, loin de réutiliser ISO/IEC 26300 pour le compléter, recouvre l'ensemble du domaine, et se pose donc comme une spécification concurrente.

L'adoption de deux standards incompatibles régissant le même domaine technologique n'est conforme ni aux usages ni au simple bon sens industriel. L'objectif du standard actuel est l'interopérabilité entre les différents dispositifs de production et de traitement de documents bureautiques ; l'existence d'un second standard serait de nature à compromettre gravement l'atteinte de cet objectif et à maintenir un cloisonnement artificiel du marché, créant de ce fait un "obstacle technique" au sens défini par l'Organisation Mondiale du Commerce.

En outre, le choix de la dénomination "Office Open XML" est de nature à susciter la confusion et l'incertitude auprès du public, sachant que le format ISO/IEC 26300 est lui-même dénommé "Open Document Format for Office Applications" et qu'il existe un logiciel bureautique libre nommé OpenOffice.org.

Conflit avec d'autres standards

Le projet ECMA 376 est une enveloppe qui, sous couvert de s'appliquer aux caractéristiques des documents bureautiques, englobe un ensemble hétérogène de spécifications particulières qui entrent en concurrence directe avec des standards déjà établis. Parmi les conflits les plus visibles, on peut citer

- un nouveau mode de représentation des dates, incompatible avec la spécification ISO 8601 ;
- une codification spécifique des langues du monde, sans rapport avec le code ISO 639 ;
- des modalités de représentation des images en conflit avec le format "Scalable Vector Graphics" (SVG) du W3C et avec le format ISO/IEC 8632 dit "Computer Graphics Metafile" (CGM) ;
- un nouveau format de représentation XML des formules mathématiques, en conflit avec la spécification MathML du W3C ;
- des algorithmes de chiffrement spécifiques, ne correspondant ni à la spécification ISO/IEC 10118-3 dite Whirlpool, ni au XML-ENC du W3C, ni à aucun des systèmes cryptographiques recommandés par les organisations gouvernementales ou professionnelles

concernées par les standards de sécurité ;

- une représentation des séquences animées qui redéfinit le domaine des présentations multimédia synchronisées en conflit avec le standard "Synchronized Multimedia Integration Language" (SMIL) du W3C.

Ces exemples démontrent que le projet ECMA 376 est susceptible de jouer un rôle déstabilisant dans de nombreux domaines, et qu'il peut être considéré comme un cheval de Troie tendant à faire adopter en bloc un vaste ensemble de spécifications qui devraient logiquement faire l'objet de décisions séparées mais dont plusieurs, si elles étaient examinées individuellement, ne pourraient qu'être rejetées. Il convient de noter à cet égard que le volume important de ce document s'explique en grande partie par la redéfinition de standards existants, et non par des apports nouveaux à l'effort de standardisation.

Incomplétude et obstacles à la mise en oeuvre

Certains éléments de la spécification ECMA 376 invoquent explicitement des fonctionnalités obligatoires qui ne sont pas décrites dans le document et qui ne correspondent ni à des standards établis ni même à des publications techniques librement accessibles. Des fonctionnalités sont prescrites par référence à certains comportements de logiciels appartenant à la société Microsoft ou à d'autres sociétés, sans autre précision.

On peut citer, entre autres exemples, certains mots-clés de la spécification tels que `autoSpaceLikeWord95`, `lineWrapLikeWord6`, `useWord97LineBreakRules` et bien d'autres, dont l'objet est de signaler que toute application conforme au standard doit, dans certaines situations, imiter le comportement de certains logiciels commerciaux (les exemples cités ici invoquent directement les logiciels Word 95, Word 6 et Word 97 de Microsoft, au lieu d'intégrer la description du comportement recherché dans la spécification publique). De même, l'application de certains aspects de la spécification en matière de représentation d'images requiert le support des "métafichiers Windows", c'est-à-dire du format Windows Metafile (WMF), dont la définition n'est pas publique.

La spécification peut, dans certains cas, avoir pour effet d'imposer au marché la dégradation consciente d'une fonctionnalité. Ainsi, elle oblige à considérer l'année 1900 comme une année bissextile pour la représentation des dates, de manière à reproduire un défaut connu du tableur Excel de Microsoft (ce qui, en l'occurrence, revient à donner valeur de standard à une déficience technique).

Or la pratique consistant à citer le comportement d'un produit commercial spécifique et à imposer le respect d'éléments de spécification non publiés est de toute évidence contraire aux règles en vigueur dans les organismes de normalisation. Cette pratique constitue en outre un obstacle technique à la mise en oeuvre des spécifications par des acteurs autres que la société Microsoft.

Avant tout processus d'admission comme standard international, le contenu d'ECMA 376 devrait, au minimum, avoir été intégralement spécifié.

Qualité technique du dossier

La qualité technique de la spécification ECMA 376 n'est pas au niveau habituellement requis pour un standard. Cette spécification est en partie alimentée par une compilation de règles découlant des comportements (nécessairement hétérogènes) de logiciels Microsoft anciens et récents.

Ainsi, la représentation des graphiques vectoriels est non seulement incompatible avec le standard SVG, mais encore fait l'objet de deux spécifications divergentes (dénommées DrawingML et VML), ce qui semble vouloir refléter des divergences de comportement entre certains logiciels Microsoft.

Les rédacteurs ne semblent avoir défini et appliqué aucune règle commune de nommage et de codification ; chaque chapitre du projet semble redéfinir son propre vocabulaire. Le manque de principes généraux de nommage est aggravé par des choix discutables dans le choix des noms des éléments et des attributs, ce qui contrevient aux recommandations de lisibilité des schémas XML, tels qu'elles sont prescrites par le W3C et par la société Microsoft elle-même. Ce vocabulaire inutilement ésotérique nuit gravement à la lisibilité de la spécification.

Le choix des unités de mesure relatives aux dimensions et à la disposition des objets est incohérent et change d'un sous-ensemble à un autre de la spécification. Par exemple, une valeur

relative de 87,5% est représenté tantôt par "pct87", tantôt par "87500", tantôt par "4375", selon l'objet auquel il s'applique. L'absence de logique commune sur ce point semble provenir de divergences historiques entre logiciels de la société Microsoft.

En résumé, nos justifications principales du « non » sont

1. Il y a **déjà un standard ISO 26300 appelé Open Document Format (ODF)**: un double standard engendre des coûts supplémentaires pour l'industrie, les gouvernements et les citoyens;
2. Il n'y a pas de **réelle implémentation de la spécification OOXML** : Microsoft Office 2007 produit une version spéciale d' OOXML, pas un format de fichier respectant la spécification OOXML;
3. Il **manque des informations dans la description du format OOXML**, par exemple, comment faire un **AutoSpaceLikeWord95** ou **useWord97LineBreakRules**;
4. Plus de **10% des exemples mentionnés dans le standard ne sont pas conformes** avec XML; (voir <http://surguy.net/articles/ooxml-validation-and-technical-review.xml> une analyse faite pour The British Standard Institute dans le cadre d'une étude semblable)
5. Il n'y a **aucune garantie que quiconque puisse développer un logiciel** implémentant la spécification OOXML **sans enfreindre une licence ou sans être redevable de frais de licence**;
6. Ce **standard est en conflit avec d'autres standards ISO**, comme ISO 8601 (Representation de Dates et Heures), ISO 639 (Codes pour la représentation des Noms et Langues), ou encore ISO:iec 10118-3 (hash cryptographique);
7. **La définition des dates dans le format de fichier du tableur interdit d'y entrer toute date antérieure à 1900**; ceci n'est apparemment pas une erreur de définition et affecte la spécification OOXML et des logiciels tels que Microsoft Excel 2000, XP, 2003 ou 2007.
8. Ce **standard n'a pas été créé en confrontant l'expérience et l'expertise des toutes les parties concernées** (comme les éditeurs de logiciels, les revendeurs, les acheteurs, les utilisateurs et les autorités de régulation), mais par Microsoft seul comme en témoigne le cahier des charges définissant la mission du Technical Committee de ECMA en charge de définir ECMA 376.
9. Non respect des critères habituels de définition des normes internationales

Compte tenu de tous ces éléments, l'AEL considère le projet ECMA 376 Office Open XML non seulement comme une source potentielle de redondances, de conflits et d'insécurité en matière de standardisation, mais encore comme un développement immature en termes de génie logiciel.

En outre, de par son caractère extraordinairement complexe et foisonnant, ECMA 376 ne peut pas, de toute évidence, être convenablement évalué et validé dans les délais et les conditions d'une procédure accélérée.

Enfin, l'adoption d'un second standard applicable aux documents bureautiques amoindrirait gravement les opportunités de convergence du marché offertes par le standard ISO/IEC 26300:2006 "OpenDocument".

Nous demandons aux membres du Comité national belge de voter « NON » à la résolution ISO DIS 29500 (format Open OfficeXML ou OOXML) notamment pour les raisons exposées ici.

Je reste à votre disposition pour toutes informations complémentaires et je vous prie d'agréer, Monsieur, l'expression de mes meilleurs sentiments.

Nicolas Pettiaux, Prof. Dr. Sc.

Annexe : démonstration du point concernant l'impossibilité de définir des dates antérieures à 1900 dans le format ECMA 376 pour les tableurs

La définition des dates dans le format de fichier du tableur interdit d'y entrer toute date antérieure à 1900; ceci n'est apparemment pas une erreur de définition et affecte la spécification OOXML et des logiciels tels que Microsoft Excel 2000, XP, 2003 ou 2007. Nous détaillons ici un élément repris en anglais dans <http://surguy.net/articles/ooxml-validation-and-technical-review.xml> avec encore plus de précisions que celles que nous fournissons.

Nous avons installé la version de démonstration de Microsoft Excel 2007, dont le format de fichier par défaut est, selon les propos de Microsoft même, le format ECMA 376. Cette version est gratuitement téléchargeable sur le site <http://www.microsoft.com> moyennant enregistrement de l'utilisateur, obtention d'un numéro de code d'utilisation et ensuite d'une validation d'installation en ligne.

	Colonne B	Colonne C	Colonne D
Ligne 2	12/06/2007		-3 B1-B4
Ligne 3	13/06/2007		
Ligne 4	14/06/2007		
Ligne 5	15/06/2007		
Ligne 6			
Ligne 7	12/06/1907		-3 B6-B9
Ligne 8	13/06/1907		
Ligne 9	14/06/1907		
Ligne 10	15/06/1907		
Ligne 11			
Ligne 12	27/02/1800		0 B11-B14
Ligne 13	27/02/1801		
Ligne 14	27/02/1802		
Ligne 15	27/02/1803		
Ligne 16			
Ligne 17	27/02/1800		-3 B16-B19
Ligne 18	28/02/1800		
Ligne 19	01/03/1800		
Ligne 20	02/03/1800		

Illustration 1: Tableau des valeurs (document introduit ici avec OpenOffice.org Calc)

Nous avons entré dans le tableur les valeurs reprises ci-dessous :

L'analyse ci-dessous décrit ce qui apparaît dans la figure [Illustration_2](#), une impression d'écran de ce qui apparaît dans l'application Microsoft Excel 2007 durant les manoeuvres.

Une première valeur est introduite dans B2 (le 12 juin 2007), puis est « tirée » avec la souris de B3 à B5, et le tableur automatiquement reconnaît la valeur comme étant une date, et calcule comme attendu la série B3 à B5 en créant les dates croissantes (13 au 15 juin 2007). Dans la cellule C2, la différence de dates est calculée entre les dates contenues dans les cellules B1 et B4. Cette valeur (-3) est la valeur attendue de la différence de jours entre les 2 dates.

La même opération est effectuée avec comme date de départ le 12 juin 1907 introduit dans la cellule B7. Celle-ci est également automatiquement et convenablement reconnue comme une date, et lors de l'opération d'extension de la série vers les cellules B8 à B10, la série est convenablement construite. La différence de dates en C7 entre les contenus des cellules B6 et B9 est faite également convenablement.

Enfin la date du 28 février 1800 est introduite dans la cellule B11. Celle-ci n'est pas reconnue comme date valable par l'application, et ne peut pas l'être même par une définition manuelle du format des données de la cellule (sélection et clic droit)

	A	B	C	D	E	F
1		12/06/2007		-3 B1-B4		
2		13/06/2007				
3		14/06/2007				
4		15/06/2007				
5						
6		12/06/1907		-3 B6-B9		
7		13/06/1907				
8		14/06/1907				
9		15/06/1907				
10						
11		27/02/1800	#VALEUR!	B11-B14		
12		27/02/1801				
13		27/02/1802				
14		27/02/1803				

Illustration 2: Test de dates avec Microsoft Excel 2007

La différence des cellules B11-B14 comme montré dans [Illustration 2](#) ne peut pas être calculée. Un examen détaillé montre que les données des cellules B11 à B14 sont considérées comme des données textuelles. Ceci ne peut pas être changé, même manuellement.

L'illustration [Illustration 3](#) obtenue avec le tableur Calc de la suite OpenOffice.org v 2.2.1 dans laquelle le fichier converti au format Excel 97-2003 fut ouvert montre que les dates sont toutes bien lues. Dans ce document, de nouvelles dates sont définies dans les cellules B16 à B19, toujours selon la même méthode (B16 est entré au clavier, reconnu automatiquement comme date par le logiciel puis étendu en série avec un « glisser-tirer ». La différence de dates dans C16 est convenablement calculée.

Ce document est ensuite sauvé au format Excel 97-2003 pour être à nouveau ouvert par l'application Excel 2007. Ceci est présenté dans la figure [Illustration 4](#). Il y apparaît clairement que les données dates définies dans les cellules B16-B19 ne sont pas reconnues comme telles.

Une comparaison détaillée des graphiques cote à cote est présentée page suivante.

Les fichiers ayant servis à cette analyse sont bien entendus disponibles.

	A	B	C	D	E
1		12/06/2007		-3 B1-B4	
2		13/06/2007			
3		14/06/2007			
4		15/06/2007			
5		12/06/1907		-3 B6-B9	
6		13/06/1907			
7		14/06/1907			
8		15/06/1907			
9		27/02/1800		0 B11-B14	
10		27/02/1801			
11		27/02/1802			
12		27/02/1803			
13		27/02/1800		-3 B16-B19	
14		28/02/1800			
15		01/03/1800			
16		02/03/1800			

Illustration 3: Ouverture du tableur avec Calc de OpenOffice.org et ajout des dates en B16-B19

	A	B	C	D	E	F
1		12/06/2007		-3 B1-B4		
2		13/06/2007				
3		14/06/2007				
4		15/06/2007				
5		12/06/1907		-3 B6-B9		
6		13/06/1907				
7		14/06/1907				
8		15/06/1907				
9		27/02/1800		#VALEUR!	B11-B14	
10		27/02/1801				
11		27/02/1802				
12		27/02/1803				
13		#####		-3 B16-B19		
14		#####				
15		#####				
16		#####				
17		#####				
18		#####				
19		#####				

Illustration 4: Réouverture du nouveau tableau dans Microsoft Excel 2007