

STRUCTURES ET MODÈLES DES CONTENUS DU WEB

Conception et transformations de langages de balisage

Charles Follet, Roland Bary

April 20, 2015

Université de Pau et des Pays de l'Adour

Introduction

Les besoins à modéliser

Conception des langages de balisage

- Hiérarchie

- Information de randonnées

- Fiches techniques et informatives

Transformations des langages

- Transformation sans contrôle

- Transformation avec contrôle

Conclusion

INTRODUCTION

- Evolution du web

- Evolution du web
- Nécessité de structurer et modéliser les contenus du web de la manière la plus indépendante possible

- Evolution du web
- Nécessité de structurer et modéliser les contenus du web de la manière la plus indépendante possible
- Manipulation de données hétérogènes voire irrégulières exprimés en XML

- Evolution du web
- Nécessité de structurer et modéliser les contenus du web de la manière la plus indépendante possible
- Manipulation de données hétérogènes voire irrégulières exprimés en XML
- Etude de cas: Conception de langages de balisages permettant la modélisation d'un type spécifique de données

- Evolution du web
- Nécessité de structurer et modéliser les contenus du web de la manière la plus indépendante possible
- Manipulation de données hétérogènes voire irrégulières exprimés en XML
- Etude de cas: Conception de langages de balisages permettant la modélisation d'un type spécifique de données
- Documents relatifs à des circuits de randonnée

LES BESOINS À MODÉLISER

- Informations de randonnée
- Hiérarchie/Structure des documents
- Fiches d'information et fiches techniques

CONCEPTION DES LANGAGES DE BAL- ISAGE

- Balisage fil de l'eau/Balisage naturel

- Balisage fil de l'eau/Balisage naturel
 - Avantage : Aucune contrainte particulière

- Balisage fil de l'eau/Balisage naturel
 - Avantage : Aucune contrainte particulière
 - Inconvénients: Difficultés rencontrées sur le traitement de certaines données (requêtes de type bd)

- Balisage fil de l'eau/Balisage naturel
 - Avantage : Aucune contrainte particulière
 - Inconvénients: Difficultés rencontrées sur le traitement de certaines données (requêtes de type bd)
- Balisage type bd/Balisage structuré Necessité de conserver la même organisation des balises

- Balisage fil de l'eau/Balisage naturel
 - Avantage : Aucune contrainte particulière
 - Inconvénients: Difficultés rencontrées sur le traitement de certaines données (requêtes de type bd)
- Balisage type bd/Balisage structuré Nécessité de conserver la même organisation des balises
 - Avantage: Possibilité de faire aisément des requêtes de type BD

- Balisage fil de l'eau/Balisage naturel
 - Avantage : Aucune contrainte particulière
 - Inconvénients: Difficultés rencontrées sur le traitement de certaines données (requêtes de type bd)
- Balisage type bd/Balisage structuré Necessité de conserver la même organisation des balises
 - Avantage: Possibilité de faire aisément des requêtes de type BD
 - Inconvénients: Perte d'information, contrainte de structure

- Balisage fil de l'eau/Balisage naturel
 - Avantage : Aucune contrainte particulière
 - Inconvénients: Difficultés rencontrées sur le traitement de certaines données (requêtes de type bd)
- Balisage type bd/Balisage structuré Nécessité de conserver la même organisation des balises
 - Avantage: Possibilité de faire aisément des requêtes de type BD
 - Inconvénients: Perte d'information, contrainte de structure
- Problème :Langage naturel vers Langage structuré \Rightarrow Perte d'informations

- Balisage fil de l'eau/Balisage naturel
 - Avantage : Aucune contrainte particulière
 - Inconvénients: Difficultés rencontrées sur le traitement de certaines données (requêtes de type bd)
- Balisage type bd/Balisage structuré Necessité de conserver la même organisation des balises
 - Avantage: Possibilité de faire aisément des requêtes de type BD
 - Inconvénients: Perte d'information, contrainte de structure
- Problème :Langage naturel vers Langage structuré ⇒ Perte d'informations
- Exemple: tourner le dos au <lieu>gave de Pau</lieu> et prendre ⇒ ...<lieu>gave de Pau</lieu>...

- Modélisation des différentes parties qui constituent le document

- Modélisation des différentes parties qui constituent le document
- Conservation de l'intégralité du document

- Modélisation des différentes parties qui constituent le document
- Conservation de l'intégralité du document
- Préambule, description, thématique, synthèse, titres, sous-titres, listes,...

- Modélisation des différentes parties qui constituent le document
- Conservation de l'intégralité du document
- Préambule, description, thématique, synthèse, titres, sous-titres, listes,...
- Avantage: Structuration précise et spécifique (s'applique bien aux documents étudiés)

- Modélisation des différentes parties qui constituent le document
- Conservation de l'intégralité du document
- Préambule, description, thématique, synthèse, titres, sous-titres, listes,...
- Avantage: Structuration précise et spécifique (s'applique bien aux documents étudiés)
- Inconvénient: ne peut structurer que le jeu de documents (jurançon, 7villages, Crabere,...) , modélisation pas assez raffinée

- Modélisation des différentes étapes qui constitue la randonnée

- Modélisation des différentes étapes qui constitue la randonnée
- L'intégralité du document n'est pas conservée

- Modélisation des différentes étapes qui constitue la randonnée
- L'intégralité du document n'est pas conservée
- Lieux, description, distance, durée, ...

- Modélisation des différentes étapes qui constitue la randonnée
- L'intégralité du document n'est pas conservée
- Lieux, description, distance, durée, ...
- Avantage: Permet des recherches temporelles et géographiques (exemple : Randonnée de moins d'une heure)

- Modélisation de fiches recapitulatives

- Modélisation de fiches recapitulatives
- Informations retenues: Fiche technique et fiche d'information (listes d'item)

- Modélisation de fiches recapitulatives
- Informations retenues: Fiche technique et fiche d'information (listes d'item)
- fiches technique, fiche d'information

- Modélisation de fiches recapitulatives
- Informations retenues: Fiche technique et fiche d'information (listes d'item)
- fiches technique, fiche d'information
- Avantage: Permet un résumé rapide, ajoute des information supplémentaires aux autres langages

- Modélisation de fiches recapitulatives
- Informations retenues: Fiche technique et fiche d'information (listes d'item)
- fiches technique, fiche d'information
- Avantage: Permet un résumé rapide, ajoute des information supplémentaires aux autres langages
- Inconvénient: On ne peut modéliser que deux types de fiches

TRANSFORMATIONS DES LANGAGES

- Langage Fil de l'eau ► langage BD : XSLT version 1

- Langage Fil de l'eau ► langage BD : XSLT version 1
- Avantage: Pas de contraintes à respecter, large choix de transformations possibles

- Langage Fil de l'eau ► langage BD : XSLT version 1
- Avantage: Pas de contraintes à respecter, large choix de transformations possibles
- Inconvénient: Inexactitude quand a la conservation des données depuis l'entrée vers la sortie

- lors du passage du langage de type BD vers l'application de sortie (L^AT_EX)

- lors du passage du langage de type BD vers l'application de sortie (\LaTeX)
- Contrôle effectué en entrée et pas en sortie (En sortie contrôle par la compilation \LaTeX)

- lors du passage du langage de type BD vers l'application de sortie (\LaTeX)
- Contrôle effectué en entrée et pas en sortie (En sortie contrôle par la compilation \LaTeX)
- Contrôle sur le contenu et les attributs des balises par les schemas XML

- lors du passage du langage de type BD vers l'application de sortie (\LaTeX)
- Contrôle effectué en entrée et pas en sortie (En sortie contrôle par la compilation \LaTeX)
- Contrôle sur le contenu et les attributs des balises par les schemas XML
- Contrôle sur les attributs des balises également grâce aux DTDs

- lors du passage du langage de type BD vers l'application de sortie (\LaTeX)
- Contrôle effectué en entrée et pas en sortie (En sortie contrôle par la compilation \LaTeX)
- Contrôle sur le contenu et les attributs des balises par les schemas XML
- Contrôle sur les attributs des balises également grâce aux DTDs
-

- lors du passage du langage de type BD vers l'application de sortie (\LaTeX)
- Contrôle effectué en entrée et pas en sortie (En sortie contrôle par la compilation \LaTeX)
- Contrôle sur le contenu et les attributs des balises par les schemas XML
- Contrôle sur les attributs des balises également grâce aux DTDs
-
- Limites:

- lors du passage du langage de type BD vers l'application de sortie (\LaTeX)
- Contrôle effectué en entrée et pas en sortie (En sortie contrôle par la compilation \LaTeX)
- Contrôle sur le contenu et les attributs des balises par les schemas XML
- Contrôle sur les attributs des balises également grâce aux DTDs
-
- Limites:
 - Etablissement d'un ordre sur une séquence de valeurs numériques en occurrence sur les identifiants des différentes étapes d'un circuit de randonnée

- lors du passage du langage de type BD vers l'application de sortie (\LaTeX)
- Contrôle effectué en entrée et pas en sortie (En sortie contrôle par la compilation \LaTeX)
- Contrôle sur le contenu et les attributs des balises par les schemas XML
- Contrôle sur les attributs des balises également grâce aux DTDs
-
- Limites:
 - Etablissement d'un ordre sur une séquence de valeurs numériques en occurrence sur les identifiants des différentes étapes d'un circuit de randonnée
 - XML Schema généré à partir d'une DTD avec une erreur d'ambiguïté quand la DTD contient des élément admettant plusieurs définitions possibles; Solution: Ecriture plus complexe de la DTD

CONCLUSION

- Découverte de la véritable utilisation du meta-langage XML et des multiples possibilités de conception de langages

- Découverte de la véritable utilisation du meta-langage XML et des multiples possibilités de conception de langages
- Limites rencontrées quand à l'utilisation des DTD pour structurer les

- Découverte de la véritable utilisation du meta-langage XML et des multiples possibilités de conception de langages
- Limites rencontrées quand à l'utilisation des DTD pour structurer les
- Possi