

w.c.s. - Development #78281

ExportToModel: perte des namespaces XML dans les documents lors de la transformation

09 juin 2023 01:58 - Pierre Ducroquet

Statut:	Fermé	Début:	09 juin 2023
Priorité:	Normal	Echéance:	
Assigné à:	Pierre Ducroquet	% réalisé:	0%
Catégorie:		Temps estimé:	0:00 heure
Version cible:		Planning:	Non
Patch proposé:	Non		
Description			
<p>Le code de génération des documents ne conserve pas l'ensemble des namespaces XML définis dans le document source. Or, LibreOffice peut stocker des éléments référençant les namespaces dans des attributs, ce qui n'est pas pris en compte par les parseurs XML, et provoque donc la perte pure et simple des namespaces en question à la génération.</p> <p>Ce bug est responsable du non-fonctionnement d'au moins deux éléments demandés sur la génération de documents:</p> <ul style="list-style-type: none">- les paragraphes masqués,- les conditions d'affichage sur les sections. <p>Le correctif ne devrait pas être trop complexe, mais impliquera certainement de passer d'ElementTree à un "vrai" parseur XML plus exhaustif comme lxml.</p>			
Demandes liées:			
Lié à w.c.s. - Development #78836: avoir le(s) namespace(s) nécessaire(s) aux...		Fermé	21 juin 2023

Révisions associées

Révision d3b3eba5 - 14 juin 2023 12:24 - Pierre Ducroquet

general: switch to lxml to have a complete XML parser (#78281)

Historique

#1 - 09 juin 2023 07:04 - Robot Gitea

- Statut changé de Nouveau à En cours

Pierre Ducroquet (pducroquet) a ouvert une pull request sur Gitea concernant cette demande :

- URL : <https://git.entrouvert.org/entrouvert/wcs/pulls/362>
- Titre : WIP: export_to_model: switch to lxml to have a complete XML parser (#78281)
- Modifications : <https://git.entrouvert.org/entrouvert/wcs/pulls/362/files>

#2 - 09 juin 2023 11:08 - Pierre Ducroquet

Après vérification dans la norme OpenDocument, ce comportement surprenant y est pourtant clairement indiqué.

http://docs.oasis-open.org/office/v1.2/cs01/OpenDocument-v1.2-cs01-part1.html#_RefHeading_1418812_253892949

The text:condition attribute specifies a condition. Conditions do not have a predefined syntax, but the attribute value should begin with a namespace prefix, followed by a ":" (U+003A, COLON) separator, followed by the text of a formula. The namespace bound to the prefix determines the syntax and semantics of the formula.

Donc avec ElementTree, on risque de casser les documents à cause de cette syntaxe qui est en dehors de la syntaxe XML "normale", seule comprise et prise en compte par ElementTree. Et ElementTree ne permet pas d'ajouter arbitrairement un ensemble de namespaces devant être présents dans le document (choix d'implémentation acceptable, mais du coup problématique ici).

#3 - 09 juin 2023 12:44 - Frédéric Péters

Il y a déjà eu un ticket sur cette affaire, [#66352](#), et bricolage pour ce cas particulier.

À propos de :

- les paragraphes masqués,
- les conditions d'affichage sur les sections.

Il y a un ticket à lier qui exposerait le problème, avec des documents, ou c'est se prémunir d'un problème sur ces éléments ?

#4 - 09 juin 2023 13:59 - Frédéric Péters

Pas de ticket, ça vient de tracim, <https://publik.tracim.fr/ui/workspaces/1/contents/thread/2229/> / <https://publik.tracim.fr/ui/workspaces/1/contents/thread/371>

#5 - 10 juin 2023 16:51 - Pierre Ducroquet

Frédéric Péters a écrit :

Pas de ticket, ça vient de tracim, <https://publik.tracim.fr/ui/workspaces/1/contents/thread/2229/> / <https://publik.tracim.fr/ui/workspaces/1/contents/thread/371>

tracim + villejuif lors du club jeudi (pendant la présentation de la roadmap, et j'étais convaincu que ça devait fonctionner en théorie, puis j'ai vu ce problème avec ElementTree)
D'ailleurs au club utilisateur ça a été mentionné dans les diapositives comme fonctionnalités à ajouter.

#6 - 11 juin 2023 09:04 - Robot Gitea

- Statut changé de *En cours* à *Solution proposée*

#7 - 12 juin 2023 10:27 - Robot Gitea

- Statut changé de *Solution proposée* à *En cours*

Frédéric Péters (fpeters) a relu et demandé des modifications sur une pull request sur Gitea concernant cette demande :

- URL : <https://git.entrouvert.org/entrouvert/wcs/pulls/362>

#8 - 12 juin 2023 10:37 - Robot Gitea

- Statut changé de *En cours* à *Solution proposée*

#9 - 12 juin 2023 10:38 - Robot Gitea

- Statut changé de *Solution proposée* à *Solution validée*

#10 - 12 juin 2023 10:39 - Thomas Noël

- Statut changé de *Solution validée* à *En cours*

Robot Gitea a écrit :

Thomas NOËL (tnoel) a approuvé une pull request

Non, erreur de manip.

#11 - 14 juin 2023 11:24 - Robot Gitea

- Statut changé de *En cours* à *Solution proposée*

#12 - 14 juin 2023 12:17 - Robot Gitea

- Statut changé de *Solution proposée* à *Solution validée*

Frédéric Péters (fpeters) a approuvé une pull request sur Gitea concernant cette demande :

- URL : <https://git.entrouvert.org/entrouvert/wcs/pulls/362>

#13 - 14 juin 2023 13:07 - Robot Gitea

- Statut changé de *Solution validée* à *Résolu (à déployer)*

Pierre Ducroquet (pducroquet) a mergé une pull request sur Gitea concernant cette demande :

- URL : <https://git.entrouvert.org/entrouvert/wcs/pulls/362>
- Titre : switch to lxml to have a complete XML parser (#78281)
- Modifications : <https://git.entrouvert.org/entrouvert/wcs/pulls/362/files>

#14 - 14 juin 2023 18:14 - Transition automatique

- Statut changé de *Résolu (à déployer)* à *Solution déployée*

#15 - 21 juin 2023 21:04 - Frédéric Péters

- Lié à Development #78836: avoir le(s) namespace(s) nécessaire(s) aux conditions sur sections & paragraphes masqués lors de la génération d'odt ajouté

#16 - 20 août 2023 04:42 - Transition automatique

Automatic expiration