

Toulouse, 3 décembre 2008



FORMATS BUREAUTIQUES

Enjeux et dérives de la normalisation



Jean-Marie Gouarné

GENICORP

jean.marie.gouarne@genicorp.fr



« Les normes, en donnant une capacité d'agir sur le comportement des autres acteurs et à les conformer à un modèle défini, constituent un instrument de pouvoir. Au sein des organismes qui les élaborent, le choix des normes (sujet et contenu) constitue donc un enjeu d'influence et de puissance. Les membres des organisations de normalisation cherchent à imposer leur vision du monde et à conforter leur propre stratégie de conquête et de protection des marchés. »

« À Armes Égales »

Rapport au Premier Ministre par Bernard Carayon, 2006



« *formats bureautiques* » essai de non-définition

- la bureautique est un concept incertain et implique l'utilisation de formats hétérogènes impliquant des problématiques distinctes (*DOC, XLS, PPT, PDF, HTML...*)
 - la bureautique au sens usuel ignore certains formats documentaires professionnels
 - l'appellation « *formats de documents révisables* » (AFNOR 2007) ne couvre que les documents produits et modifiés par *certain types d'applications*
-
-



« *formats bureautiques* » périmètre actuel de la normalisation

- formats de documents **révisables** (FDR) dans un sens restrictif
- domaine arbitrairement défini à partir de trois types d'applications nommément désignées : **traitement de textes, tableur, présentation**

*un périmètre marqué par la bureautique
« grand public » de la fin du 20ème siècle
et ses limites*



« *formats bureautiques* » un territoire complexe et mal cerné

- un FDR est un format-enveloppe : un document bureautique combine des objets textuels, graphiques et sonores
 - il peut inclure des contenus dynamiques et des procédures de traitement
 - la couverture intégrale du contenu n'est pas envisageable dans une seule norme
 - la définition d'un périmètre à partir des trois seules applications classiques est insuffisante pour les cas d'utilisation récents et à venir (ex: *Web x.x*)
-
-



« *formats bureautiques* » un très ancien nouveau débat

- les premiers efforts de normalisation des FDR remontent aux années 1980 (cf. *ODA*)
 - ces efforts ont pris une ampleur sans précédent dans la décennie actuelle
 - le processus en cours est marqué par une controverse (sans issue prévisible) relative à la position dominante d'un fournisseur
 - l'opposition d'objectifs et d'intérêts inconciliables compromet la qualité technique et la validité à long terme des travaux
-
-



définition de la cible

- un consensus existe sur des objectifs abstraits de **pérennité et d'interopérabilité**
- les positions divergent quant au contenu pratique de ces notions
- le devenir de la base installée complique le débat



pérennité

**assurance de la possibilité future d'accès
au document au-delà de la disparition
des outils ayant servi à le créer**

mais encore...

- pérennité du contenu ?
- pérennité de la structure ?
- pérennité de la présentation ?



pérennité des limites congénitales

- la préservation intégrale de la *présentation* est hors de portée des FDR normalisés actuels
- inversement, la normalisation tend à se complexifier et à couvrir des aspects autres que la préservation des contenus statiques (ex : formules)





interopérabilité

- au sens industriel général, l'interopérabilité est la capacité d'un composant ou d'un système à fonctionner en interaction avec un autre composant ou système
- dans le débat sur les FDR, le mot est utilisé pour désigner la possibilité pour des utilisateurs dotés d'équipements bureautiques concurrents d'échanger des documents sans perte de contenu ou de présentation (=> **compatibilité**)



une perspective à recadrer

- **interopérabilité** (au sens strict) et **compatibilité** sont des conditions de la **pérennité**
- un malentendu persiste sur la notion d'interopérabilité
- la pérennité est au *second plan* dans les travaux actuels sur les FDR



rationalisation du passé ou normalisation pour l'avenir ?

- les FDR du passé n'ont pas été conçus pour la pérennité ; cet héritage pèse sur la normalisation
 - le « standard de fait » est en réalité le produit d'une accumulation de formats hétérogènes
 - la rétro-ingénierie intégrale de l'existant à des fins de normalisation ne serait pas économiquement justifiée
 - compte tenu des implications commerciales, les acteurs sont fondamentalement divisés sur le rapport entre anciens FDR et nouvelles normes
-
-



Histoire d'une course à la normalisation

ODF/OOXML : même technologie, visions divergentes

Limites de la normalisation dans le domaine des formats bureautiques



formats bureautiques

une histoire en 3 temps

- cloisonnement et hétérogénéité (1980 ...)
- le monopole en guise de standard (1990 ...)
- la mode des « standards ouverts » (2000 ...)





le prix de l'absence de norme

- la multiplicité des formats est impraticable dans un monde interconnecté et compromet la préservation à long terme
 - sans norme, l'effet de réseau mène au monopole
 - les « standards de fait » entretiennent des conditions de marché économiquement coûteuses et stratégiquement dangereuses ; ils ne garantissent aucune pérennité
-
-



normes et standards

conventions terminologiques

- *standard de fait* : spécification privée contrôlée par un acteur en position dominante
 - *standard* : spécification homologuée par une profession ou un organisme privé multipartite et dont l'autorité est (plus ou moins) reconnue : ECMA, OASIS, W3C, AIIM, etc
 - *norme* (standard de droit) : spécification homologuée par un organisme de normalisation sous contrôle étatique (ex: AFNOR) ou une fédération d'agences nationales (ex: ISO, CEI, UIT)
-
-



normes et standards cadre institutionnel actuel



JTC1/SC34



**Sous-comité commun de l'ISO et de la CEI
consacré aux formats de documents et aux
langages de traitement**

Correspondant français : AFNOR



normes et standards cadre institutionnel actuel

- la conduite de processus de normalisation est une prestation de service payante offerte par les agences de normalisation
- les normes résultent généralement d'un consensus entre les principaux « sponsors » industriels
- les entreprises utilisatrices souhaitent la normalisation mais en laissent le contrôle aux fournisseurs et en subissent les résultats





un précédent oublié : ODA

Open Document Architecture

- défini comme *format d'échange* de documents
 - standard ECMA publié dès 1985
 - encouragé par la Commission Européenne
 - norme internationale depuis 1989 (UIT T.410, ISO 8613)
 - homologué AFNOR : Norme Française
 - toujours en vigueur, jamais utilisé : première alerte sur l'efficacité de la normalisation des FDR
-
-



ODF – l'origine

le format OpenOffice.org

- issu d'un logiciel « propriétaire » puis libre mais...
- architecture logicielle facilitant le découplage entre application et format d'enregistrement
- un rôle de challenger logiquement intéressé par une large interopérabilité





ODF – les jalons (1)

- 2001 : publication du *format OpenOffice.org XML avant disponibilité du logiciel OpenOffice.org*
- 11-2002 : soumission au consortium OASIS ; **OpenOffice.org XML devient Open Office XML**
- 12-2004 : **Open Office XML devient Open Document Format for Office Applications (ODF)**
- 5-2005 : ODF 1.0 validé par l'OASIS



ODF – les jalons (2)

- 10-2005 : ODF 1.0 soumis au JTC1 (ISO/IEC)
 - 5-2006 : ODF 1.0 approuvé, norme internationale **ISO/IEC 26300:2006** (publication 12-2006)
 - 10-2006 : ODF 1.1 validé par l'OASIS
 - 2006-... : ODF 1.2 en cours
 - 2007-... : ODF 1.1 utilisé par OpenOffice.org
-
-



ODF – les jalons (3)

- 9-2007 : ISO/IEC 26300:2006 refusé par l'AFNOR en tant que Norme Française
 - 2007-... : influence grandissante d'IBM dans l'évolution des spécifications
 - 2008-... : entrée et participation active de Microsoft au comité technique ODF de l'OASIS
 - 2009 : prise en charge (annoncée) d'ODF par Microsoft Office
-
-



ODF – impact direct

- la « preuve du concept » d'un standard ouvert **exploitable** est faite
- la pression des pouvoirs publics en faveur d'un « standard de droit » devient explicite
- 2004 : l'UE insiste pour la *normalisation* des FDR





ODF – influence politique

« Because of its specific role in society, the public sector must avoid that a specific product is forced on anyone interacting with it electronically. Conversely, any document format that does not discriminate against market actors and that can be implemented across platforms should be encouraged. »

« Microsoft should consider the merits of submitting XML formats to an international standards body of their choice »



IDABC

European eGovernment Services

5-2004 <http://ec.europa.eu/idabc/en/document/2592/5588>



ODF – la grande illusion ?

- la phénomène ODF a suscité trop d'attentes
 - la viabilité technique d'une spécification ne détermine pas son adoption effective
 - la normalisation ne s'impose qu'à ceux qui le veulent bien
 - la suite des événements va relativiser le « triomphe »
 - en 2005 la « guerre des formats » dans le Massachusetts donne une mesure des enjeux et des luttes d'influence en perspective
-
-



Schémas Office 2003 : **Microsoft en route vers la standardisation**

- 9-2003 : publication et liberté d'utilisation des schémas XML de *Microsoft Office 2003*
 - les spécifications sont séparées par produits (*WordprocessingML, SpreadsheetML, FormTemplate XML, etc*); *Powerpoint* n'est pas couvert
 - faible succès, dû aux limites fonctionnelles et aux incertitudes juridiques
 - la pression de l'UE et « l'affaire » du Massachusetts vont déterminer un engagement plus fort
-
-



OOXML – le lancement

- 11-2005 : les schémas *Office 2003* sont modifiés, complétés (ex: prise en charge de *Powerpoint*, ajout de parties communes), associés à une spécification de stockage (*Open Packaging Conventions*) et soumis à l'ECMA
- 12-2005 : naissance du projet **Office Open XML** et lancement du *Comité Technique 45* (TC45), responsable de la spécification au sein de l'*ECMA*



ECMA TC45

Apple, Barclays Capital, BP, British Library, Essilor, Intel, Microsoft, NextPage, Novell, Statoil, Toshiba, United States Library of Congress

- pluridisciplinarité
- compétence technique
- MS + utilisateurs, constructeurs & éditeurs
- ... *mais où est la concurrence ?*





OOXML – la ruée

- 9-2006 : spécification OOXML complète soumise par le TC45 à l'A.G. de l'ECMA
- 12-2006 : spécification validée comme standard **ECMA 376** malgré opposition d'IBM, et soumise au JTC1 (ISO/IEC)
- 1-2007 : déploiement commercial d'Office 2007



OOXML – l'adoption

- 9-2007 : ajournement de la procédure, demande de correctifs substantiels par de nombreuses agences nationales de normalisation
- 2-2008 : réingénierie lourde (correction d'erreurs techniques et éditoriales, repositionnement des spécifications dites transitoires)
- 4-2008 : OOXML adopté, norme internationale **ISO/IEC 29500:2008** (publication 11-2008)



point de situation en décembre 2008 (1)

- ODA : norme internationale et norme française, toujours en vigueur, jamais utilisée
- ODF 1.0 (ISO/IEC 26300) : norme internationale non homologuée par la France, utilisée
- ODF 1.1: (OASIS ODF), standard privé, utilisé (OpenOffice.org, support MS-Office annoncé)
- ODF 1.2 : en voie de développement, projet de soumission au JTC1 (2009 ?), partiellement utilisé (OpenOffice.org 3.0)

.../...



point de situation en décembre 2008 (2)

- OOXML (ECMA 376) : utilisé (Office 2007, OpenOffice.org...), standard privé
 - OOXML (ISO/IEC 29500) : normalisé, partiellement utilisé (partie commune ISO/ECMA et fonctions transitoires)
 - formats propriétaires Microsoft récents : non destinés à être normalisés, utilisés (Office 2007)
 - formats propriétaires Microsoft « classiques » : non destinés à être normalisés mais *massivement dominants*
-
-



et pendant ce temps...

... le JTC1/SC34 et les agences affiliées (dont AFNOR) commercialisent une prestation d'harmonisation et de convergence entre ODF et OOXML !!!



problèmes soulevés

- normalisation décalée de la réalité des besoins
- dérapage de la normalisation institutionnelle et concurrence des normes
- deux visions de l'interopérabilité
- deux conceptions d'un format bureautique



précipitation et dérive

- les procédures accélérées sont adaptées aux projets de normes *simples et/ou proches de leur version finale*, recueillant un *large consensus* et correspondant à des *opportunités industrielles immédiates*
- le dossier Office Open XML *ne répondait pas à ces conditions* ; son adoption en mode accéléré était donc une *anomalie institutionnelle* qui reste à expliquer



L'étrange dossier ISO/IEC 29500

- nécessité d'un "deuxième tour" et ampleur sans précédent des correctifs : *le dossier était-il prêt ?*
- ISO/IEC 29500 a divergé d'Office 2007 et n'est pas immédiatement exploitable comme plate-forme de compatibilité; *où est l'urgence industrielle ?*
- aucun concurrent direct* de Microsoft dans le domaine concerné ne soutient ISO/IEC 29500; plusieurs pays ont fait appel contre la décision; *où est le consensus ?*

* IBM, Sun/OpenOffice.org, Adobe, Google, Oracle, KDE, Red Hat...



concurrency des normes

- l'absence de définition du mot *contradiction* ouvre les portes de l'ISO sans logique précise (autre que l'augmentation des « ventes » de normes)
 - la concurrence des standards est normale quand elle découle d'un historique ancien et chargé
 - elle est étrange quand il s'agit de deux spécifications *récentes* soumises à *un an d'intervalle* aux *mêmes instances...*
 - pourquoi dédoubler une norme au lieu de l'enrichir ?
-
-



concurrence des normes **un point de vue français abandonné**

« Le référentiel général d'interopérabilité (RGI), défini par la direction générale de la modernisation de l'État (DGME), préconise ce format ODF sous sa référence ISO. [...] Pour éviter un conflit de normes, qui serait préjudiciable aux gains attendus du cadre normatif du RGI à moyen et long terme, les représentants de l'État privilégient une position qui n'aboutisse pas à deux normes concurrentes mais à l'enrichissement de la norme existante. »

Extrait de la réponse à la question écrite n°117669 – 10 avril 2007
<http://www.questions.assemblee-nationale.fr/q12/12-117669QE.htm>



concurrence des normes un point de vue français attentiste

« Les promoteurs du projet de norme OOXML ont pris des engagements en vue de permettre l'interopérabilité avec la norme ODF. [...] Ces engagements répondaient aux principales objections mises en avant par la France pour justifier son vote négatif en septembre 2007. [...] Ce projet de norme a donc été adopté. AFNOR sera désormais très vigilante à ce que les engagements pris soient effectivement tenus. Elle participera aux travaux menés à l'ISO sur le sujet et tiendra les membres de la commission de normalisation française régulièrement informés de leur état d'avancement. »

Extrait de la réponse à la question écrite n°23299 – 9 septembre 2008
<http://questions.assemblee-nationale.fr/q13/13-23299QE.htm>



complémentarité ODF/OOXML ?

- selon ses promoteurs, OOXML ne correspond pas aux mêmes besoins qu'ODF, mais...
- la liste des besoins correspondant à chacun des deux formats n'a jamais été écrite
- les deux normes se définissent elles-mêmes comme visant le même périmètre applicatif





complémentarité ODF/OOXML ?

ODF

« This document defines an XML schema for office applications and its semantics. The schema is suitable for office documents, including text documents, spreadsheets, charts and graphical documents like drawings or presentations, but is not restricted to these kinds of documents. »

ISO/IEC 26300:2006 §1.1 "Introduction"



complémentarité ODF/OOXML ?

OOXML

« ISO/IEC 29500 defines a set of XML vocabularies for representing word-processing documents, spreadsheets and presentations. »

ISO/IEC 29500:2008 §1.1 "Scope"



interopérabilité

l'option OOXML

*« Par défaut, les documents créés avec Microsoft Office 2007 utiliseront le format de fichier Open XML d'Office. Ce format de fichier est basé sur XML et permet une meilleure **intégration des données entre les documents et les systèmes centraux.** La spécification d'un format de fichier ouvert et libre de royalties est le garant d'une **interopérabilité accrue dans les environnements hétérogènes.** Les fournisseurs de technologie se voient ainsi offrir la possibilité d'**intégrer des documents Microsoft Office dans leurs solutions.** »*



interopérabilité

l'option ODF

*« Afin de simplifier les **transformations entre documents** et maximiser l'**interopérabilité**, ODF utilise des standards reconnus. »*

*« La transparence du format OpenDocument permet à chaque acteur ou organisation d'**échanger tous les documents** nécessaires à leur fonctionnement avec la parfaite maîtrise de tous les éléments de la chaîne de ces **échanges**. »*



deux visions de l'interopérabilité

- les administrations qui encouragent la normalisation recherchent avant tout la compatibilité dans un but de mise en concurrence des fournisseurs, mais...
 - cet objectif est difficile à concilier avec le modèle d'un fournisseur dont la position dominante s'appuie en grande partie sur les marchés publics
 - peut-on vraiment demander à ce fournisseur de soutenir une vision de l'interopérabilité susceptible de faciliter la fuite vers des offres alternatives ?
-
-



architecture ODF

- spécification intégrée
 - cohérence syntaxique et non-redondance
 - absence de couverture de la logique applicative (formules, macros, schémas métier)
 - effort déclaré de cohérence externe ; réutilisation (mais partielle et imparfaite) de standards existants
 - compatibilité Microsoft *non prioritaire* et laissée sous la responsabilité des *applications*
-
-



architecture OOXML

- spécifications séparées selon la logique des trois composantes classiques d'une suite bureautique (tableur, ttx, présentation) et faible cohérence interne (syntaxe hétérogène)
 - couverture partielle de la logique applicative (formules, schémas métier, mais pas macros)
 - compatibilité *Microsoft* prioritaire, références externes privées, cohérence externe non prioritaire
-
-



adhérence au monde Microsoft

une priorité affirmée

« OpenXML was designed from the start to be capable of faithfully representing the pre-existing corpus of word-processing documents, presentations, and spreadsheets that are encoded in binary formats defined by Microsoft Corporation. »



adhérence au monde Microsoft conséquences sur la conception

- reflet de la division *Word/Excel/Powerpoint* :
incohérence syntaxique et redondance
- importance élevée de la rétro-ingénierie syntaxique
de bas niveau au détriment de la conception d'un
modèle de données « orienté métier »
- la norme est encombrée d'une proportion
anormalement élevée de spécifications dites
« transitoires »



adhérence au monde Microsoft **une inconnue majeure**

Malgré le souci de continuité avec l'existant, la correspondance entre Open XML et les anciens formats Microsoft n'est pas publiquement spécifiée. Les seuls outils de conversion disponibles dépendent de l'environnement Microsoft Office.

Ni ODF, ni OOXML n'assure la reprise de l'existant hors contexte propriétaire.



un bilan mitigé

- la valeur ajoutée des 5 dernières années de normalisation des FDR reste à démontrer
- la normalisation s'est transformée en une chasse à la labellisation sans perspective réelle de convergence
- le système institutionnel de normalisation s'est avéré incapable de produire un consensus et sa crédibilité est affaiblie





dommages collatéraux

« The days of open standards development are fast disappearing. Instead we are getting “standardization by corporation”... »

Martin Bryan, Coordinateur ISO/IEC JTC1/SC34 WG1
novembre 2007

<http://www.jtc1sc34.org/repository/0940.htm>



dommages collatéraux

- l'IEEE s'est retirée du débat dès 2007
 - la Suède a annulé son propre vote pour ISO/IEC 29500
 - quatre pays font (vainement) appel auprès du JTC1 pour faire annuler la procédure ISO/IEC 29500 (Brésil, Cuba, Inde, Vénézuéla)
 - de grands acteurs publics et *tous* les compétiteurs de Microsoft dans le domaine bureautique expriment publiquement leur défiance envers le JTC1
-
-



conclusion

La gouvernance des processus de normalisation en vigueur n'est pas adaptée au contexte stratégique des FDR. La normalisation institutionnelle reflète l'état présent des ententes et des rapports de forces entre fournisseurs.

Les industriels utilisateurs en quête de garanties de pérennité des documents bureautiques n'obtiendront gain de cause qu'en usant de leurs propres canaux d'influence et de leur capacité de mise en concurrence des fournisseurs, indépendamment des « labels » officiels