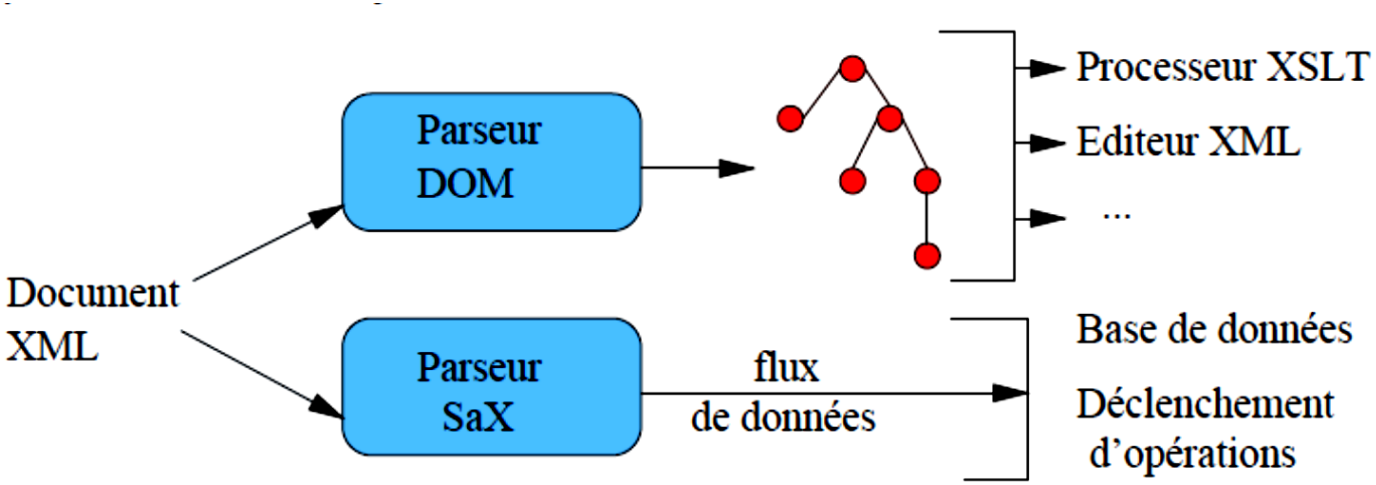
DOM

Les deux principales interfaces de programmation **XML** :

- **DOM** (Document Object Model), basé sur une représentation **hiérarchique**
  - ➢ Construit une représentation du document en mémoire sous forme d'arbre
- **SaX** (Simple **API** for XML), basé sur des déclencheurs (événements/action)
  - ➢ Définit des triggers qui se déclenchent sur certaines balises

DOM

Toutes les applications XML passent par une phase préalable d'analyse



DOM

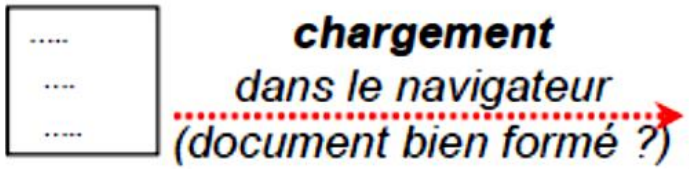
- DOM est l'acronyme de **D**ocument **O**bject **M**odel
- C'est une spécification du **W3C** pour proposer une **API** qui permet de parcourir et de manipuler un document **XML**
- DOM fournie une représentation mémoire d'un document XML sous la forme d'un arbre d'objets et d'en permettre la manipulation (parcours, recherche et mise à jour)

DOM

DOM permet:

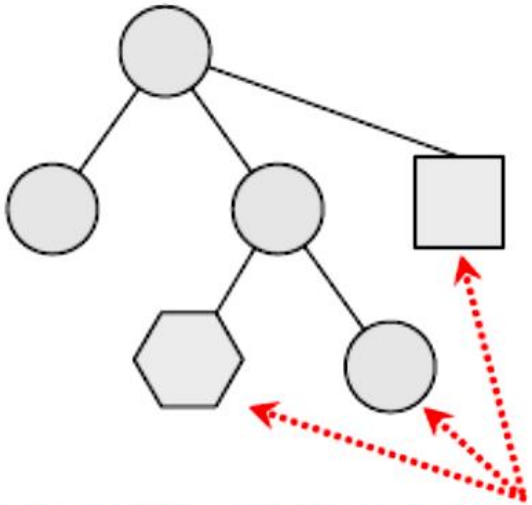
- Fournir une **API** indépendante des langages de programmation.
- Une représentation structurée et orientée objet des éléments et du contenue
- La modification des propriétés de ces objets par l'intermédiaire de méthodes
- L'ajout et la suppression d'objets
- La gestion des évènements du navigateur

Construction de l'arbre XML



document **XML**

chargement
dans le navigateur
(document bien formé ?)



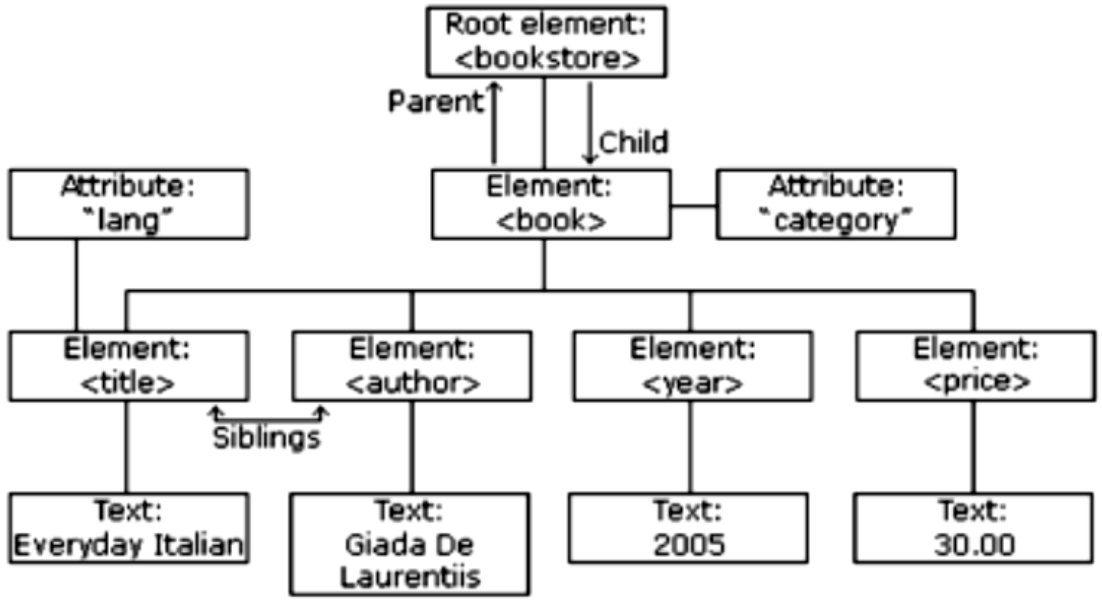
arbre XML = hiérarchie de nœuds de types différents

- **une racine** (le document lui-même),
- **instruction de traitement** (<? ….?>)
- **commentaire**, (<! --….-->)
- **élément**, (balises)
- **texte**, (le texte PCDATA d'une balise)
- **attribut**, (attribut dans une balise)
- ….. etc.

Exemple de la hiérarchie des nœuds DOM

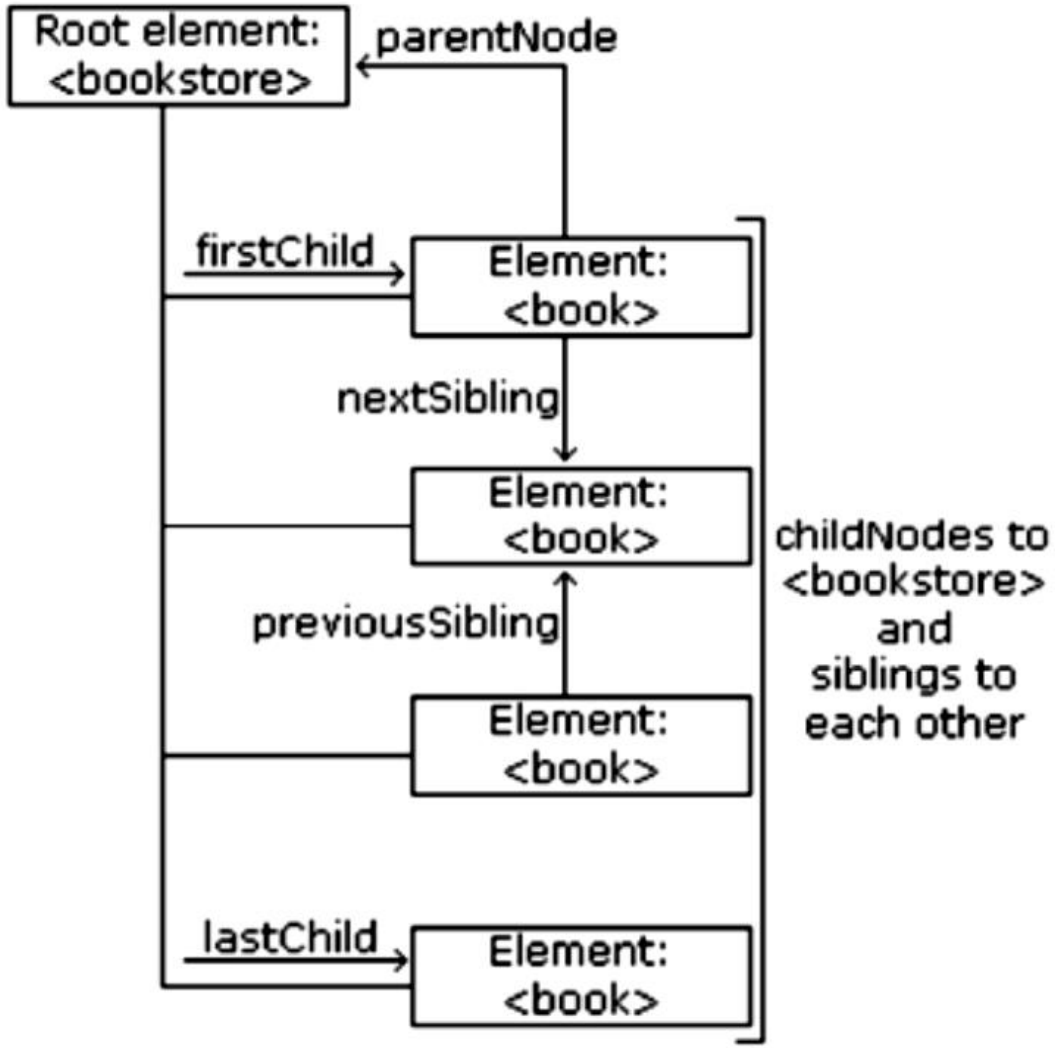
```
<?xml version="1.0" encoding="ISO-8859-1"?>
<bookstore>
 <book category="COOKING">
   <title lang="en">Everyday Italian</title>
   <author>Giada De Laurentiis</author>
   <year>2005</year>
   <price>30.00</price>
</book>
<book category="CHILDREN">
   <title lang="en">Harry Potter</title>
   <author>J K. Rowling</author>
   <year>2005</year>
   <price>29.99</price>
 </book>
  ...
<bookstore>
```

The corresponding Tree



Parcours d'un arbre de noeuds XML à l'aide de DOM

- **On peut naviguer entre les noeuds en utilisant les différentes relations qui peuvent exister entre eux:**
  - parentNode
  - childNodes
  - firstChild
  - lastChild
  - nextSibling
  - previousSibling



Structure de l'arbre XML en Noeuds

- **Un seul nœud "Document" contenant des nœuds fils :**
  - **"commentaire"** → *zéro ou plusieurs*
  - **"instructions de traitement"** → *zéro ou plusieurs*
  - **"DOCTYPE"** → *au plus un, la déclaration DOCTYPE dans le prologue XML*
  - **"racine"** → *un, le corps du document XML*

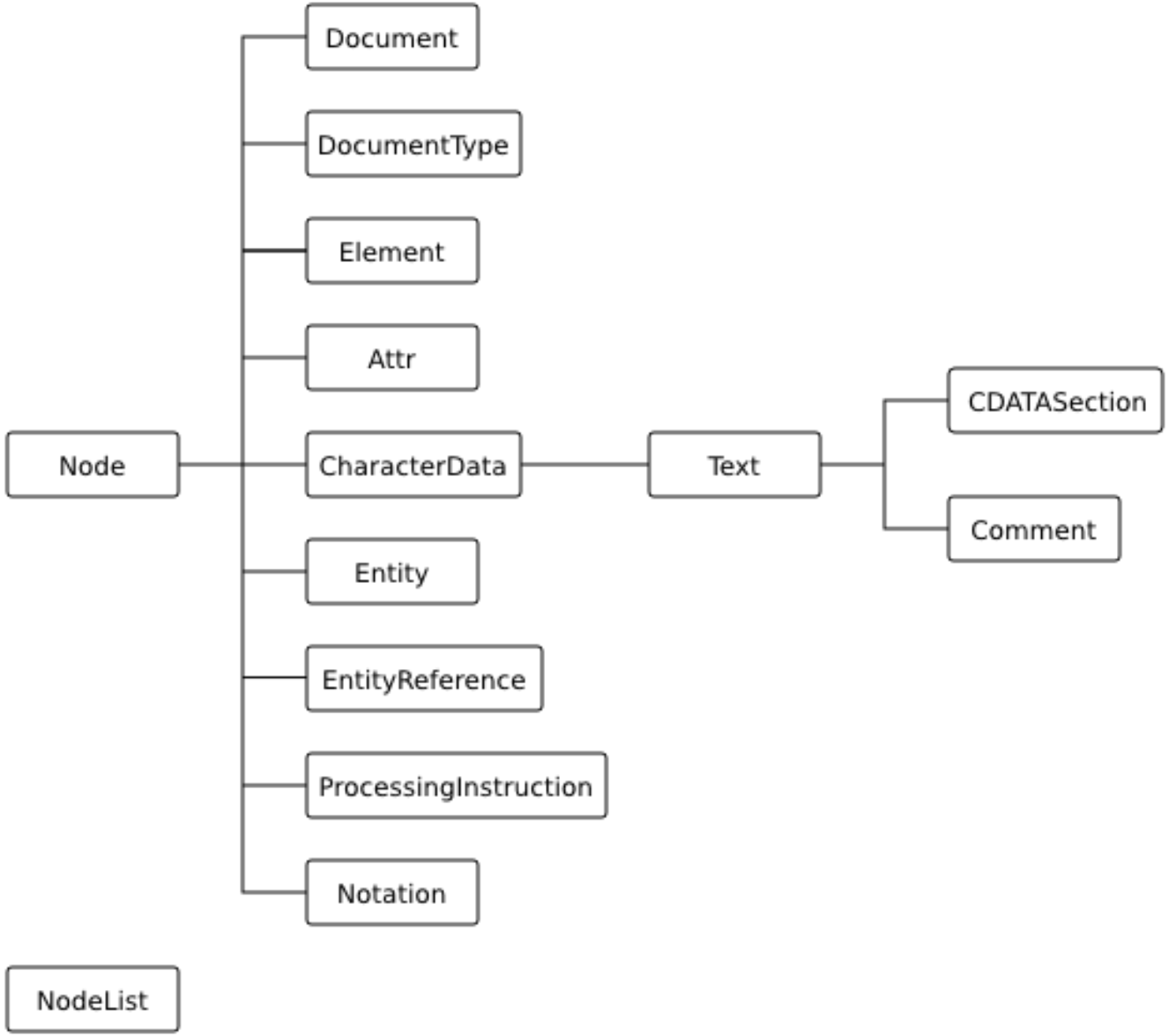
- **Les nœuds fils d'une "racine" sont de type :**
  - **"element"** → balises filles
  - **"texte"** → le texte entre les deux balises

    NB: Il peut y avoir plusieurs nœuds texte si le contenu est mélangé avec des balises
  - **"commentaire"** → *zéro ou plusieurs*
  - **"instructions de traitement"** → *zéro ou plusieurs*
  - **"section CDATA"** → *texte non interprété entre* [….]

DOM

- DOM définie une hiérarchie de classes pour le traitement des nœuds XML dont **Node** est la classe DOM principale
- Chaque objet DOM définit :
  - Des propriétés pour les nœuds (accès en lecture seule ou lecture-écriture),
  - Des méthodes de traitement.

Principales classes DOM



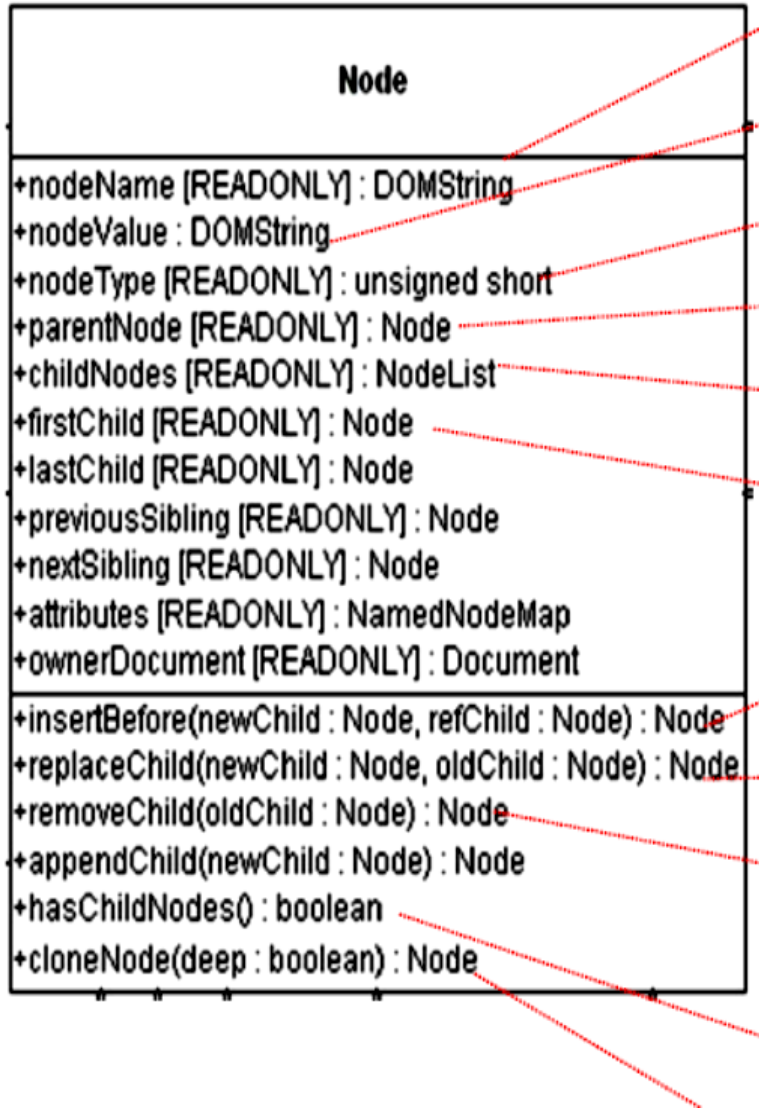
DOM

- L’interface **Node**: est la classe de base de la structure. Tous les nœuds de l'arbre sont basés sur cette classe

- Elle comprend un certain nombre d'attributs permettant de décrire le nœud, mais aussi de naviguer dans la structure:

  - nodeName, nodeValue, nodeType, parentNode, childNodes, firstChild, lastChild, previousSibling, nextSibling …etc.

DOM



*Le **nom du nœud** (i.e. nom de la balise, de l'attribut etc.)*

*La **valeur du nœud** (i.e. contenu de la balise, l'attribut etc.)*

*Le **type du nœud** (1=Element; 2=Attribut; 3=Text; 4=CDATASection; 9=Document)*

*Le **nœud parent***

*La **liste des nœuds fils***

*Le **premier nœud fils dans la liste** ……etc.*

Insérer un nœud avant

Remplacer un nœud fils

Effacer un nœud fils

a-t-il des nœuds fils ?

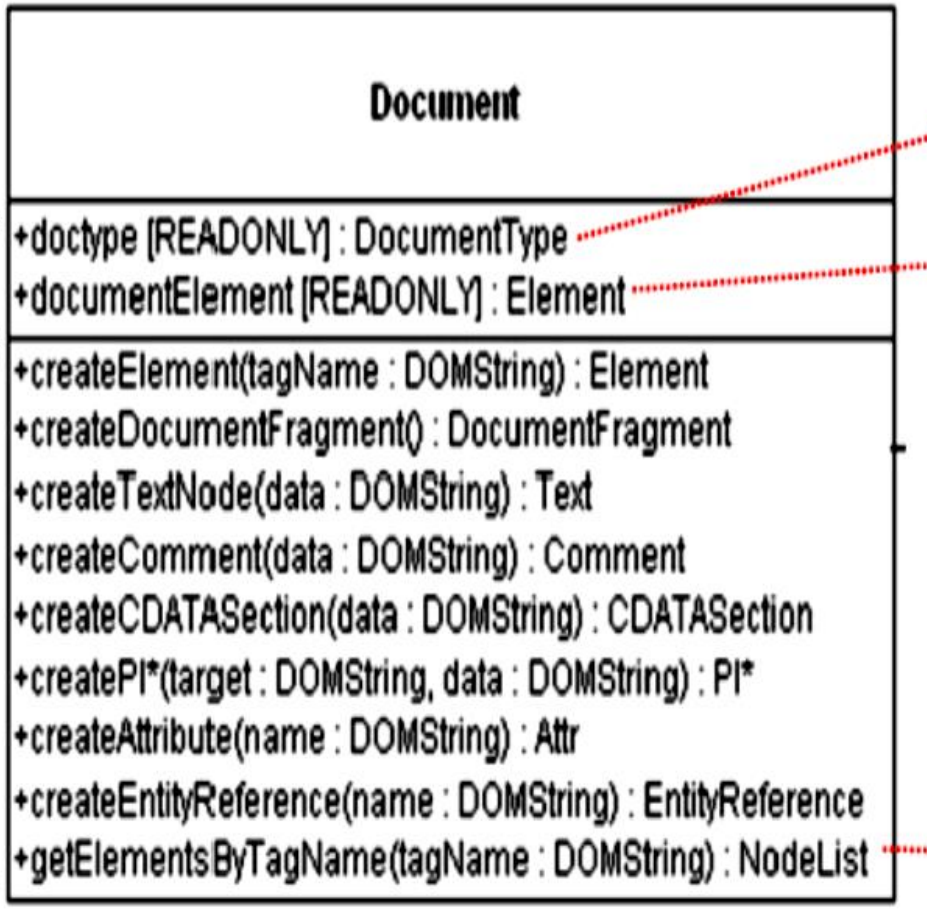
DOM

- L’interface **Document:** cette interface caractérise l’ensemble du document. Voici quelques méthodes d’usage courant :
  - getDocumentElement(),getElementById(String Id), createElement(String tagName), createTextNode(String data), createAttribute(String name), createComment(String data)…etc,
- L’interface **Element:** voici quelques méthodes d’usage courant :
  - getAttribute(String name) / setAttribute(String name), getElementsByTagName(String name),

DOM

attributs
(READONLY = en lecture seule)

méthodes



Type du document

La racine du document

Rechercher une balises par nom