
WEB DES DONNÉES

TP N° 1 – XML

Romain LELONG – romain.lelong@gmail.com

Avant de commencer ...

Outre le présent sujet, les fichiers suivants doivent être récupérés avant de débiter le TP : `xml-en-xml.xml`, `exponentielle.mml`, `smiley.svg`.

Les quatre parties de ce TP sont indépendantes. Chacune d'elles fait l'objet d'un fichier au format XML à modifier ou à créer de toute pièce.

Ce TP ne sera pas noté et aucun compte-rendu n'est demandé. Vous pouvez le réaliser en binôme ou seul. Les divers éléments abordés dans celui-ci pourront néanmoins faire l'objet d'une micro-évaluation prochaine.

1 Installation d'xmllint

Vous aurez besoin au cours du module de Web des Données d'utiliser l'outil `xmllint`.

Il s'agit d'un outil en ligne de commande. Celui-ci offre une gamme de fonctionnalités variée dans le contexte de la manipulation des fichiers XML (e.g. Analyse syntaxique, Formatage, Validation (e.g. DTD et XSD, etc)).

Cette outil est déjà installé sur les machines de l'université. Si vous travaillez avec ces dernières vous pouvez passer à la section suivante. En revanche, si vous souhaitez réaliser votre TP avec votre machine personnelle vous pouvez suivre la procédure d'installation suivante.

Installation sur Ubuntu

Il suffit d'entrer dans un terminal les deux ligne de commande suivante :

```
sudo apt-get update
sudo apt-get install libxml2-utils
```

Installation sur Windows

`xmllint` est également disponible sur Windows en utilisant le gestionnaire de packages `Chocolatey`. Il convient donc de procéder à son installation avant celle d'`xmllint`. Voici donc la procédure à suivre :

1. Chercher « Windows Powershell » et l'ouvrir en tant qu'administrateur ;
2. Vérifier l'exécution en entrant la commande `Get-ExecutionPolicy`. Si celle-ci renvoie `Restricted`, faites `Set-ExecutionPolicy AllSigned` puis entrez `Y` ;
3. Entrer la commande suivante :


```
Set-ExecutionPolicy Bypass -Scope Process -Force;
[System.Net.ServicePointManager]::SecurityProtocol = [System.Net.ServicePointManager]::SecurityProtocol -bor 3072;
iex ((New-Object System.Net.WebClient).DownloadString('https://community.chocolatey.org/install.ps1'))
```

Chocolatey devrait maintenant être installé.
4. Utiliser la commande `choco install xsltproc` pour installer `xmllint`.

`xmllint` est donc installé et devrait automatiquement être disponible dans l'invite de commandes.

2 XML en XML

Certaines des informations de la « fiche info Google » affichée pour le mot clé "XML" (voir copie d'écran ci-contre) ainsi que de celles obtenue pour le World Wide Web Consortium (W3C) ont été colligées dans le fichier `xml-en-xml.xml`. Malheureusement ce document XML est mal formé et comporte des erreurs.

Étape n° 1 : Ouvrir le fichier `xml-en-xml.xml` à l'aide du navigateur  Firefox.

Étape n° 2 : Corriger pas-à-pas chacune des erreurs. On veillera à ce que ces corrections n'engendrent aucune perte d'information.

Étape n° 3 : Si l'on se base sur le document XML d'origine, avec quel espace de noms les éléments ou attributs suivants sont-ils qualifiés :

— l'élément `<normes>` (ligne 10) ?

Extensible Markup Language

Langage de programmation

L'Extensible Markup Language, généralement appelé XML, « langage de balisage extensible » en français, est un métalangage informatique de balisage générique qui est un sous-ensemble du Standard Generalized Markup Language. [Wikipédia](#)

Norme : 1.0 (5^e édition) 1.1 (2^e édition)

Spécification : Format ouvert

Basé sur : [Standard Generalized Markup Language](#), [Unicode](#)


Développé par : [World Wide Web Consortium](#)

PUID : fmt/101


Type de format : Langage de balisage

Version initiale : 1999


Recherches associées Voir d'autres éléments (plus de 10)




JavaScript Object Notation



Hypertext Markup Language




JavaScript



Comma-separated values

3 Format MathML

Le Mathematical Markup Language (MathML) est un langage au format XML. Il a pour objectif de permettre l'écriture et l'affichage de symboles mathématiques sur le Web. Ce langage fait en outre partie des recommandations du W3C.

Étape n° 4 : Ouvrir le fichier `exponentielle.mml` avec le navigateur  Firefox.

Étape n° 5 : Qu'est ce qui permet au navigateur d'interpréter ce fichier comme une équation et non comme du simple code XML ? (Objectiver votre réponse)

Étape n° 6 : Trouver la ligne de commande de l'outil `xmlint` qui permet de ré-indenter le fichier `exponentielle.mml` ?

Étape n° 7 : Compléter le fichier `exponentielle.mml` afin que ce dernier affiche l'équation :

$$e^x = \sum_{n=0}^{+\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!} + o(x^n)$$

Voici une description de quelques unes des balises employées dans MathML :

`<mrow>` : Utilisé pour regrouper n'importe quel nombre de sous-expressions de manière horizontale.

`<mi>` : Utilisé pour spécifier un identifiant tel que le nom d'une variable, une fonction ou une constante, etc. (e.g. `<mi>PI</mi>`).

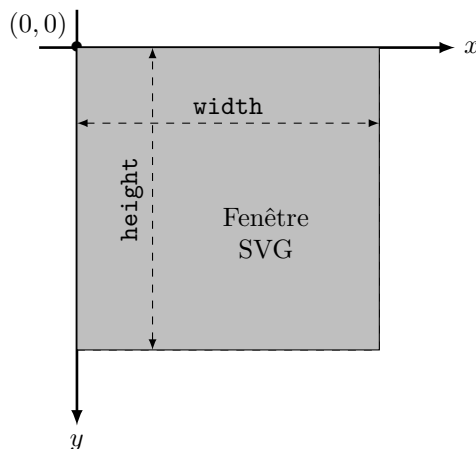
`<mn>` : Utilisé pour spécifier un littéral numérique comme 3.14 etc. (e.g. `<mn>3.14</mn>`).

`<mo>` : Utilisé pour spécifier un opérateur comme +, - etc. (e.g. `<mo>+</mo>`).

`<mfenced>` : Fournit la possibilité d'ajouter des parenthèses et des séparateurs d'ouverture et de fermeture personnalisés à une expression.

4 Format SVG

Le Scalable Vector Graphics (SVG) est un format d'image vectorielle basé sur XML. En pratique, il permet de « *déssiner* » des graphiques bidimensionnels à l'intérieur d'une fenêtre préalablement définie. Le système de coordonnées orienté utilisé en SVG est schématisé ci-dessous :



Étape n° 8 : Ouvrir le fichier `smiley.svg` avec le navigateur 🦊 Firefox.

Étape n° 9 : Corriger cette erreur. Cette correction ne devra pas alourdir inutilement le fichier mais au contraire le rendre plus compact et concis.

Étape n° 10 : Ouvrir à nouveau le fichier dans le navigateur 🦊 Firefox.

Étape n° 11 : Remplacer dans le fichier `smiley.svg`, la valeurs de quatre attributs par l'une des valeurs littérales suivantes : `"zombie-mouth"`, `"zombie-right-eye"`, `"zombie-left-eye"`, `"zombie-head"`.

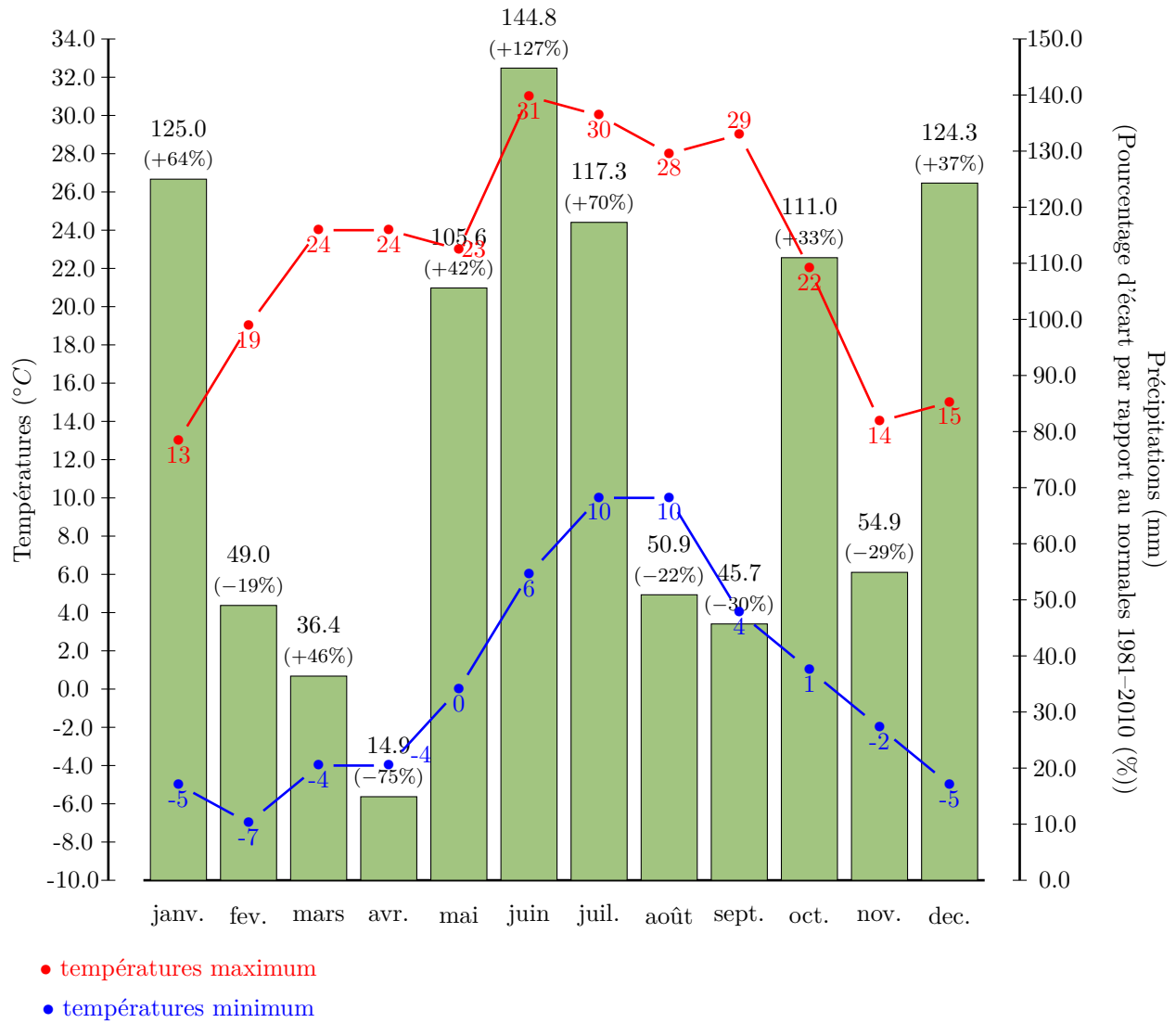
Étape n° 12 : Laisser votre âme d'artiste s'exprimer et opérer une « *dézombification* » de votre smiley en lui redonnant un teint et une expression faciale plus joviale.

5 XML

En page 5 sont données diverses informations climatiques enregistrées par la station météorologique de Rouen-Boos en 2021.

Étape n° 13 : Proposer un fichier XML intitulé `climat.xml` permettant de structurer ces informations en XML. On se contentera d'inclure dans ce fichier uniquement les données relatives aux 3 premiers mois de l'année (i.e. janvier, février et mars).

Températures et précipitations en 2021 – Rouen-Boos



Épisodes neigeux en 2021 – Rouen-Boos

Mois	Jour	heures
Janvier	04	9h00
	16	7h00
		9h00
	23	9h00
		10h00
	24	10h00
		11h00
Février	07	0h00
Novembre	27	4h00
		6h00