

# OMPI



SCIT/SDWG/5/9

ORIGINAL : anglais

DATE : 17 septembre 2004

F

ORGANISATION MONDIALE DE LA PROPRIÉTÉ INTELLECTUELLE  
GENÈVE

COMITÉ PERMANENT DES TECHNIQUES DE L'INFORMATION

GRUPE DE TRAVAIL SUR LES NORMES  
ET LA DOCUMENTATION

Cinquième session  
Genève, 8 – 11 novembre 2004

NORME ST.36 DE L'OMPI

*Document établi par le Secrétariat*

1. Le Groupe de travail sur les normes et la documentation (SDWG) du Comité permanent des techniques de l'information (SCIT) a adopté, pendant sa quatrième session tenue en janvier 2004, un calendrier pour l'examen de la norme ST.36 de l'OMPI proposée (Recommandation concernant le traitement des documents de brevet utilisant une norme en XML (Extensible Markup Language)) (paragraphe 86 et 87 du document SCIT/SDWG/4/14). Conformément à ce calendrier, l'Équipe d'experts chargée des normes relatives au traitement et à l'échange des données électroniques (ci-après dénommée "Équipe d'experts ED PES") a étudié le projet de norme proposée, qui est joint en annexe au présent document aux fins de son examen et de son approbation par le SDWG.

2. L'annexe contient la norme proprement dite. Des éléments supplémentaires relatifs à la norme ST.36 de l'OMPI proposée tels que des définitions de type de document (DTD), des éléments communs internationaux (ICE) et des exemples qui ne font pas partie intégrante de la norme figurent à l'adresse suivante : <http://www.wipo.int/en/taskforce/edpes/dtds.html>. La liste correspondante sera actualisée à mesure que le Bureau international recevra ces éléments complémentaires. Les annexes A et B de la norme seront également disponibles à l'adresse susmentionnée.

3. Compte tenu du caractère technique et approfondi du présent document, il est demandé, à titre exceptionnel, que les commentaires et les questions d'ordre technique concernant ce document soient adressés par écrit au Bureau international avant la session du SDWG qui se tiendra en novembre 2004, afin de permettre l'élaboration de réponses techniques détaillées.

4. *Le SDWG est invité :*

a) *à adopter officiellement le titre "Norme ST.36 de l'OMPI (Recommandation concernant le traitement des documents de brevet utilisant une norme XML (Extensible Markup Language))" pour la norme proposée;*

b) *à examiner et approuver la norme ST.36 de l'OMPI qui figure dans l'annexe du présent document.*

[L'annexe suit]

ANNEXE

**NORME ST.36**

RECOMMANDATION CONCERNANT LE TRAITEMENT DES DOCUMENTS DE BREVET  
UTILISANT UNE NORME EN XML (EXTENSIBLE MARKUP LANGUAGE)

**TABLE DES MATIÈRES**

<b>Norme ST.36</b> .....	2
INTRODUCTION.....	2
DÉFINITIONS.....	3
PORTÉE DE LA NORME.....	3
PRESCRIPTIONS DE LA NORME.....	4
Généralités .....	4
Caractères.....	5
Nommage des éléments communs internationaux .....	6
Nommage des éléments propres aux offices .....	6
Attributs .....	7
Adjonction, dépréciation ou modification d'éléments .....	7
Conventions relatives aux éléments et aux attributs.....	7
Conventions relatives aux DTD.....	8
Conventions relatives aux documents .....	11
Entités externes .....	12
TIFF.....	12
JPEG .....	13
Norme ST.33 de l'OMPI.....	13
Norme ST.35 de l'OMPI.....	13
PDF.....	13
MÉGACONTENU.....	13
DTD couramment utilisées dans l'industrie.....	14
DTD type pour les publications de brevet.....	14
RÉFÉRENCES.....	14
 ANNEXE A : xx-patent-document.dtd.....	
 ANNEXE B : Example XML document instance.....	

**NORME ST.36**

RECOMMANDATION CONCERNANT LE TRAITEMENT DES DOCUMENTS DE BREVET  
UTILISANT UNE NORME EN XML (EXTENSIBLE MARKUP LANGUAGE)

INTRODUCTION

1. La présente norme recommande l'utilisation de ressources XML (eXtensible Markup Language) pour le dépôt, le traitement, la publication et l'échange de tout type d'information en matière de brevets. Elle s'inspire largement de l'appendice I de l'annexe F de la septième partie des Instructions administratives du Traité de coopération en matière de brevets, (ci-après dénommée "Annexe F"). Le terme "ressources XML" vise tout élément utilisé pour créer et faire fonctionner une application XML. Bien qu'en principe les ressources XML incluent les feuilles de style, les schémas du World Wide Web Consortium (Consortium W3C) et d'autres objets, cette norme ne contient pour l'instant que des définitions de type de document (DTD), des modèles de contenu, des éléments et un jeu restreint d'entités caractères. Pour de plus amples informations sur le Consortium W3C, consulter le site <http://www.w3c.org/>.

2. La présente norme est une application de la norme *Extensible Markup Language (XML) 1.1*.  
Voir : <http://www.w3.org/TR/2004/REC-xml11-20040204/>:

*"Le langage de balisage extensible (Extensible Markup Language, XML) est un sous-ensemble de SGML qui est complètement décrit dans ce document. Son but est de permettre au SGML générique d'être transmis, reçu et traité sur le Web de la même manière que l'est HTML aujourd'hui. XML a été conçu pour être facile à mettre en œuvre et interopérable avec SGML et HTML."*

3. Le balisage d'un document XML conforme à la présente norme est un exemple de représentation du contenu d'un document utilisant une norme en XML, dans lequel *"les documents se composent d'unités de stockage appelées entités, qui contiennent des données analysables ou non. Les données analysables se composent de caractères, certains formant les données textuelles, et le reste formant le balisage. Le balisage décrit les structures logiques et de stockage du document. XML fournit un mécanisme pour imposer des contraintes à ces structures."* (Consortium W3C).

4. Le langage XML ne peut pas être utilisé en tant que tel comme base pour le traitement des documents de brevet – *"Cette spécification ne contraint pas la sémantique, l'utilisation ou (au-delà de la syntaxe) les noms des types d'éléments ou d'attributs ..."* (Consortium W3C).

5. Par conséquent, la présente norme définit les éléments et leurs identificateurs génériques, ou "balises", et les attributs pour le balisage des documents de brevet. C'est-à-dire qu'elle indique dans une certaine mesure la sémantique (signification), l'utilisation et les noms des éléments et attributs qui constituent les différents types de documents visés.

*"[Définition : chaque document XML contient un ou plusieurs éléments, dont les limites sont marquées soit par des balises ouvrantes et fermantes, soit, pour les éléments vides, par une balise d'élément vide. Chaque élément a un type, identifié par un nom, parfois appelé son "identificateur générique" (IG), on peut y associer un jeu de spécifications d'attribut.] Chaque spécification d'attribut comprend un nom et une valeur."* (Consortium W3C)

Note : pour une description complète et des définitions, voir la spécification XML sur le site <http://www.w3.org/TR/2004/REC-xml11-20040204/>.

6. La présente norme a pour but de fournir des structures logiques, indépendantes de tout système, pour le traitement des documents de brevet, qu'il s'agisse de texte ou de données images. Cela signifie qu'elle peut être utilisée à la place des normes ST.30, ST.32, ST.33 et ST.35 de l'OMPI pour le dépôt, le traitement, la publication et l'échange d'informations bibliographiques, d'abrégés ou de tous types de documents de brevet en texte intégral. La norme fournit des ressources XML pour les données suivantes :

a) Le texte complet ou partiel des documents de brevet, y compris les données bibliographiques, enregistré sous forme codée caractère par caractère.

b) Des pages entières de documents représentées sous la forme d'une seule image, quel que soit leur contenu (données bibliographiques, texte ou images).

c) À l'intérieur des documents (texte complet), des données qui ne peuvent pas être codées caractère par caractère telles que les dessins, les formules chimiques, en particulier les tableaux complexes (appelées des "images incrustées").

7. Les documents XML conformes à la présente norme doivent être bien formés et correspondre à l'une des définitions de type de document (DTD) qui figurent dans l'annexe F ou à une DTD propre à un office elle-même conforme à la présente norme. Une DTD conforme à la présente norme doit être constituée à partir des éléments figurant dans les principes directeurs relatifs à cette norme. Les DTD de l'annexe F sont publiées sur le site <http://www.wipo.int/pct-safe/epct/schemadocs/> et actualisées sur ce site dès qu'une modification est approuvée. Dès qu'une DTD actualisée est publiée sur le site Web, elle peut être utilisée officiellement.

## DÉFINITIONS

8. Aux fins de la présente norme, les définitions suivantes sont précisées :

a) Par **documents de brevet** il faut entendre les brevets d'invention, les brevets de plante, les brevets de dessin ou modèle, les certificats d'utilité, les modèles d'utilité, les documents additionnels correspondants, les demandes et les mémoires descriptifs publiés, les types de documents relatifs au traitement des demandes de brevet, y compris les activités postérieures à la délivrance du brevet, le maintien des droits de propriété et toutes les communications entre offices et déposants et entre offices.

b) Le **balisage** est défini comme étant du texte ajouté au contenu d'un document, qui en décrit la structure et d'autres attributs d'une manière indépendante de tout système et de tout traitement dont ce document peut faire l'objet.

c) Pour d'autres définitions, consulter la spécification XML à l'adresse <http://www.w3c.org/TR/2004/REC-xml11-20040204/>.

## PORTÉE DE LA NORME

9. Bien que les DTD figurant dans l'annexe F aient été conçues à des fins d'utilisation dans le cadre du Traité de coopération en matière de brevets, la présente norme vise à généraliser leur utilisation par tous les offices de brevets pour le dépôt électronique. La DTD type ci-dessous vise à donner des indications pour l'utilisation des éléments communs internationaux (ICE) aux fins de la publication des documents de brevet. D'autres DTD pourront être ajoutées à la liste au fur et à mesure de l'évolution de la norme.

### *Liste des DTD de l'annexe F*

(voir <http://www.wipo.int/pct-safe/epct/xml/canon.htm>)

amendment-request.dtd  
application-body.dtd  
application-receipt-list.dtd  
indication-bio-deposit.dtd  
declaration.dtd  
demand.dtd  
dispatch-list.dtd  
ex-officio-correction.dtd  
fee-sheet.dtd  
ipea-demand-receiving-info.dtd  
lprp.dtd  
package-data.dtd  
pkgheader.dtd  
power-of-attorney.dtd  
priority-doc.dtd  
receiving-office-request-info.dtd  
request.dtd  
search-report.dtd  
table-external.dtd  
xmit-receipt.dtd

### *DTD type*

xx-patent-document.dtd

### *DTD courantes dans l'industrie, incorporées par renvoi*

mathml2.dtd  
soextblx.dtd (aussi désignée sous la forme  
calstblx.dtd)

SCIT/SDWG/5/9  
Annexe, page 4

10. On trouvera ci-après certaines DTD de l'annexe F avec le type d'opération correspondant pour illustrer leur utilisation proposée. Le tableau ne constitue qu'une indication de l'utilisation possible de ces DTD dans le cadre des opérations relatives aux brevets, les offices pouvant avoir des besoins différents.

Nom de la DTD	Type d'opération						
	Dépôt	Publication	Instruction	Délivrance	Après-délivrance	Nouvelle publication	Correspondance
amendment-request	★					★	
application-body	★	★		★		★	
bio-deposit	★	★		★		★	
declaration	★						★
demand	★						★
dispatch	★						
fee-sheet	★						★
iprp			★				★
package-data	★	★	★	★	★	★	★
pkgheader	★	★	★	★	★	★	★
power-of-attorney	★		★	★	★		★
priority-doc	★						★
request	★						
search-report		★				★	★
table-entity	★	★		★		★	
xmit-receipt	★		★		★		★
xx-patent-document		★		★		★	

PRESCRIPTIONS DE LA NORME

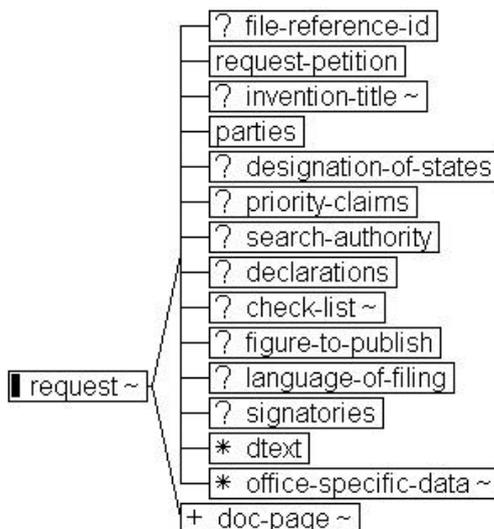
*Généralités*

11. Les éléments communs internationaux constituent le fondement de la présente norme. Ils découlent de l'annexe F, de la norme ST.32 de l'OMPI et d'autres sources. Voir [http://www.wipo.int/pct-safe/epct/xml\\_canon.htm/](http://www.wipo.int/pct-safe/epct/xml_canon.htm/).

12. Les éléments communs internationaux doivent être utilisés selon les modalités prévues dans la présente norme, c'est-à-dire qu'ils doivent avoir le même nom, le même contenu, les mêmes attributs et la même signification que ce qui est indiqué dans la liste des éléments communs internationaux. Il est entendu que cette norme et l'annexe F ne peuvent pas contenir tous les éléments demandés par l'ensemble des offices de brevets; le cas échéant, des éléments propres à un office peuvent être utilisés comme cela est indiqué ci-dessous.

13. Les informations propres à un office peuvent être traitées ainsi :

a) Isolées dans une DTD distincte appelée par exemple de la "DTD request" par l'élément `office-specific-data` (recommandé).



b) Incluses directement dans l'élément `office-specific-data`, auquel cas l'élément vide peut être modifié et inclure `#PCDATA` ou d'autres modèles de contenu, le cas échéant. Le préfixe constitué du code de pays à deux lettres est ajouté à `office-specific-data` : par exemple, `wo-office-specific-data`. Le modèle de contenu `office-specific-data` ne doit pas être modifié sans adjonction du préfixe de l'office.

c) Compte tenu de la convention sur les espaces de nommage dans XML. Les espaces de nommage XML offrent une méthode simple pour qualifier les noms des éléments et des attributs utilisés dans des documents XML, en associant ceux-ci avec des espaces de nommage désignés par des références d'URI (système universel d'identification des ressources Internet). (Voir : <http://www.w3.org/TR/REC-xml-names/>)

14. À un niveau supérieur, ces données peuvent figurer dans des documents distincts appelés de la DTD `package-data` par l'élément `other-documents`. Les DTD ou balises appelées par `office-specific-data` ou `other-documents` sont totalement sous le contrôle de l'office responsable.

15. Le nom des DTD ou des éléments propres à un office commence par le code de pays à deux lettres de l'office correspondant indiqué dans la norme ST.3, suivi d'un séparateur (trait d'union ou deux points) et du nom de l'entité. Tous les autres noms seront considérés comme des DTD internationales ou des éléments internationaux (génériques). Il est donc conseillé de limiter exclusivement l'emploi de noms commençant par un mot à deux lettres à ceux qui représentent un code de pays valable. Par exemple, `request` devient `ep-request` lorsqu'il est modifié pour être utilisé par l'OEB, pour le nom du fichier DTD et l'élément racine.

16. Aux fins de l'interfonctionnement en matière de dépôt entre les offices de brevets, il est nécessaire d'utiliser les DTD suivantes telles qu'elles sont définies dans l'annexe F, à l'exclusion de toute autre solution propre aux offices : `application-body`, `table-external`, `pdoc-certificate`, `package-header`, `package-data` et `xmit-receipt`.

17. Lorsque des documents contiennent des éléments propres à un office ou mentionnent des DTD propres à un office, l'administration qui publie des documents établit à l'intention des autres offices et des utilisateurs un avis contenant des explications sur le contenu et la signification de ces éléments et des DTD. Cet avis doit être publié sur un site Web géré par l'office et facilement accessible ou sur le site Web de l'OMPI. Il doit contenir les DTD et une description complète de chacun des éléments.

#### Caractères

18. Bien que le langage XML permette d'autres codages des caractères, la présente norme recommande d'utiliser exclusivement Unicode. Il peut être utile d'ajouter des entités de caractères pour les caractères qui ne figurent pas encore dans Unicode, tels que ceux qui sont énumérés dans `wipo.ent` (voir [http://www.wipo.int/pct-safe/epct/xml\\_canon.htm/](http://www.wipo.int/pct-safe/epct/xml_canon.htm/)). Le fichier de cette entité contient des noms d'entité généraux qui peuvent être utilisés dans des documents à la place des points de code des codages avec lesquels ils sont mis en correspondance dans `wipo.ent`. L'utilisation de ces entités nécessite, pour la présentation, la création de glyphes qui n'existent pas encore. Voir <http://www.w3.org/XML/Core/2002/10/charents-20021023> pour de plus amples informations sur les entités de caractères.

SCIT/SDWG/5/9  
Annexe, page 6

19. Les documents doivent inclure l'instruction de traitement suivante à la première ligne du fichier. Noter que cette norme ne s'applique qu'à UTF-8.

```
<?xml version='1.1' encoding='utf-8' ?>
```

20. Cependant, dans le cas d'écritures logographiques, Unicode en UTF-8 peut créer des fichiers exceptionnellement volumineux puisque le codage peut utiliser quatre, six ou même huit octets par caractère. Les offices nationaux peuvent alors choisir une police et un codage qui ramènent les fichiers à des tailles gérables. Les offices qui choisissent ce mode d'action doivent être prêts à consulter les partenaires avec lesquels ils font des échanges et à informer le public en conséquence.

21. Les caractères qui peuvent figurer dans un document XML sont précisés dans la *Recommandation XML 1.1 du Consortium W3C* et sont repris dans la présente norme avec l'exception ci-après. Les caractères utilisés dans des noms d'éléments ou d'attributs décrits dans la présente norme se limitent à la série suivante :

```
{abcdefghijklmnopqrstuvwxyz1234567890-}.
```

22. Les offices sont vivement encouragés à créer, à des fins de publication et d'échange, des documents qui ont été "normalisés" conformément au *Character Model for the World Wide Web* (Modèle de caractères pour le World Wide Web) (<http://www.w3.org/TR/2003/WD-charmod-20030822/>). Les programmes d'analyse syntaxique qui utilisent XML 1.1 peuvent servir à effectuer des essais de normalisation. De tels essais permettront de renforcer sensiblement la cohérence des opérations de tri et de comparaison de chaînes en garantissant que certaines options de codage de caractères disponibles sous Unicode ont été mises en œuvre systématiquement dans l'ensemble de la communauté internationale des brevets.

#### *Nommage des éléments communs internationaux*

23. Tous les noms d'éléments doivent être constitués de mots anglais.

24. Lorsqu'un nom d'élément nécessite plus d'un mot, les mots utilisés sont séparés par un tiret (-).

25. Dans la présente norme, les noms d'éléments sont écrits en alphabet latin uniquement, limité au jeu de caractères suivant : {abcdefghijklmnopqrstuvwxyz1234567890-}. Les caractères accentués et les majuscules ne sont pas utilisés. Pour des raisons historiques, les noms d'éléments figurant dans l'élément SDOBI tiré de la norme ST.32 conservent le B majuscule et d'autres noms d'éléments en majuscules.

26. Dans la mesure du possible, les noms sont descriptifs, et non mnémotechniques ni abrégés. Quiconque doit pouvoir comprendre la signification du nom d'élément en n'ayant guère ou pas du tout besoin de consulter une autre documentation. Parmi les exceptions notables figurent les éléments les plus courants figurant dans les demandes de brevet, comme p pour paragraphe et d'autres éléments issus, par exemple, du langage HTML, et certains autres éléments de formatage largement utilisés (par exemple, voir la DTD *application-body*). Il est peu probable que d'autres exceptions seront nécessaires.

27. La signification du nom et le contenu prévu de l'élément sont précisés en une ou deux phrases. Le texte correspondant doit mentionner toutes les règles applicables et résumer très brièvement leur contenu. Dans une DTD, les grandes lignes de ce texte doivent figurer dans un commentaire précédant immédiatement l'élément ou l'attribut auquel s'applique la DTD.

28. Pour des raisons historiques, certains composants des éléments communs internationaux figurent dans la colonne nom de la norme ST.32 sous la forme B<sub>n</sub>n<sub>n</sub> dans laquelle n est un nombre. Ces noms d'élément sont les éléments correspondants indiqués dans la norme ST.32 de l'OMPI – *Recommandation concernant le balisage des documents de brevet selon le SGML (langage normalisé de balisage généralisé)*. Cette colonne pourra être supprimée en temps utile lorsque les offices cesseront d'utiliser les noms de balise B<sub>n</sub>n<sub>n</sub>.

#### *Nommage des éléments propres aux offices*

29. Les règles indiquées pour les éléments communs internationaux (section précédente) s'appliquent.

30. Tous les noms d'éléments propres à un office sont des mots anglais dans la mesure du possible.

31. Chaque nom d'élément propre à un office est précédé du code de la norme ST.3 pour l'office qui détient l'élément en question. Le code de la norme ST.3 est séparé du nom de l'élément par un tiret (-) ou deux points (:). Par exemple, jp:fterm, ou ep-printer-name. Les deux points ne sont utilisés que lorsque l'office qui détient l'élément utilise les espaces de nommage XML du Consortium W3C (voir *Namespaces in XML*, <http://www.w3.org/TR/REC-xml-names/>).

#### *Attributs*

32. Si un office souhaite ajouter des attributs à un élément commun international ou modifier certains attributs, il doit présenter une demande de modification.
33. Les éléments propres aux offices peuvent être assortis de tous les attributs nécessaires, à condition qu'ils ne soient pas incompatibles avec les attributs définis dans la présente norme (voir les paragraphes suivants).
34. Les noms d'attribut ne doivent pas être redéfinis dans une DTD. Le nom doit donc toujours avoir la même signification, quel que soit l'élément auquel il s'applique. Un commentaire doit expliquer les valeurs possibles que cet attribut peut avoir, ce qu'elles signifient et, le cas échéant, comment les interpréter. Le commentaire relatif à l'attribut doit figurer dans le commentaire relatif à l'élément auquel l'attribut s'applique.
35. Les noms d'attribut doivent être réutilisés lorsque les valeurs de l'attribut autorisées sont identiques ou lorsque la valeur de l'attribut a la même signification. Si les valeurs des attributs ne sont pas identiques, les noms doivent aussi être différents. Par exemple, l'attribut `file` doit être utilisé chaque fois que la valeur de l'attribut est le nom d'un fichier informatique. Aucun autre nom d'attribut ne doit être utilisé à cette fin.

#### *Adjonction, dépréciation ou modification d'éléments*

36. L'adjonction, la dépréciation ou la modification d'éléments communs internationaux doit être approuvée conformément aux procédures établies par l'OMPI. En général, les ressources XML publiées sur le site <http://www.wipo.int/pct-safe/epct/xml/canon.html>, y compris celles qui sont citées dans la présente norme, sont révisées lors de l'examen et de l'approbation des propositions par l'Équipe d'experts EDPEs du Groupe de travail sur les normes et la documentation.
37. Une fois établi qu'un élément commun international est déprécié (il ne doit plus être utilisé), le changement de statut sera indiqué dans le répertoire des éléments communs internationaux, avec la date à partir de laquelle l'élément n'est plus utilisé.
38. L'adjonction, la dépréciation ou la modification des éléments propres à un office ne requiert que l'approbation de l'office de propriété industrielle qui crée l'élément. Toutefois, les offices sont censés veiller à ce que les éléments qui leur sont propres aient une signification et une description distinctes de celles des éléments communs internationaux.
39. Un élément propre à un office qui est jugé utile par d'autres offices que celui qui l'a créé peut être considéré comme un élément commun. Le passage au statut d'élément commun suppose normalement la dépréciation de l'élément propre à l'office et l'adjonction d'un nouvel élément commun international qui a le même nom que l'élément propre à l'office moins le préfixe constitué du code de la norme ST.3.
40. Lorsqu'un office a décidé qu'un élément qui lui est propre est déprécié (il n'utilisera plus cet élément), il demandera que soient notées dans le répertoire la modification du statut de l'élément et la date à partir de laquelle il n'utilisera plus l'élément.
41. Lorsque cela est possible, les demandes de modification relatives à des DTD concernées par l'adjonction, la dépréciation ou la modification d'un élément commun ou d'un élément propre à un office doivent être présentées en même temps que la demande relative à la création, la dépréciation ou la modification de l'élément. L'entrée en vigueur des éléments et des DTD modifiés devra coïncider le plus souvent possible.

#### *Conventions relatives aux éléments et aux attributs*

42. Certains éléments et attributs ont une signification particulière dans le cadre des éléments communs internationaux. La présente section explique leur finalité en raison des avantages importants qui en découlent s'ils sont traités de façon uniforme dans l'ensemble de la communauté de la propriété industrielle.
43. L'élément `date` est largement utilisé dans l'ensemble des modèles de contenu de `xx-patent-document` pour chaque occurrence d'une date. Par conséquent, il doit toujours apparaître comme l'enfant d'un autre élément qui définit le contexte et explique à quel événement correspond la date. Même si le mot `date` apparaît dans l'élément parent, l'élément `date` doit être utilisé dans cet élément pour indiquer la date réelle. Le contenu de l'élément doit toujours être exprimé sous la forme d'une série de chiffres dont les quatre premiers correspondent à l'année, les deux suivants au mois (celui de gauche étant remplacé par zéro si nécessaire) et les deux derniers au jour du mois (celui de gauche étant remplacé par zéro si nécessaire). Par exemple, 20040717 correspond à 2004 juillet 17.
44. L'élément `document-id` est utilisé pour identifier tous les types de document de propriété industrielle auxquels la présente norme s'applique. Par conséquent, il doit toujours apparaître comme l'enfant d'un autre élément qui définit le contexte et explique quel document est identifié. L'élément `date` dans `document-id` correspond à la date de publication des documents ou à la date de dépôt des documents non publiés. Aucune autre date ne doit être

SCIT/SDWG/5/9  
Annexe, page 8

utilisée. L'élément `doc-number` ne doit pas contenir de code de type, ni de date, ni de pays, chacune de ces données étant stockée séparément. Lorsqu'il est important d'indiquer l'identification du document conformément à la norme ST.10/C de l'OMPI ou à une autre norme de l'OMPI, utiliser une feuille de style pour présenter le contenu selon la norme. L'élément `name` est utilisé pour l'une ou l'autre des parties associées au document.

45. Les offices sont encouragés à baliser chaque occurrence de chaque numéro de type de document dans un document publié. Cela sera très positif pour les systèmes de recherche.

46. L'attribut `id` n'est jamais requis, mais son utilisation peut être particulièrement utile pour les documents. La norme XML exige que la valeur conférée à un attribut `id` soit unique dans un document. Cet attribut n'a pas d'autre but que l'identification univoque de l'objet auquel il est rattaché dans le document et il ne doit pas servir à d'autres fins. Bien que la valeur d'un attribut `id` puisse être constituée de n'importe quelle séquence de caractères, il peut être utile, à des fins de lisibilité, de définir cette valeur pour représenter le type d'élément auquel elle est attribuée. Par exemple, des valeurs `id` telles que "p0001", "p0002" ... "p0101", etc. peuvent être attribuées aux paragraphes d'une divulgation.

47. Il convient d'être prudent lorsque l'on attribue des valeurs de cette façon. Par exemple, si plus d'une série de revendications sont publiées (éventuellement dans plusieurs langues), plusieurs revendications porteront le numéro "1". Les valeurs `id` doivent alors rester uniques et la méthode décrite ici doit donc être modifiée pour garantir ce caractère unique, par exemple grâce à l'adjonction d'une indication de la langue, comme "cl-en-0001", "cl-de-0001", "cl-fr-0001", etc.

48. Si des modifications doivent être traitées automatiquement, il peut être utile de spécifier des valeurs uniques pour l'attribut `id` dans un dossier. Dans ce cas, lorsqu'un attribut `id` est associé à un paragraphe, par exemple, il ne doit jamais être modifié quel que soit le nombre de fois où le document lui-même est modifié. Lorsqu'un paragraphe est supprimé, la valeur `id` doit être supprimée pour ce dossier. Lorsqu'un nouveau paragraphe est ajouté, une valeur `id` nouvelle doit être attribuée au dossier.

49. Un attribut `id` est souvent la cible d'un attribut `idref` associé à `claim-ref`, `crossref` ou `figref`. L'appariement de valeurs `id` et `idref` permet l'établissement de liens hypertextes lorsque le document est affiché par le navigateur. L'élément `crossref` vise à désigner tout objet arbitraire dans le document, à l'exclusion d'une revendication ou d'un schéma. Pour éviter toute incompatibilité avec la pratique établie dans les navigateurs, il est important de ne pas utiliser ces attributs dans un autre but.

50. Les DTD de base et les modèles de contenu des éléments qui figurent dans ces DTD ne peuvent pas être modifiés en vue d'une utilisation nationale (on parle alors d'"éléments de base"). Même lorsqu'un élément de base est utilisé dans une DTD autre qu'une DTD de base, il ne peut pas être modifié. Si un office souhaite introduire des informations complémentaires dans un élément de base, les attributs `id` et `idref` peuvent être utilisés à cette fin sans que l'élément de base soit modifié. Rattacher un attribut `id` à l'élément en question. Créer un élément propre à l'office avec un attribut `idref` désignant l'élément de base. Dans l'élément propre à l'office, ajouter tous les éléments complémentaires nécessaires aux fins de l'utilisation au niveau national.

51. L'attribut `lang` contient normalement un code à deux lettres fondé sur les normes ISO pour la langue du contenu de l'élément auquel il est rattaché. Lorsque le code à deux lettres ne convient pas, les offices sont encouragés à appliquer les conventions établies par l'*Internet Engineering Task Force* et décrites dans le document *Tags for the Identification of Languages* (balises pour l'identification des langues) (<http://www.rfc-editor.org/rfc/rfc3066.txt>).

#### *Conventions relatives aux DTD*

52. Le nom de fichier d'une DTD indique sa version. Le nom de fichier assorti d'un numéro de version est une chaîne résultant de la concaténation du type de document suivie d'un tiret lui-même suivi de la lettre "v", d'un chiffre indiquant le numéro de la version principale et d'une section facultative constituée d'un tiret et d'un chiffre indiquant le numéro d'une révision mineure.

Exemple : `request-v1-2.dtd`

53. Chaque DTD doit contenir un numéro et une date de version dans l'identifiant public. Exemples :

```
PUBLIC "-//WIPO//DTD APPLICATION BODY 1.3//EN" "application-body-v1-3.dtd"
```

```
Reference this DTD as PUBLIC "-//USPTO//DTD us-patent-grant v4.0 2004-07-30//EN"
```

```
Alias: Grant Red Book (GRB)
```

SCIT/SDWG/5/9  
Annexe, page 9

54. Au début de la DTD, un historique des révisions indique les modifications en précisant la date et la portée de chaque modification dans l'ordre chronologique inverse. Exemples :

```
<!--
***** Revision History *****
2004-04-15 H. Li
. Added num attribute to li element.
. Changed content model maths from (img | math) to (img | (math,img?)).
. Added date attribute to element us-patent-application.
. Added math, chemistry, table, program-listing to img-content allowed values
.. and changed dna to DNA.
. Add figref to li, dd, claim-text elements and table cell entry.
. Changed subname? to subname* in serial, article and online elements.
. Renamed table-external element to table-external-ref.
2004-03-04 B. Cox
. Renamed us-continued-prosecution-application to
.. us-issued-on-continued-prosecution-application
.. and renamed attribute from cpa-text to grant-cpa-text
.. and changed attribute value (see below).
***** End Revision History *****
-->
```

55. Indiquer des interlocuteurs dans le prologue. Exemples :

```
<!--
Contacts:
EPO: Paul Brewin pbrewin-z@not-epo.org
JPO: Shiro Ankyu anky-shiro-z@not-jpo.go.jp
USPTO: Bruce B. Cox bruce.cox-z@not-uspto.gov
WIPO: Hideto Tanaka hideto.tanaka-z@not-wipo.int
-->
```

```
<!-- Contact: Bruce B. Cox
U.S. Patent and Trademark Office
Crystal Park 8, Suite 1032
Washington, DC 89231
+1-783-906-6062
bruce.cox@uspto.gov-->
```

56. Les DTD qui contiennent, par renvoi, des DTD courantes dans l'industrie sont souvent illisibles (elles ne peuvent pas être analysées) si les DTD externes ne sont pas disponibles sur l'ordinateur local. L'exemple ci-après illustre une méthode permettant de masquer facilement ces renvois externes de façon temporaire, et de mettre en forme et de visualiser la DTD lorsqu'il n'est pas facile d'accéder aux renvois externes. Cette solution suppose que l'on connaisse les sections balisées figurant dans les DTD (pour de plus amples explications, voir <http://etext.lib.virginia.edu/TEI/p4/tei.cgi?div=div2&id=SG17BIS> et des sites Web similaires).

```
<!--to include mathml2.dtd change MATHML2_DTD value to "INCLUDE",
change MATH_PLACEHOLDER value to "IGNORE", the same for the TABLE_DTD,
TABLE_PLACEHOLDER, and WIPO_ENT
INCLUDE
IGNORE
-->
<!ENTITY % UNICODE_PLANE1D_ESCAPE "IGNORE">
<!ENTITY % WIPO_ENT "IGNORE">
<!ENTITY % MATHML2_DTD "IGNORE">
<!ENTITY % TABLE_DTD "IGNORE">
<!ENTITY % MATH_PLACEHOLDER "INCLUDE">
<!ENTITY % TABLE_PLACEHOLDER "INCLUDE">
<![%UNICODE_PLANE1D_ESCAPE; [
<!ENTITY % plane1D "&#38;#38;#xE"
]]>
<![%WIPO_ENT; [
<!--
import character entity set. Download from:
http://pcteasy.wipo.int/efiling_standards/schemaDocs/wipo.ent
Note that nsgmls-based parsers (SP, Near & Far Designer, etc.)
may not be able to process this file for reasons described below
in MathML comments.
-->
<!ENTITY % wipo PUBLIC "-//WIPO//ENTITIES WIPO 1.0//EN" "wipo.ent">
%wipo;
]]>
<![%MATHML2_DTD; [
<!-- DTD MathML2: maintained by W3C. Download from:
http://www.w3.org/TR/MathML2/DTD-MathML-20010221.zip
```

## SCIT/SDWG/5/9 Annexe, page 10

If using nsgmls-based parser (SP, Near & Far Designer, etc.)  
Uncomment 'mathml-charent-module' switch below or replace the  
Referenced MathML2 DTD with the version downloadable from:  
<http://www.w3.org/Math/DTD/dtd-sp.zip>  
This notice copied from: <http://www.w3.org/Math/DTD/>  
"DTD for nsgmls  
Some systems (including the popular nsgmls parser) may not be able  
to process files using 'plane 1' characters which have Unicode  
numbers higher than #xFFFF. The versions of the DTD provided here  
incorporate the modifications mentioned above, but the high  
characters are replaced by the equivalent mchar construct  
<mchar name="..." /> this allows the DTD to be read and for MathML  
files to be validated using such systems."

```
-->
<!--ENTITY % mathml-charent.module "IGNORE" -->
<!ENTITY % MATHML.prefixed "IGNORE">
<!ENTITY % MATHML.xmlns "">
<!--import MathML2 dtd -->
<!ENTITY % mathml2 PUBLIC "-//W3C//DTD MathML 2.0//EN" "mathml2.dtd">
%mathml2;
]]>

<![%TABLE_DTD; [
<!-- DTD OASIS Open XML Exchange Table Model.
      Maintained by OASIS; download from:
      http://oasis-open.org/specs/soextblx.dtd
      Note that the FPI in soextblx.dtd refers to itself as 'calstblx'.
      That convention has been followed here.
-->
<!-- create content for title element in table -->
<!ENTITY % title "<!ELEMENT title (#PCDATA | b | i | u | sup | sub | smallcaps)* > ">
%title;
<!--override OASIS Exchange <entry> model -->
<!ENTITY % tbl.entry.mdl "(#PCDATA | b | i | u | sup | sub | smallcaps | br
      | patcit | nplcit | bio-deposit | crossref | figref | img
      | dl | ul | ol | chemistry | maths)* ">

<!--import OASIS Exchange model -->
<!ENTITY % calstblx PUBLIC "-//OASIS//DTD XML Exchange Table Model 19990315//EN"
      "soextblx.dtd">
<!ENTITY % yesorno "NMTOKEN" >
<!ENTITY % tbl.table.att          " pgwide          %yesorno;          #IMPLIED
      orient          (port | land)      #IMPLIED
      tabstyle       NMTOKEN            #IMPLIED">

%calstblx;
]]>
<![%MATH_PLACEHOLDER; [
<!--(PLACEHOLDER:w3c math dtd)-->
<!ELEMENT math (#PCDATA)>
]]>
<![%TABLE_PLACEHOLDER; [
<!--(PLACEHOLDER:cals table dtd)-->
<!ELEMENT table (#PCDATA)>
]]>
```

57. La DTD doit normalement énumérer les éléments à partir de l'élément racine et branche par branche, dans l'ordre naturel du document, les attributs suivant immédiatement l'élément auquel ils s'appliquent. Cependant, toute configuration aidant les utilisateurs à comprendre la structure logique du document est acceptable.

58. Comme cela est indiqué ci-dessus, insérer un commentaire immédiatement avant chaque élément figurant dans la DTD lorsque cela est nécessaire pour cerner le sens ou comprendre l'utilisation des éléments. Le commentaire doit contenir des informations suffisantes pour permettre une utilisation appropriée de l'élément ou de l'attribut et renvoyer à une documentation plus complète si nécessaire. Exemples :

```
<!--The problem the invention purports to solve (Rule 5.1(a)(iii))-->
<!ELEMENT tech-problem (heading* , p)+>
<!ATTLIST tech-problem id ID #IMPLIED >
```

59. Les règles applicables aux DTD des éléments communs internationaux s'appliquent aux DTD propres aux offices.

60. Les DTD doivent utiliser les éléments communs internationaux autant que possible et réserver les éléments propres aux offices aux seuls cas dans lesquels la pratique nationale s'écarte sensiblement de la pratique internationale.

SCIT/SDWG/5/9  
Annexe, page 11

61. Lorsque cela est nécessaire, les éléments propres à un office peuvent être insérés dans une DTD qui par ailleurs est conforme à une norme de l'OMPI. Cette DTD modifiée devient une DTD propre à l'office et doit être dénommée en conséquence, c'est-à-dire que l'élément racine, le nom de fichier et l'identifiant public pour une DTD propre à l'office doivent commencer par le code de l'office correspondant prévu dans la norme ST.3.

62. Toute DTD qui contient au minimum un élément propre à un office est, par définition, une DTD propre à un office.

*Conventions relatives aux documents*

63. La création de documents conformément à la présente norme doit présenter les caractéristiques indiquées ci-après.

64. Les documents conformes à la présente norme sont bien formés et respectent la norme XML 1.1.

65. Il est recommandé que les documents soient conformes à la convention sur le nommage de fichiers présentée dans l'annexe F (et à l'adresse <http://www.wipo.int/scit/fr/standards/index.htm>) pour le dépôt et le traitement. Par exemple, la liste de fichiers suivante correspond à un brevet américain publié contenant sept pages d'images.

```
US06282717-20010904.xml
US06282717-20010904-D00000.TIF
US06282717-20010904-D00001.TIF
US06282717-20010904-D00002.TIF
US06282717-20010904-D00003.TIF
US06282717-20010904-D00004.TIF
US06282717-20010904-D00005.TIF
US06282717-20010904-D00006.TIF
```

66. Pour les demandes internationales, chaque office récepteur du PCT doit créer des documents conformes à l'intégralité des spécifications de l'annexe F.

67. Les documents XML commencent par une section prologue. Le prologue doit contenir une instruction de traitement XML qui précise la version de la norme XML et le schéma de codage. Il doit aussi contenir une déclaration de type de document, également appelée déclaration DOCTYPE. La déclaration de type de document peut éventuellement contenir un identifiant public (introduit par le mot clé PUBLIC) et doit obligatoirement contenir un identifiant système (introduit par le mot clé SYSTEM, à moins que le mot clé PUBLIC ait été utilisé), qui indique un URI (Uniform Resource Identifier) mentionnant la DTD par rapport à laquelle le document peut être validé. Les identifiants publics peuvent se trouver dans les commentaires figurant au début de chaque fichier de DTD.

Exemple - prologue avec identifiant public :

```
<?xml version="1.1" encoding="UTF-8"?>
<!DOCTYPE request PUBLIC "-//WIPO//DTD REQUEST 1.1//EN" "request-v1-1.dtd">
```

Exemple - prologue avec identifiant système :

```
<?xml version="1.1" encoding="UTF-8"?>
<!DOCTYPE request SYSTEM "request-v1-1.dtd">
```

68. Chaque document indique de façon non ambiguë la DTD à laquelle il correspond, sans exception. Cette déclaration n'impose aucune contrainte en ce qui concerne le traitement du document par un office.

Exemples d'éléments racines avec attribut de version :

```
<request dtd-version="request-v1-1.dtd">
<request dtd-version="1.1">
```

Exemples de prologues avec déclaration DOCTYPE :

```
<?xml version="1.1" encoding="UTF-8"?>
<!DOCTYPE request SYSTEM "request-v1-1.dtd" >
```

```
<?xml version="1.1" encoding="UTF-8"?>
<!DOCTYPE request PUBLIC "-//WIPO//DTD REQUEST 1.1 2003-06-02//EN"
SYSTEM "request-v1-1.dtd" >
```

69. À titre facultatif, deux instructions de traitement dans chaque document devraient indiquer le logiciel utilisé pour créer le document, l'une portant sur le nom du logiciel et l'autre sur la version. Ces instructions de traitement apparaissent dans le prologue du document. Voir <http://www.w3.org/TR/2000/REC-xml-20001006#sec-pi> pour une description de la syntaxe utilisée dans les instructions de traitement.

```
<?software_name [name of the software]?>  
<?software_version [version of the software]?>
```

70. Si des feuilles de style peuvent être utilisées, la DTD ou chaque document conforme à la DTD indique les feuilles de style, conformément à la recommandation du W3C *Recommendation Associating Style Sheets with XML documents* (recommandation sur l'association de feuilles de style et de documents XML version 1.0) publiée à l'adresse <http://www.w3.org/TR/xml-stylesheet/>.

Exemple de déclaration d'une feuille de style :

```
<?xml-stylesheet href="..\..\stylesheet_factory\pap.xsl" type="text/xsl"?>
```

71. L'organisation des espaces dans les documents peut largement contribuer à la lisibilité des données brutes. Par exemple, les documents peuvent contenir un saut de ligne après la balise de clôture de chaque paragraphe, de sorte que chaque paragraphe commence sur une nouvelle ligne.

72. Les feuilles de style s'appuient sur des hypothèses concernant le contenu situé entre les balises d'ouverture et de clôture. Bien que les processeurs XML comprennent les espaces dans certaines circonstances, d'autres logiciels utilisés pour traiter les documents peuvent fonctionner différemment. Éviter d'ajouter des espaces non pertinents qui ne font pas partie du contenu. Le contenu doit commencer immédiatement après la balise d'ouverture et s'achever juste avant la balise de clôture.

Par exemple,

```
<postcode>20231</postcode>
```

et non

```
<postcode> 20231 </postcode>.
```

#### *Entités externes*

73. Une entité externe est un élément qui accompagne un document XML et qui est appelé dans le document. Les entités externes font partie intégrante du document de brevet. Sans elles, le document XML ne peut pas être analysé, présenté ou compris de façon satisfaisante.

74. En ce qui concerne les documents de brevet, les entités externes sont le plus souvent des pages de dessin, mais elles peuvent aussi contenir des images incrustées, des listages informatiques, des tableaux, des listages de séquences, des caractères non définis ou des entités de caractères. Les images incrustées sont des entités externes couramment utilisées : un renvoi à un fichier image externe est introduit dans un document à l'endroit où l'image doit apparaître lors de la présentation du document.

75. Dans les documents de brevet, les images peuvent se présenter comme des pages entières lues au scanner ou comme des images dites incrustées. Les images incrustées forment le plus souvent des parties de document qui ne peuvent être codées ni stockées à l'aide d'un jeu de caractères. Il peut s'agir de dessins, de formules chimiques, de tableaux complexes, de caractères non définis, etc. On entend par caractères non définis des caractères qui ne sont pas prévus dans un jeu de caractères ou qui n'existent pas en tant qu'appels d'entité de caractères (voir ci-dessus). À l'heure actuelle, l'annexe F prévoit quatre types de données images pour les éléments communs internationaux <img>: TIFF, JPEG, ST.33 et ST.35; ces quatre catégories sont examinées ci-dessous.

76. Les sous-sections ci-après donnent des indications sur l'utilisation des types de fichiers images prévus dans l'annexe F. Si un office décide de publier des documents qui contiennent des renvois à des entités externes utilisant d'autres codages, ou si l'office s'écarte des recommandations figurant ci-dessous, il doit fournir aux utilisateurs des informations suffisantes pour une divulgation satisfaisante des documents.

#### TIFF

77. Les images incrustées figurent sous un en-tête TIFF (tagged image file format) et sont conformes au profil publié à l'adresse <http://www.wipo.int/scit/fr/standards/index.htm>. Pour une explication complète du format TIFF, voir <http://partners.adobe.com/asn/developer/pdfs/tn/TIFF6.pdf>.

78. Le schéma de codage recommandé pour les données images TIFF est fondé sur la technique de compression de données READ II modifié pour les télécopieurs UIT-T (CCITT) du groupe 4 telle qu'elle est décrite dans la recommandation T.6 de l'UIT-T (CCITT) relative aux télécopieurs du groupe 4. Pour plus de détails sur cette recommandation et le contenu éventuel de l'information figurant dans l'en-tête TIFF, voir les annexes 3 et 4 de la norme ST.35.

#### JPEG

79. Les images incrustées peuvent aussi figurer dans un en-tête JPEG et correspondre au profil publié à l'adresse <http://www.wipo.int/scit/fr/standards/index.htm>. Pour une explication complète de la norme JPEG, voir <http://www.jpeg.org/>.

80. Si elle est utilisée, la norme JPEG doit être conforme aux règles propres aux offices. Par exemple, le PCT n'autorise que les images en noir et blanc (règle 11.13 du PCT : "Les dessins doivent être exécutés en lignes et traits durables, noirs, suffisamment denses et foncés, uniformément épais et bien délimités, sans couleurs ni lavis"). Si un autre office autorise d'autres options en JPEG, ces options doivent être précisées et publiées.

#### Norme ST.33 de l'OMPI

81. Lors du formatage des fichiers images externes conformément à la norme ST.33, l'office doit renvoyer les utilisateurs à la norme ST.33 de l'OMPI intitulée "Format recommandé pour l'échange de documents de brevet sous forme de "fac-similés"", disponible à l'adresse <http://www.wipo.int/scit/fr/standards/standards.htm>.

#### Norme ST.35 de l'OMPI

82. Lorsqu'un office formate des fichiers images externes conformément à la norme ST.35, il doit renvoyer les utilisateurs à la norme ST.35 de l'OMPI intitulée "Format recommandé pour l'échange d'information sur les documents de brevet publiés, enregistrée en mode mixte sur bande en bobine ou en cartouches IBM 3480/90", disponible à l'adresse <http://www.wipo.int/scit/fr/standards/standards.htm>.

83. La présente norme ne porte que sur l'utilisation des données images de la norme ST.35, pas sur les autres types de données mentionnées dans cette norme.

#### PDF

84. Les entités externes qui sont des fichiers PDF doivent correspondre au profil publié à l'adresse <http://www.wipo.int/scit/fr/standards/index.htm>. Pour de plus amples informations sur le format PDF, voir <http://www.adobe.com/products/acrobat/adobe.pdf.html>.

85. Si un office décide de publier des documents qui appellent des entités externes utilisant des codages exclusifs, il doit au minimum fournir des informations suffisantes pour une présentation satisfaisante des documents.

#### MÉGACONTENU

86. Certaines demandes de brevet et les publications qui en découlent ont un contenu si volumineux (ci-après dénommé "mégacontenu") qu'il empêche ou entrave considérablement le traitement informatique habituel. Une méthode adoptée par certains offices pour surmonter cette difficulté consiste à traiter le mégacontenu comme une entité externe. Cette section décrit la façon de mettre en œuvre cette méthode pour différents types de mégacontenus.

87. En tant qu'entité externe, le mégacontenu reste une partie intégrante du document (demande telle qu'elle a été déposée ou publication), mais il se trouve dans un fichier distinct qui peut être ignoré pour certains types de traitement ou traité séparément à partir du document dont il fait partie. Habituellement, une entité externe est appelée dans un document à l'endroit où elle figurerait normalement lors de la présentation, c'est-à-dire à sa position initiale logique dans le document. Cette section donne des indications aux offices qui décident de traiter les mégacontenus comme des entités externes.

88. Lorsqu'une partie du contenu, par exemple un listage de séquence ou un tableau, dépasse une limite fixée par un office de propriété industrielle, celui-ci peut exiger que le mégacontenu soit traité comme une entité externe. Par exemple, un office peut exiger que les tableaux, qui occupent au moins 300 pages imprimées, soient traités comme une entité externe. Publié comme une entité externe, un grand tableau peut être plus facilement exclu de l'impression automatique, ce qui évite une utilisation de papier imprévue et excessive par les examinateurs et le public non averti. Il peut aussi être exclu du téléchargement, à moins qu'un utilisateur ou un examinateur ne le demande expressément, afin d'éviter l'encombrement des réseaux et des serveurs.

89. En ce qui concerne les listages de séquences, la norme ST.25 définit le format de l'entité externe. Pour la norme ST.25, voir <http://www.wipo.int/scit/fr/standards/index.htm>.

90. En ce qui concerne les tableaux, le format XML utilisé pour les tableaux dans un document est aussi utilisé pour les entités externes. Une DTD de tableau couramment utilisée dans l'industrie et légèrement modifiée à cette fin est : `table-external.dtd`.

91. En ce qui concerne les listages informatiques, le format de l'entité externe est généralement un texte ASCII simple, le formatage étant déterminé par la syntaxe du langage du programme informatique. Pour cela, il faudra placer la liste de programme dans un tableau à une cellule appelé dans `table-external-doc` dans une DTD `table-external.dtd`.

92. En ce qui concerne les structures chimiques ou les formules mathématiques, dans lesquelles le mégacontenu doit contenir normalement de nombreux éléments par ailleurs relativement petits, le contenu est introduit dans un ou plusieurs tableaux et les tableaux qui en découlent sont donc traités comme des entités externes, au moyen de la DTD `table-external.dtd`.

#### *DTD couramment utilisées dans l'industrie*

93. Lorsque cela convient au contenu du document, c'est-à-dire lorsque le contenu n'est pas propre au domaine de la propriété industrielle, utiliser des DTD courantes dans l'industrie. Ce ne sont pas des DTD internationales ou propres aux offices; elles sont incorporées par renvoi (par exemple `application-body.dtd`).

94. Pour les formules mathématiques, utiliser MathML, version 2. Voir <http://www.w3.org/Math/> pour un exposé complet.

95. Pour les tableaux, utiliser la DTD des tableaux OASIS XML. Voir <http://www.oasis-open.org/specs/tm9901.html>, pour accéder au memorandum technique d'OASIS TR 9901:1999, *XML Exchange Table Model Document Type Definition*.

96. Lorsqu'un office souhaite utiliser d'autres DTD courantes dans l'industrie, une demande de modification doit être présentée.

97. Il est recommandé de n'incorporer les DTD couramment utilisées dans l'industrie que par renvoi.

98. Les DTD couramment utilisées dans l'industrie ne peuvent être modifiées en aucune façon lorsqu'elles sont appelées dans des DTD propres aux offices. Il existe toutefois une exception : une intervention ponctuelle pour ajouter des attributs à l'élément racine de la DTD des tableaux OASIS dans la DTD `xx-patent-publication`. La DTD des tableaux OASIS est créée expressément pour permettre la définition locale du contenu de la cellule. D'autres exceptions peuvent être présentées sous forme de demandes de modification.

#### *DTD type pour les publications de brevet*

99. La DTD type `xx-patent-document.dtd` n'est pas destinée à être utilisée sous sa forme publiée. Elle est censée constituer une base à partir de laquelle chaque office peut créer sa propre DTD avec un minimum d'efforts et en utilisant au mieux les éléments communs et les structures logiques de haut niveau.

100. Les règles applicables aux noms propres aux offices qui sont indiquées ci-dessus doivent être observées lorsqu'on modifie la DTD type pour les publications en matière de brevets (`xx-patent-document.dtd`). Les offices sont incités à ne pas apporter de modifications qui ne sont pas essentielles au but poursuivi, afin de ne pas nuire à l'interfonctionnement.

101. Les contraintes propres aux offices qui ne sont pas exprimées dans les DTD visées dans la présente norme doivent être exprimées dans une autre spécification publique propre à l'office.

#### RÉFÉRENCES

102. Les normes et documents ci-après présentent un intérêt en ce qui concerne la présente norme :

- a) Norme ST.3 de l'OMPI - Norme recommandée concernant les codes à deux lettres pour la représentation des États, autres entités et organisations intergouvernementales.
- b) Norme ST.9 de l'OMPI - Recommandation concernant les données bibliographiques qui figurent sur les brevets ou qui se rapportent aux brevets ou aux CCP.
- c) Norme ST.16. de l'OMPI - Code normalisé recommandé pour l'identification de différents types de documents de brevet.
- d) Norme ST.25 de l'OMPI - Norme relative à la présentation du listage des séquences de nucléotides et d'acides aminés dans les demandes de brevet.

SCIT/SDWG/5/9  
Annexe, page 15

- e) Norme ST.32 de l'OMPI - Recommandation concernant le balisage des documents de brevet selon le SGML (langage normalisé de balisage généralisé).
- f) Norme ST.33 de l'OMPI – Norme recommandée pour l'échange de documents de brevet sous forme de "fac-similés".
- g) Norme ST.35 de l'OMPI - Format recommandé pour l'échange d'information sur les documents de brevet publiés, enregistrée en mode mixte sur bande en bobine ou en cartouches IBM 3480/90.
- h) Instructions administratives du Traité de coopération en matière de brevets : septième partie, Instructions relatives au dépôt et au traitement électroniques des demandes internationales.
- i) Tim Bray, Jean Paoli, C. M. Sperberg-McQueen, Eve Maler. Extensible Markup Language (XML) 1.1. Recommandation du Consortium W3C du 4 février 2004, éditée au sein du W3C le 15 avril 2004. Voir : <http://www.w3.org/TR/2004/REC-xml11-20040204>.
- j) Norme internationale ISO/IEC 10646-1:1993(E) : Information Technology – Universal Multiple – Octet Coded Character Set (UCS) – Part 1 : Architecture and Basic Multilingual Plane. Organisation internationale de normalisation, Genève, 1993.
- k) Norme internationale ISO/IEC 10646-2, Information Technology – Universal Multiple – Octet Coded Character Set (UCS) – Part 2 : Supplementary Planes. Première édition, Organisation internationale de normalisation, Genève, 2001.

[Fin de l'annexe et du document]