

DTD (Document Type Definition)

16 janvier 2014

emeline.dorey@gmail.com

Sommaire

DTD (Document
Type Definition)

Introduction

DTD interne

DTD externe

Déclarer les
éléments

Indicateurs
d'occurrence

Déclarer les
attributs

Limites des DTD

- 1 Introduction
- 2 DTD interne
- 3 DTD externe
- 4 Déclarer les éléments
- 5 Indicateurs d'occurrence
- 6 Déclarer les attributs
- 7 Limites des DTD

Introduction

Introduction

DTD (Document
Type Definition)

Introduction

DTD interne

DTD externe

Déclarer les
éléments

Indicateurs
d'occurrence

Déclarer les
attributs

Limites des DTD

- Une définition de document type, appelée DTD (Document Type Definition), permet de décrire la structure d'un document SGML et XML.
- XML permet d'utiliser un fichier qui vérifie qu'un document XML est conforme à une syntaxe donnée.
- Lorsque l'on écrit un document XML qui respecte ses règles de syntaxe, on obtient un document XML "bien formé".
- Dans ce cours, nous allons voir comment spécifier des contraintes propres à nos documents XML, grâce aux DTD.

Introduction

Terminologie

DTD (Document
Type Definition)

Introduction

DTD interne

DTD externe

Déclarer les
éléments

Indicateurs
d'occurrence

Déclarer les
attributs

Limites des DTD

- Un **document** est “**bien formé**” lorsqu'un document XML répond aux règles de base du XML et ne comporte pas de DTD.
- Un **document** est “**valide**” par rapport à une DTD si ce document XML est bien formé et conforme à cette DTD.

Introduction

Intérêt

DTD (Document
Type Definition)

Introduction

DTD interne

DTD externe

Déclarer les
éléments

Indicateurs
d'occurrence

Déclarer les
attributs

Limites des DTD

- La DTD permet de spécifier une grammaire pour un langage et de tester automatiquement son respect par un document donné.
- L'avantage est de :
 - faciliter l'échange et la mise en commun de documents produits par différentes personnes ;
 - aider les développeurs qui conçoivent des outils automatiques pour traiter les documents respectant la même DTD.

Introduction

Définition

DTD (Document
Type Definition)

Introduction

DTD interne

DTD externe

Déclarer les
éléments

Indicateurs
d'occurrence

Déclarer les
attributs

Limites des DTD

- Une DTD définit la structure d'un document :
 - les éléments et les attributs qui y sont autorisés ;
 - les types de contenus qui y sont permis.
- Elle permet de faire la différence entre un document bien formé et un document valide.
- Le document bien formé répond aux exigences de la spécification.
- Le document valide se conforme strictement aux règles établies par la DTD à laquelle il fait référence.

Introduction

Définition

DTD (Document
Type Definition)

Introduction

DTD interne

DTD externe

Déclarer les
éléments

Indicateurs
d'occurrence

Déclarer les
attributs

Limites des DTD

- Une DTD peut être définie de deux manières :
 - **interne** : la grammaire est incorporée au sein même du document ;
 - **externe** : la DTD est un fichier à part. Elle est appelée depuis le document XML.
- Le stockage externe de la DTD permet de la partager entre différents documents XML.
- On peut mêler DTD interne et externe.

DTD interne

Présentation

DTD (Document
Type Definition)

Introduction

DTD interne

DTD externe

Déclarer les
éléments

Indicateurs
d'occurrence

Déclarer les
attributs

Limites des DTD

- Entre une DTD interne et une DTD externe, le contenu ne change pas.
- Les déclarations d'une DTD interne sont écrites juste après le prologue du document XML.
- La spécification de la DTD sera dans l'entête du document XML.
- Les déclarations doivent être faites dans l'ordre :
 - mot-clé **DOCTYPE** ;
 - élément racine du document ;
 - contenu de la DTD elle-même, entre crochets.

DTD interne

Exemple de déclaration

DTD (Document
Type Definition)

Introduction

DTD interne

DTD externe

Déclarer les
éléments

Indicateurs
d'occurrence

Déclarer les
attributs

Limites des DTD

```
<?xml version="1.0" encoding="ISO-8859-1" standalone="yes"?>
<!DOCTYPE bibliotheque[
  <!ELEMENT bibliotheque (livre)*>
  <!ELEMENT livre (titre, auteur)>
  <!ELEMENT titre (#PCDATA)>
  <!ELEMENT auteur (#PCDATA)>
]>
<bibliotheque>
  <livre>
    <titre>XML : Cours et exercices</titre>
    <auteur>Alexandre Brilliant</auteur>
  </livre>
</bibliotheque>
```

DTD interne

Exemple de déclaration

DTD (Document
Type Definition)

Introduction

DTD interne

DTD externe

Déclarer les
éléments

Indicateurs
d'occurrence

Déclarer les
attributs

Limites des DTD

Attention !

Si la DTD est interne, il faut déclarer dans le prologue du fichier XML *standalone="yes"*.

Rappel

Dans le prologue du fichier XML, si :

- **standalone="yes"**, le document est auto-contenu ⇒ la DTD est interne.
- **standalone="no"**, le document n'est pas auto-contenu ⇒ la DTD est externe. Déclaration par défaut ;

DTD externe

Présentation

DTD (Document
Type Definition)

Introduction

DTD interne

DTD externe

Déclarer les
éléments

Indicateurs
d'occurrence

Déclarer les
attributs

Limites des DTD

- Une DTD est un document texte avec l'extension **.dtd**, ce n'est pas un document XML \Rightarrow ne commence pas par une déclaration XML.

DTD externe

Présentation

DTD (Document
Type Definition)

Introduction

DTD interne

DTD externe

Déclarer les
éléments

Indicateurs
d'occurrence

Déclarer les
attributs

Limites des DTD

- Les DTD externes peuvent être privées ou publiques.
- Les DTD privées sont :
 - accessibles uniquement en local, sur la machine de développement ;
 - de type **SYSTEM** ;
- Les DTD publiques sont :
 - disponibles pour tout le monde, sur un serveur distant. On y accède grâce à une URI (Uniform Resource Identifier) ;
 - de type **PUBLIC** ;

DTD externe

Exemple de déclaration

DTD (Document
Type Definition)

Introduction

DTD interne

DTD externe

Déclarer les
éléments

Indicateurs
d'occurrence

Déclarer les
attributs

Limites des DTD

Exemple de déclaration de DTD externe privée, à placer dans le fichier XML après le prologue :

```
<!DOCTYPE bibliotheque SYSTEM "bibliotheque.dtd">
```

Exemple de déclaration de DTD externe publique :

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.1//EN" "  
http://www.w3.org/TR/xhtml11/DTD/xhtml11.dtd">
```

DTD externe

Exemple de déclaration

DTD (Document
Type Definition)

Introduction

DTD interne

DTD externe

Déclarer les
éléments

Indicateurs
d'occurrence

Déclarer les
attributs

Limites des DTD

Exemple de contenu pour le fichier "bibliotheque.dtd" :

```
<!ELEMENT bibliotheque (livre)*>  
<!ELEMENT livre (titre, auteur)>  
<!ELEMENT titre (#PCDATA)>  
<!ELEMENT auteur (#PCDATA)>
```

Déclarer les éléments

Présentation

DTD (Document
Type Definition)

Introduction

DTD interne

DTD externe

Déclarer les
éléments

Indicateurs
d'occurrence

Déclarer les
attributs

Limites des DTD

Modèle de déclaration d'un élément dans une DTD :

```
<!ELEMENT nom contenu_element>
```

- *nom* : nom de l'élément \Rightarrow nom d'une balise dans le fichier XML associé ;
- *contenu_element* : type auquel il est associé. Les valeurs possibles sont :
 - texte \Rightarrow **(#PCDATA)**
 - vide \Rightarrow **EMPTY**
 - séquence d'éléments \Rightarrow **(elt1, elt2, elt3, ...)**
 - choix d'éléments \Rightarrow **(elt1 | elt2 | elt3 | ...)**
 - libre \Rightarrow **ANY**
 - mixte \Rightarrow mélange de texte et d'éléments (enfants)

Déclarer les éléments

Élément texte

DTD (Document
Type Definition)

```
<!ELEMENT nom (#PCDATA)>
```

- Cet élément doit contenir du texte.

Recommandation

Pour éviter les éventuelles erreurs du parseur, mieux vaut mettre le mot clé **#PCDATA** entre parenthèses.

Exemple d'utilisation :

```
<!ELEMENT titre (#PCDATA)>
```

Se traduira par exemple dans le document XML :

```
<titre>XML par la pratique</titre>
```

Introduction

DTD interne

DTD externe

Déclarer les
éléments

Indicateurs
d'occurrence

Déclarer les
attributs

Limites des DTD

Déclarer les éléments

Élément vide

DTD (Document
Type Definition)

Introduction

DTD interne

DTD externe

Déclarer les
éléments

Indicateurs
d'occurrence

Déclarer les
attributs

Limites des DTD

```
<!ELEMENT nom EMPTY>
```

- L'élément vide n'a aucun contenu : pas de texte, ni même d'autres éléments.
- C'est une balise auto-fermante.

Exemple d'utilisation :

```
<!ELEMENT br EMPTY>
```

Se traduira par exemple dans le document XML :

```
<br />
```

Déclarer les éléments

Élément vide

DTD (Document
Type Definition)

Introduction

DTD interne

DTD externe

Déclarer les
éléments

Indicateurs
d'occurrence

Déclarer les
attributs

Limites des DTD

Remarque

Un élément vide peut tout à fait posséder des attributs. Par exemple :

```

```

Déclarer les éléments

Séquence d'éléments

DTD (Document
Type Definition)

Introduction

DTD interne

DTD externe

Déclarer les
éléments

Indicateurs
d'occurrence

Déclarer les
attributs

Limites des DTD

```
<!ELEMENT pere (fils1, fils2)>
```

- Une séquence d'éléments est une **liste ordonnée** d'éléments qui apparaîtront comme des éléments enfants de l'élément principal que l'on définit ici.
- Cet élément principal ne pourra contenir aucun autre élément que ceux déclarés dans la séquence.
- Les éléments enfants sont placés entre parenthèses et séparés par des virgules.

Déclarer les éléments

Séquence d'éléments

DTD (Document
Type Definition)

Introduction

DTD interne

DTD externe

Déclarer les
éléments

Indicateurs
d'occurrence

Déclarer les
attributs

Limites des DTD

Attention

- Chaque élément enfant doit être déclaré plus loin dans la DTD (peu importe sa position par rapport à la déclaration de la séquence).
- Les éléments enfants doivent apparaître dans le fichier XML dans l'ordre de déclaration de la séquence.

Déclarer les éléments

Séquence d'éléments

DTD (Document
Type Definition)

Introduction

DTD interne

DTD externe

Déclarer les
éléments

Indicateurs
d'occurrence

Déclarer les
attributs

Limites des DTD

Exemple d'utilisation :

```
<!ELEMENT auteur (nom, prenom)>
```

Se traduira par exemple dans le document XML :

```
<auteur>  
  <nom>Jules</nom>  
  <prenom>Dupond</prenom>  
</auteur>
```

Exemple invalide :

```
<auteur>  
  <prenom>Dupond</prenom>  
</auteur>
```

Déclarer les éléments

Choix d'éléments

DTD (Document
Type Definition)

Introduction

DTD interne

DTD externe

Déclarer les
éléments

Indicateurs
d'occurrence

Déclarer les
attributs

Limites des DTD

```
<!ELEMENT pere (fils1 | fils2)>
```

- Un choix d'éléments permet de définir dans une liste les éléments enfants possibles.
- L'utilisation précise dépend des indicateurs d'occurrence (abordés plus loin).
- Les éléments enfants sont placés entre parenthèses et séparés par des "|".

Attention

- Chaque élément enfant doit être déclaré dans la DTD.

Déclarer les éléments

Élément à contenu mixte

DTD (Document
Type Definition)

Introduction

DTD interne

DTD externe

Déclarer les
éléments

Indicateurs
d'occurrence

Déclarer les
attributs

Limites des DTD

- L'élément à contenu mixte est une liste de choix, avec des indicateurs d'occurrence bien sélectionnés.
- Il peut contenir aussi bien du texte que des éléments enfants.

Exemple de déclaration :

```
<!ELEMENT identite (#PCDATA | nom)*>
```

Exemple d'utilisation de cette déclaration :

```
<identite>  
  M. <nom>Dupond</nom>  
</identite>
```

Indicateurs d'occurrence

Présentation

DTD (Document
Type Definition)

Introduction

DTD interne

DTD externe

Déclarer les
éléments

Indicateurs
d'occurrence

Déclarer les
attributs

Limites des DTD

- Pour les séquences, les choix d'éléments et les éléments à contenu mixte, la notion d'indicateurs d'occurrence est mentionnée.
- Ces indicateurs permettent de définir les éléments XML qu'un élément peut ou doit contenir.
- Ils définissent des règles d'utilisation, grâce à une syntaxe spécifique.

Indicateurs d'occurrence

Notations

DTD (Document
Type Definition)

Introduction

DTD interne

DTD externe

Déclarer les
éléments

Indicateurs
d'occurrence

Déclarer les
attributs

Limites des DTD

- ? : un seul ou rien
- + : un ou plusieurs
- * : rien, un ou plusieurs (définition d'une option)
- | : l'un ou l'autre mais pas les deux
- , : obligatoires (dans l'ordre)
- () : regroupement \Rightarrow les parenthèses permettent de regrouper les éléments pour leur appliquer les autres opérateurs

Indicateurs d'occurrence

Exemples commentés

DTD (Document
Type Definition)

Introduction

DTD interne

DTD externe

Déclarer les
éléments

Indicateurs
d'occurrence

Déclarer les
attributs

Limites des DTD

- **A?** : l'élément A peut être présent (une seule fois) ou non
- **A+** : l'élément A doit être présent au moins une fois
- **A*** : l'élément A peut être présent plusieurs fois ou pas du tout
- **A|B** : l'élément A ou l'élément B peuvent être présents (mais pas les deux)
- **A,B** : l'élément A doit être présent, suivi de l'élément B (obligatoire également)
- **(A,B)+** : un ou plusieurs (A suivi de B)

Déclarer les attributs

Présentation

DTD (Document
Type Definition)

Introduction

DTD interne

DTD externe

Déclarer les
éléments

Indicateurs
d'occurrence

Déclarer les
attributs

Limites des DTD

Modèle de déclaration d'un attribut dans une DTD :

```
<!ATTLIST nom_element nom_attribut type_attribut  
mode>
```

- *nom_element* : nom de l'élément auquel cet attribut appartient ;
- *nom_attribut* : nom de l'attribut en cours de définition ;

Déclarer les attributs

Présentation

DTD (Document
Type Definition)

Introduction

DTD interne

DTD externe

Déclarer les
éléments

Indicateurs
d'occurrence

Déclarer les
attributs

Limites des DTD

- *type_attribut* : type de donnée de l'attribut :
 - **CDATA** ⇒ on affecte une chaîne de caractères à l'attribut
 - énumération
 - on définit une liste de valeurs possibles pour l'attribut (permet de limiter le choix de l'utilisateur) ⇒ **(Choix1 | Choix2 | ...)**
 - pour définir une valeur par défaut il faut faire suivre l'énumération par la valeur désirée entre guillemets ⇒ **(Choix1 | Choix2 | ...) "valeur par défaut"**
 - **NMTOKEN** ⇒ un mot, donc pas de blanc (chiffre autorisé). Exemple : utile pour déclarer un code postal.
 - **NMTOKENS** ⇒ liste de mots, séparés par des espaces (suite de NMTOKEN séparés par des espaces)

Déclarer les attributs

Présentation

DTD (Document
Type Definition)

Introduction

DTD interne

DTD externe

Déclarer les
éléments

Indicateurs
d'occurrence

Déclarer les
attributs

Limites des DTD

- **ENTITY** ⇒ nom d'entité non XML
- **ENTITIES** ⇒ liste de noms d'entités non XML, séparés par des espace (suite de ENTITY, séparés par des espaces))
- **ID** ⇒ on définit un identifiant unique pour chaque élément
- **IDREF(S)** ⇒ renvoi vers un (des) ID utilisé(s) ailleurs dans le document

Attention

- Un **ID** ne peut être que **#REQUIRED** ou **#IMPLIED**.
- Un élément ne peut disposer au maximum que d'un seul attribut **ID**.

Déclarer les attributs

Présentation

DTD (Document
Type Definition)

Introduction

DTD interne

DTD externe

Déclarer les
éléments

Indicateurs
d'occurrence

Déclarer les
attributs

Limites des DTD

- *mode* : précisions sur le type d'attribut :
 - **#IMPLIED** ⇒ attribut facultatif
 - **#REQUIRED** ⇒ attribut obligatoire
 - **#FIXED valeur** ⇒ attribut à valeur fixe (la valeur est déjà fixée dans la DTD)
 - **#DEFAULT valeur** ⇒ valeur par défaut

Déclarer les attributs

Exemples

DTD (Document
Type Definition)

Introduction

DTD interne

DTD externe

Déclarer les
éléments

Indicateurs
d'occurrence

Déclarer les
attributs

Limites des DTD

Exemple de fraction d'une DTD

```
<!ELEMENT date (#PCDATA)>  
<!ATTLIST date format (EN | FR) #REQUIRED>
```

- L'élément *date* est du texte.
- Cet élément dispose d'un attribut *format* obligatoire, ne pouvant prendre que la valeur *EN* ou *FR*.

Exemple de déclaration correcte dans le document XML

```
<date format="FR">8 octobre 2010</date>
```

Déclarer les attributs

Exemples

DTD (Document
Type Definition)

Introduction

DTD interne

DTD externe

Déclarer les
éléments

Indicateurs
d'occurrence

Déclarer les
attributs

Limites des DTD

Exemple de fraction d'une DTD

```
<!ELEMENT auteur (#PCDATA)>  
<!ATTLIST auteur numero ID #REQUIRED>  
<!ELEMENT livre (#PCDATA)>  
<!ATTLIST livre reference IDREF #REQUIRED>
```


Déclarer les attributs

Exemples

DTD (Document
Type Definition)

Introduction

DTD interne

DTD externe

Déclarer les
éléments

Indicateurs
d'occurrence

Déclarer les
attributs

Limites des DTD

Exemple d'une fraction de déclaration correcte dans le document XML

```
<auteur numero="a1">Thierry Boulanger</auteur>
<auteur numero="a2">Alexandre Brillant</auteur>
<livre reference="a1">XML par la pratique - Bases
    indispensables, concepts et cas pratiques</livre>
<livre reference="a2">XML : Cours et exercices</livre
>
```

Déclarer les attributs

Attributs multiples

DTD (Document
Type Definition)

Introduction

DTD interne

DTD externe

Déclarer les
éléments

Indicateurs
d'occurrence

Déclarer les
attributs

Limites des DTD

- Un élément peut posséder plusieurs attributs.
- Tous les attributs propres à un élément sont déclarés dans la même instruction

Exemple de déclaration XML :

```
<livre titre="Les DTD" genre="Technique" pages="60"  
  />
```

Exemple de fraction de DTD pouvant correspondre à la déclaration de ces attributs :

```
<!ATTLIST livre titre CDATA #REQUIRED  
  genre (Technique | Roman | Documentaire) "Roman"  
  pages CDATA #IMPLIED>
```

Limites des DTD

DTD (Document
Type Definition)

Introduction

DTD interne

DTD externe

Déclarer les
éléments

Indicateurs
d'occurrence

Déclarer les
attributs

Limites des DTD

- Le nombre d'apparitions d'un élément ne peut pas être contraint précisément. On ne dispose que des quantifieurs $?$, $*$ et $+$. On ne peut pas dire qu'un élément doit apparaître plus de 3 fois mais toujours moins de 7.
- On ne dispose pas de possibilité pour typer les contenus des éléments.
- On ne dispose que d'un typage faible pour les valeurs des attributs.
- On ne peut pas contraindre la forme de ces contenus (par exemple, entre 5 et 20 caractères, contenant un signe $@$, ...).

Limites des DTD

DTD (Document
Type Definition)

Introduction

DTD interne

DTD externe

Déclarer les
éléments

Indicateurs
d'occurrence

Déclarer les
attributs

Limites des DTD

- Il n'est pas possible de typer les références (elles peuvent référencer n'importe quel identifiant du document).

Pour pallier ces manques, d'autres propositions ont été faites, permettant de spécifier un langage XML de manière plus précise. Par exemple : XML Schema et Relax NG.