

# XML

## eXtensible Markup Language

Bérénice Lemoine et Olivier Montet

Faculté des Sciences  
Université de Montpellier  
Département Informatique

TCCP 2017

# Un petit peu d'histoire...

- Au début d'internet la transmission d'information était fastidieuse.

# Un petit peu d'histoire...

- Au début d'internet la transmission d'information était fastidieuse.
- D'où l'invention du SGML (Standard Generalized Markup Language).

# Un petit peu d'histoire...

- Au début d'internet la transmission d'information était fastidieuse.
- D'où l'invention du SGML (Standard Generalized Markup Language).
- Celui-ci étant trop complexe et ayant des échanges de données difficiles avec le web, le XML 1.0 (eXtensible Markup Language) vu le jour en 1998.

# Un petit peu d'histoire...

- Au début d'internet la transmission d'information était fastidieuse.
- D'où l'invention du SGML (Standard Generalized Markup Language).
- Celui-ci étant trop complexe et ayant des échanges de données difficiles avec le web, le XML 1.0 (eXtensible Markup Language) vu le jour en 1998.
- La même année le XML est devenu une recommandation du W3C ("World Wide Web Consortium")

# Un petit peu d'histoire...

- Au début d'internet la transmission d'information était fastidieuse.
- D'où l'invention du SGML (Standard Generalized Markup Language).
- Celui-ci étant trop complexe et ayant des échanges de données difficiles avec le web, le XML 1.0 (eXtensible Markup Language) vu le jour en 1998.
- La même année le XML est devenu une recommandation du W3C ("World Wide Web Consortium")
- Malgré la nouvelle version XML 1.1 sortie en 2004, XML 1.0 reste celle la plus utilisée.

- 1 Qu'est-ce que le XML ?
  - Un langage de description
  - Un langage de balisage générique
- 2 Particularités du XML
  - Créer un fichier XML bien formé
  - Créer un fichier XML valide
- 3 Utilisation du XML
  - Généralités
  - Exemples d'applications
  - Application avec Java
- 4 Le XML n'a pas que des avantages...
  - Inconvénients
- 5 Conclusion

- 1 Qu'est-ce que le XML ?
  - Un langage de description
  - Un langage de balisage générique
- 2 Particularités du XML
  - Créer un fichier XML bien formé
  - Créer un fichier XML valide
- 3 Utilisation du XML
  - Généralités
  - Exemples d'applications
  - Application avec Java
- 4 Le XML n'a pas que des avantages...
  - Inconvénients
- 5 Conclusion



# Qu'est-ce que le XML ?

Un langage de description

- Le XML est un langage qu'on peut appeler langage de description, il permet de décrire et de structurer un ensemble de données selon un jeu de règles et des contraintes définies. On peut par exemple l'utiliser pour décrire l'ensemble des livres d'une bibliothèque, ou encore la liste des chansons d'un CD, etc.

- 1 Qu'est-ce que le XML ?
  - Un langage de description
  - Un langage de balisage générique
- 2 Particularités du XML
  - Créer un fichier XML bien formé
  - Créer un fichier XML valide
- 3 Utilisation du XML
  - Généralités
  - Exemples d'applications
  - Application avec Java
- 4 Le XML n'a pas que des avantages...
  - Inconvénients
- 5 Conclusion

# Qu'est-ce que le XML ?

## Langage de Balisage Générique

- Un *langage de balisage* est un langage qui s'écrit grâce à des balises. Ces balises permettent de structurer de manière hiérarchisée et organisée les données d'un document. Chaque balise peut-elle-même avoir différentes propriétés grâce aux attributs.

# Qu'est-ce que le XML ?

## Langage de Balisage Générique

- Un *langage de balisage* est un langage qui s'écrit grâce à des balises. Ces balises permettent de structurer de manière hiérarchisée et organisée les données d'un document. Chaque balise peut-elle-même avoir différentes propriétés grâce aux attributs.
- Le terme *générique* signifie que nous allons pouvoir créer nos propres balises. Nous ne sommes pas obligés d'utiliser un ensemble de balises existantes comme c'est par exemple le cas en HTML.

# Qu'est-ce que le XML ?

## Langage de Balisage Générique

- Un *langage de balisage* est un langage qui s'écrit grâce à des balises. Ces balises permettent de structurer de manière hiérarchisée et organisée les données d'un document. Chaque balise peut-elle-même avoir différentes propriétés grâce aux attributs.
- Le terme *générique* signifie que nous allons pouvoir créer nos propres balises. Nous ne sommes pas obligés d'utiliser un ensemble de balises existantes comme c'est par exemple le cas en HTML.
- exemple : `<telephone type="fixe">`

- 1 Qu'est-ce que le XML ?
  - Un langage de description
  - Un langage de balisage générique
- 2 Particularités du XML
  - Créer un fichier XML bien formé
  - Créer un fichier XML valide
- 3 Utilisation du XML
  - Généralités
  - Exemples d'applications
  - Application avec Java
- 4 Le XML n'a pas que des avantages...
  - Inconvénients
- 5 Conclusion

# Créer un fichier XML bien formé

## L'entête

### OBLIGATOIRE !

Tout document XML doit commencer par la ligne :

```
<?xml version="1.0" encoding="UTF-8"?>
```

# Créer un fichier XML bien formé

## Les balises I

### Obligations

On peut créer toutes les balises que l'on souhaite, à **condition** qu'elles respectent certaines règles :

- Toute balise ouverte doit être correctement fermée :  
`<balise> Voici une balise </balise>`



# Créer un fichier XML bien formé

## Les balises I

### Obligations

On peut créer toutes les balises que l'on souhaite, à **condition** qu'elles respectent certaines règles :

- Toute balise ouverte doit être correctement fermée :  
`<balise> Voici une balise </balise>`
- Il n'y a qu'une balise racine :  
`<racine>`

# Créer un fichier XML bien formé

## Les balises I

### Obligations

On peut créer toutes les balises que l'on souhaite, à **condition** qu'elles respectent certaines règles :

- Toute balise ouverte doit être correctement fermée :  
`<balise> Voici une balise </balise>`
- Il n'y a qu'une balise racine :  
`<racine>`
- Les attributs de la balise sont entre guillemets :  
`<balise attribut="A1" >`

# Créer un fichier XML bien formé

## Les balises II

### Interdictions

Les balises interdites :

- `<8balise>`  
`<bal-isé>`  
`<balise une>`

### Interdictions

Les balises interdites :

- `<8balise>`  
`<bal-isé>`  
`<balise une>`
- Les chevauchements sont interdits :  
`<baliseUne<baliseDeux>`  
`<baliseUne ><baliseDeux ></baliseUne >`

# Créer un fichier XML bien formé

## Balises et attributs

### Attributs d'une balise

- Le langage XML est générique, on a vu que l'on peut choisir les balises que l'on veut. C'est également le cas pour leurs attributs qui vont nous permettre de donner des détails supplémentaires sur l'information que contient la balise en question.

# Créer un fichier XML bien formé

## Balises et attributs

### Attributs d'une balise

- Le langage XML est générique, on a vu que l'on peut choisir les balises que l'on veut. C'est également le cas pour leurs attributs qui vont nous permettre de donner des détails supplémentaires sur l'information que contient la balise en question.
- `<prix devise="euro" moyenpaiement="chèque">25.3 </prix>`

# Créer un fichier XML bien formé

## Exemple de fichier bien formé

### *Exemple de fichier bien formé*

```
<?xml version="1.0" encoding="UTF-8"?>
<personne sexe="Feminin">
  <nom>CONNOR</nom>
  <prenom>Sarah</prenom>
  <telephones>
    <telephone type="fixe">none</telephone>
    <telephone type="portable">06 72 15 95 62</telephone>
  </telephones>
</personne>
```

- 1 Qu'est-ce que le XML ?
  - Un langage de description
  - Un langage de balisage générique
- 2 Particularités du XML
  - Créer un fichier XML bien formé
  - Créer un fichier XML valide
- 3 Utilisation du XML
  - Généralités
  - Exemples d'applications
  - Application avec Java
- 4 Le XML n'a pas que des avantages...
  - Inconvénients
- 5 Conclusion



# Créer un fichier XML valide

## Définition

### *Définition du mot Valide*

C'est un document XML bien formé, qui en plus, respecte les règles d'un fichier de définition DTD (Définition de Type de Document) ou schémas XML. Pour un fichier DTD il faut mettre en entête la ligne suivante :

```
<!DOCTYPE AML SYSTEM "nomfichier.dtd" >
```

### *Définition d'un document XML*

C'est un ensemble de règles qu'on impose au document qui permettent de décrire la façon dont le document XML doit être construit.

# Créer un fichier XML valide

## Exemple

*Voici un exemple de définition de document XML de l'exemple plus haut :*

```
<!ELEMENT personne (nom, prenom, telephones)>  
<!ATTLIST personne sexe (masculin | feminin) REQUIRED>  
<!ELEMENT nom (PCDATA)>  
<!ELEMENT prenom (PCDATA)>  
<!ELEMENT telephones (telephone+)>  
<!ELEMENT telephone (PCDATA)>  
<!ATTLIST telephone type CDATA REQUIRED>
```

- 1 Qu'est-ce que le XML ?
  - Un langage de description
  - Un langage de balisage générique
- 2 Particularités du XML
  - Créer un fichier XML bien formé
  - Créer un fichier XML valide
- 3 Utilisation du XML
  - Généralités
  - Exemples d'applications
  - Application avec Java
- 4 Le XML n'a pas que des avantages...
  - Inconvénients
- 5 Conclusion

- Seul le XML n'est pas très utile au final...
- Il trouve tout son intérêt et succès grâce à la galerie de langages qui gravitent autour de lui ! Les DTD par exemple font partie de cette galerie. Ces langages permettent de manipuler les documents XML, les transformer, mettre en forme, créer des liens ou comme nous l'avons vu précédemment, décrire leur contenu.

- Seul le XML n'est pas très utile au final...
- Il trouve tout son intérêt et succès grâce à la galerie de langages qui gravitent autour de lui ! Les DTD par exemple font partie de cette galerie. Ces langages permettent de manipuler les documents XML, les transformer, mettre en forme, créer des liens ou comme nous l'avons vu précédemment, décrire leur contenu.
- Beaucoup de logiciels peuvent utiliser le XML. Les langages tels que Java, C++, etc. utilisent des fichiers XML pour décrire les programmes et leurs paramètres de sécurité. De même les logiciels comme OpenOffice, FreeMind, etc. utilisent le format XML pour différentes tâches mais notamment pour stocker leur configuration.

- 1 Qu'est-ce que le XML ?
  - Un langage de description
  - Un langage de balisage générique
- 2 Particularités du XML
  - Créer un fichier XML bien formé
  - Créer un fichier XML valide
- 3 Utilisation du XML
  - Généralités
  - Exemples d'applications
  - Application avec Java
- 4 Le XML n'a pas que des avantages...
  - Inconvénients
- 5 Conclusion

# Utilisation du XML

## Exemple d'applications

- Le SVG (*Scalable Vector Graphics*), est un langage qui permet d'écrire des graphiques vectoriels 2D en XML.

# Utilisation du XML

## Exemple d'applications

- Le SVG (*Scalable Vector Graphics*), est un langage qui permet d'écrire des graphiques vectoriels 2D en XML.
- Le Xlink est une spécification du W3C (*World Wide Web*). Il permet de créer des hyperliens dans des documents XML mais contrairement aux liens en HTML, ces liens sont multidirectionnels, c'est-à-dire qu'il peuvent pointer vers plusieurs documents. Le SVG utilise XLink.



- 1 Qu'est-ce que le XML ?
  - Un langage de description
  - Un langage de balisage générique
- 2 Particularités du XML
  - Créer un fichier XML bien formé
  - Créer un fichier XML valide
- 3 Utilisation du XML
  - Généralités
  - Exemples d'applications
  - **Application avec Java**
- 4 Le XML n'a pas que des avantages...
  - Inconvénients
- 5 Conclusion

# Utilisation du XML

## Exemple Java I

### *Exemple de fichier bien formé*

```
<?xml version="1.0" encoding="UTF-8"?>
<personne sexe="Feminin">
  <nom>CONNOR</nom>
  <prenom>Sarah</prenom>
  <telephones>
    <telephone type="fixe">none</telephone>
    <telephone type="portable">06 72 15 95 62</telephone>
  </telephones>
</personne>
```

*Les premières étapes pour créer un programme Java permettant la lecture d'un fichier XML sont :*

- Récupérer une instance de la classe "DocumentBuilderFactory", c'est-à-dire un DOM (Document Object Memory) permettant de fournir une représentation mémoire du document XML sous la forme d'un arbre d'objets et d'en permettre la manipulation (parcours, recherche et mise à jour).

*Les premières étapes pour créer un programme Java permettant la lecture d'un fichier XML sont :*

- Récupérer une instance de la classe "DocumentBuilderFactory", c'est-à-dire un DOM (Document Object Memory) permettant de fournir une représentation mémoire du document XML sous la forme d'un arbre d'objets et d'en permettre la manipulation (parcours, recherche et mise à jour).
- Création d'un parseur, c'est-à-dire, un analyseur syntaxique XML qui permet de récupérer dans une structure XML, des balises, leurs contenus, leurs attributs et de les rendre accessibles.

*Pour la suite du code il y a deux options :*

- Faire un code général dans le cas où l'on ne connaît pas la structure du document XML.

*Pour la suite du code il y a deux options :*

- Faire un code général dans le cas où l'on ne connaît pas la structure du document XML.
- Ou bien faire un code correspondant à un fichier XML propre (On ne pourra donc pas utiliser ce code pour lire un autre fichier XML que celui voulu.

*Pour la suite du code il y a deux options :*

- Faire un code général dans le cas où l'on ne connaît pas la structure du document XML.
- Ou bien faire un code correspondant à un fichier XML propre (On ne pourra donc pas utiliser ce code pour lire un autre fichier XML que celui voulu.
- L'exemple que l'on va vous montrer correspond exactement au fichier XML vu plus haut.

- 1 Qu'est-ce que le XML ?
  - Un langage de description
  - Un langage de balisage générique
- 2 Particularités du XML
  - Créer un fichier XML bien formé
  - Créer un fichier XML valide
- 3 Utilisation du XML
  - Généralités
  - Exemples d'applications
  - Application avec Java
- 4 Le XML n'a pas que des avantages...
  - Inconvénients
- 5 Conclusion



# Le XML n'a pas que des avantages...

## Inconvénients I

*Le XML est un langage très verbeux :*

```
<NOMBREEXEMPLAIRES>5 </NOMBREEXEMPLAIRES>
```

- L'information n'occupe qu'un octet mais les balises qui l'entourent en occupent 39 (octets).

# Le XML n'a pas que des avantages...

## Inconvénients I

*Le XML est un langage très verbeux :*

```
<NOMBREEXEMPLAIRES>5 </NOMBREEXEMPLAIRES>
```

- L'information n'occupe qu'un octet mais les balises qui l'entourent en occupent 39 (octets).
- **Enorme gaspillage de la mémoire.**

# Le XML n'a pas que des avantages...

## Inconvénients II

### *Le XML est un langage lent :*

Ce type de document étant un format texte, il oblige l'ordinateur à tout convertir en binaire (pour s'en faire une "représentation" en mémoire), faire ses traitements, puis tout reconverter en texte dans l'autre sens.

# Le XML n'a pas que des avantages...

## Inconvénients III

### *Autres inconvénients :*

- Il est très mal adapté à la représentation d'autres types de données, comme les données relationnelles.

# Le XML n'a pas que des avantages...

## Inconvénients III

### *Autres inconvénients :*

- Il est très mal adapté à la représentation d'autres types de données, comme les données relationnelles.
- Il est également incapable de stocker des informations binaires, comme des images, vidéos ou musiques.

# Le XML n'a pas que des avantages...

## Inconvénients III

### *Autres inconvénients :*

- Il est très mal adapté à la représentation d'autres types de données, comme les données relationnelles.
- Il est également incapable de stocker des informations binaires, comme des images, vidéos ou musiques.
- De plus, dans la pratique, les fichiers XML que nous rencontrons sont fortement imbriqués, ce qui rend la lecture difficile pour l'homme.

# Le XML n'a pas que des avantages...

## Inconvénients III

### *Autres inconvénients :*

- Il est très mal adapté à la représentation d'autres types de données, comme les données relationnelles.
- Il est également incapable de stocker des informations binaires, comme des images, vidéos ou musiques.
- De plus, dans la pratique, les fichiers XML que nous rencontrons sont fortement imbriqués, ce qui rend la lecture difficile pour l'homme.
- Et enfin, les technologies qui entourent le XML sont nombreuses et complexes, il est donc difficile de bien les maîtriser.

- Le XML est langage de balisage générique qui permet la structuration de données et l'échanges de celles-ci avec le web assez facilement. De plus, il est lisible aussi bien par un ordinateur que par un humain.



- Le XML est langage de balisage générique qui permet la structuration de données et l'échanges de celles-ci avec le web assez facilement. De plus, il est lisible aussi bien par un ordinateur que par un humain.
- De nombreuses technologies l'utilisent pour structurer leur données.

- Le XML est langage de balisage générique qui permet la structuration de données et l'échanges de celles-ci avec le web assez facilement. De plus, il est lisible aussi bien par un ordinateur que par un humain.
- De nombreuses technologies l'utilisent pour structurer leur données.
- Même si c'est un langage facilitant les échanges de données simplement, il possède des inconvénients non négligeables surtout au niveau de la place mémoire nécessaire.

Merci de votre attention.  
Si vous avez des questions n'hésitez pas à les poser.



## Open ClassRooms

<https://openclassrooms.com/courses/structurez-vos-donnees-avec-xml>

<https://openclassrooms.com/courses/le-point-sur-xml>



## sebsauvage.net

<http://sebsauvage.net/comprendre/xml/>



## Wikipedia

[https://fr.wikipedia.org/wiki/Extensible\\_Markup\\_Language](https://fr.wikipedia.org/wiki/Extensible_Markup_Language)



## The XML Family of Specifications : The Big Picture

<http://www.daisy.org/samples/daisypedia/bigpix-filer/bigpix22.gif>