

TIRÈME SARL

OASIS Tables, XML Exchange Table Model Document Type Definition

Rédaction : Pierre Attar

Issu du Département de la défense américaine et repris ensuite par toute la communauté SGML, le modèle *OASIS Tables* est maintenant adapté à XML. Il est principalement utilisé chez les éditeurs et pour la documentation technique. Sa structure est peu différente de celle des tableaux HTML, qui s'en est beaucoup inspiré.

Son avenir ? Si les notions de colonnes (`col` et `colgroup`) et de têtes (`thead` et `tfoot`) sont implémentées dans les outils XML, ce modèle pourra sans problème disparaître au profit du modèle de HTML. Dans l'attente, les bases de documents structurées avec *OASIS Tables* pourront continuer à exister, la conversion vers HTML étant triviale.

Objectifs

Le tableau fait partie des constructeurs indispensables pour augmenter la capacité d'expression du discours écrit : support à la communication. Son intérêt est de mettre en relation des idées dans un but de clarté et de concision. Discursif, le tableau s'appuie sur une *spatialisation* de l'information sur un espace bidimensionnel constitué par la page (ou par toute autre sorte de "fenêtre"), dès lors que l'information est consultée de façon électronique. Les idées émises pèsent alors le poids relatif que leur confère leur position relative dans le tableau.

Dans le discours écrit, le tableau apparaît toujours juxtaposé au texte il trouve sa place entre des paragraphes. Cette juxtaposition est spécifique et ne relève pas de l'illustration (comme l'image ou le graphique) ; elle relève beaucoup plus d'une coopération avec le mode textuel. Le tableau n'est pas autosuffisant ; il nécessite très souvent le mode textuel pour expliquer les données que lui-même présente.

Typologie des structures tableaux

D'un point de vue pratique, l'utilisation d'une structuration de l'information sous forme de tableau(x) répond à différents objectifs :

ordonnement : le tableau est un mode de classement. Les fameuses structures de listes unidimensionnelles, qui hiérarchisent entre elles des séries d'informations, en sont une illustration ;

représentation graphique de données et, plus généralement, de faits : le tableau n'est alors qu'une représentation graphique d'un ensemble d'informations par elle-mêmes organisées selon un modèle sous-jacent. La valeur du tableau résidera, de fait, dans sa capacité de présentation. Des outils graphiques évolués seront nécessaires pour permettre à l'utilisateur de manipuler directement la sémantique des données sous-jacentes :

organisations complexes d'informations : ces tableaux "libres" ne sont pas structurés par leurs données, mais plutôt par le sens relatif que donne le concepteur aux différents composants. Les tableaux mettent en regard des éléments d'information ;

outil de mise en page : "libre", là encore, le tableau est devenu et est utilisé sur les supports de consultation électroniques comme l'outil privilégié de toute mise en page. Il pallie l'absence de définition de ce que peut être une page électronique.

Les tableaux peuvent donc être utilisés, d'une part, pour représenter, de façon ergonomique et lisible, des données fortement structurées par elles-mêmes et, d'autre part, pour supporter un mode discursif de l'expression écrite. En conséquence, il convient d'adapter la terminologie pour différencier ces modes d'usage relatifs aux "tableaux à structure de contenu" de ceux à "structure libre".

Le tableau a parfois une valeur graphique intrinsèque qui ne dépend pas d'un quelconque modèle de données, mais beaucoup plus des intentions de l'auteur. Cela ne signifie pas pour autant que toutes les sémantiques sous-jacentes, si elles existent, ne soient pas utiles. Cela veut simplement dire qu'il est nécessaire de prendre en compte, dans les formats de codage textuels, cette double nature des sources d'information représentées par les tableaux. Cette prise en compte pourra alors être réalisée au niveau du codage des données elles-mêmes - pour les structures "libres" - ou au niveau des spécifications de présentation des données, pour les "structures de contenu".

Avec ses feuilles de styles **FOSI**, la communauté **CALS** a très bien compris cette idée de double nature de l'information tabulaire. Le même concept est également repris dans **XSL** et l'avait été, auparavant, dans **DSSSL**. En effet, grâce à ces technologies, des constructeurs graphiques génériques de création de tableaux sont utilisables lors de la définition des feuilles de styles. Ils s'appuient sur les notions de colonnes, de lignes et de cellules. Charge alors au concepteur d'une feuille de styles particulière d'utiliser ces constructeurs pour "mettre en tableau" des structures de contenu ou pour les appliquer à des structures "libres", saisies sous forme de tableau dès l'origine.

Fonctionnalités et formats de codage des tableaux à structure "libre"

Tous les modèles utilisés par les communautés **SGML** puis **XML** ont un ensemble de caractéristiques communes. Leur objectif est toujours de proposer un codage de données permettant d'émuler les informations nécessaires à un système générique de composition de tableaux. Les fonctionnalités sont essentiellement graphiques, calquées sur la façon dont sont actuellement utilisés les tableaux sur un support exclusivement papier. Parmi celles-ci, on distingue :

- mise en colonne, ligne et cellule,
- graphisme d'encadrement du tableau,
- filets autour d'une cellule ou d'un ensemble de cellules,
- prise en compte des recouvrements horizontaux et verticaux,
- prise en compte d'en-tête et de pied de tableau,
- prise en compte des alignements horizontaux et verticaux à l'intérieur d'une cellule,
- prise en compte des couleurs de fonds de cellules.

D'un point de vue plus structurel, un tableau comporte un titre et un ensemble de sous-tableaux, pour définir un tableau de tête, un corps et, parfois, un tableau de fin.

Un sous-tableau comporte :

un ensemble de définitions de colonnes : celles-ci n'apparaissent pas en tant que telles dans le tableau présenté. Elles permettent de donner au système de formatage des informations sur le nombre et la taille de chaque colonne. À titre indicatif, seul le modèle proposé par la société GRIF, dans son éditeur/composeur [SGML](#) prenait en compte, de façon complète, la notion de colonne. Dans les autres modèles, la colonne ne sert que de référence pour définir la position d'une cellule ;

des lignes contenant des cellules : pour une cellule, il est possible de définir des recouvrements verticaux ou horizontaux ("spanning" en anglais), des filets et des alignements horizontaux ou verticaux.

Positionnements de cellules dans un tableau "libre"

Cellule simple	Cellule simple
cellule s'étendant sur trois colonnes en largeur	
cellule s'étendant sur deux lignes et deux colonnes, en hauteur et largeur	
Cellule s'étendant sur deux lignes en hauteur	

Le format de tableaux OASIS Open

Le modèle du Consortium est une simplification et une amélioration du modèle CALS. Son propos est de rendre ce dernier modèle plus adéquat à l'échange d'informations documentaires saisies sous forme tabulaires, tout en garantissant la représentation des données saisies avec différents outils en environnement hétérogène. Proposé début 1996 pour [SGML](#), il existe aujourd'hui adapté à [XML](#). Son utilisation est généralisée dans toute l'industrie et dans beaucoup de secteurs de l'édition.

Principes

Le modèle de tableaux proposé est actuellement certainement le modèle le plus utilisé par la communauté [SGML](#) et, de façon plus générale, dans les applications traitant de documents structurés [XML](#) (bon nombre d'applications [SGML](#) ont maintenant migré sur [XML](#)). Il est reconnu et repris dans toute la documentation technique du Département de la défense américaine, mais aussi par beaucoup de [DTD](#) industrielles sectorielles, comme celles de l'aéronautique civile (DTD ATA – Air Transport Association), de l'aéronautique militaire européenne (AECMA), des constructeurs de matériels informatiques (DocBook), de l'industrie automobile (J2008), des constructeurs de matériaux semi-conducteurs (Pinnacles) ou encore des télécommunications (Telecom/Technical Interchange Markup – TIM). Les éditeurs et, notamment, les éditeurs de logiciels, lorsqu'ils intègrent un modèle de tableau, utilisent très souvent ce modèle.

Ces tableaux "libres" sont bidimensionnels et contiennent des lignes et des colonnes. Des règles de formatage de base sont proposées à tous les niveaux : colonnes, lignes et cellules, ensemble du tableau. Ces règles s'intéressent aussi bien aux contenus (alignements) qu'aux filets d'encadrement.

D'un point de vue structurel, la structure proposée est hiérarchisée. Une table contient un titre et le contenu du tableau lui-même :

- *spécifications de colonne* (`colspec`) : le modèle présuppose un mécanisme d'héritage de styles et, plus généralement, de présentation. Ainsi, les spécifications de colonne s'appliqueront par défaut à l'ensemble des trois sous-tableaux, sauf spécification contradictoire à un niveau inférieur ;

- *tête de tableau* (`tthead`) : la tête du tableau est identifiée de façon séparée du corps du tableau. Cette identification permet, sur papier, une présentation différenciée des contenus. Sur un support électronique, elle peut aussi être traitée de façon différenciée pour, par exemple, rester présente dans une sous-fenêtre de consultation du tableau ;

- *corps de tableau* (`tbody`) : le corps du tableau contient des données. Celles-ci sont organisées sous forme de lignes contenant des cellules. Le corps est composé de lignes (`row`) qui elles-mêmes contiennent des cellules (`entry`).

D'un point de vue graphique, des fonctionnalités sont proposées, au niveau des cellules, pour mettre en place :

- *des recouvrements* (utilisation de `namest` et `nameend`, en conjonction avec les définitions de colonnes) : ces chevauchements peuvent être horizontaux ou verticaux ;

- *des filets* (`rowsep` et `colsep`) : ce sont des filets encadrant l'ensemble du tableau ou une cellule seulement ;

- *des alignements* : à l'intérieur d'une cellule, des alignements verticaux ou horizontaux peuvent être mis en œuvre.

À de rares exceptions près, ces fonctionnalités sont largement suffisantes pour la plupart des documentations techniques et, plus généralement, pour les tableaux à structure libre. C'est ce qui participe à l'acceptation généralisée de ce modèle. En revanche, les utilisateurs d'**HTML** reprocheront à ce modèle la pauvreté des types de filets et des couleurs de fonds... ceci est certainement lié à cet héritage qui nous vient du monde de la documentation technique à qui l'on doit la création de ce modèle. Quoi qu'il en soit, ce modèle ne se définit pas comme étant un modèle de création de structures de pages : seulement de "tableaux libres".

Recommandations(s)

-  *XML Exchange Table Model Document Type Definition - OASIS Technical Memorandum*

Recommandation, version TR 9901:1999, du 29-09-1999
Document sur <http://www.oasis-open.org/html/tm9901.htm>

-  *Exchange Table Model Document Type Definition - OASIS Technical Resolution*

Recommandation, version TR 9503:1995, du 08-05-1996
Document sur <http://www.oasis-open.org/html/a503.htm>

-  *CALS Table Model Document Type Definition - OASIS Technical Memorandum*

Recommandation, version TM 9502:1995, du 19-10-1995
Document sur <http://www.oasis-open.org/html/a502.htm>