

INU3011 Documents structurés

Cours 12

Chaînes de traitement XML
Création de documents XML
Conférencière invitée

Plan

- Préparation au Mini-quiz 3
- Chaînes de traitement XML (*workflow*)
 - Méthodes de création de documents XML
 - *Single-source publishing*
- Conférence Edith Cannel

Préparation au Mini-quiz 3 (1/3)

- Principes de modélisation (C6)
 - Par exemple :
 - Balisage logique
 - Découpage adéquat
 - [050-Ex-utilite-validite](#) (messages-telephoniques-XSLT)
- XPath
 - Modèle de données XPath
 - Exemples détaillés [200-XPath](#)
 - Exercices [exerc-XPath](#)

Préparation au Mini-quiz 3 (2/3)

- XSLT
 - Exemples [240-Ex-XSLT-pas-a-pas](#)
 - 1 et 2
 - Exercices XSLT de base [exerc-XSLT-1](#)
 - *Pas l'Exercice XSLT avancé exerc-XSLT-2*
- Correspondances HTML ↔ LaTeX (C10)

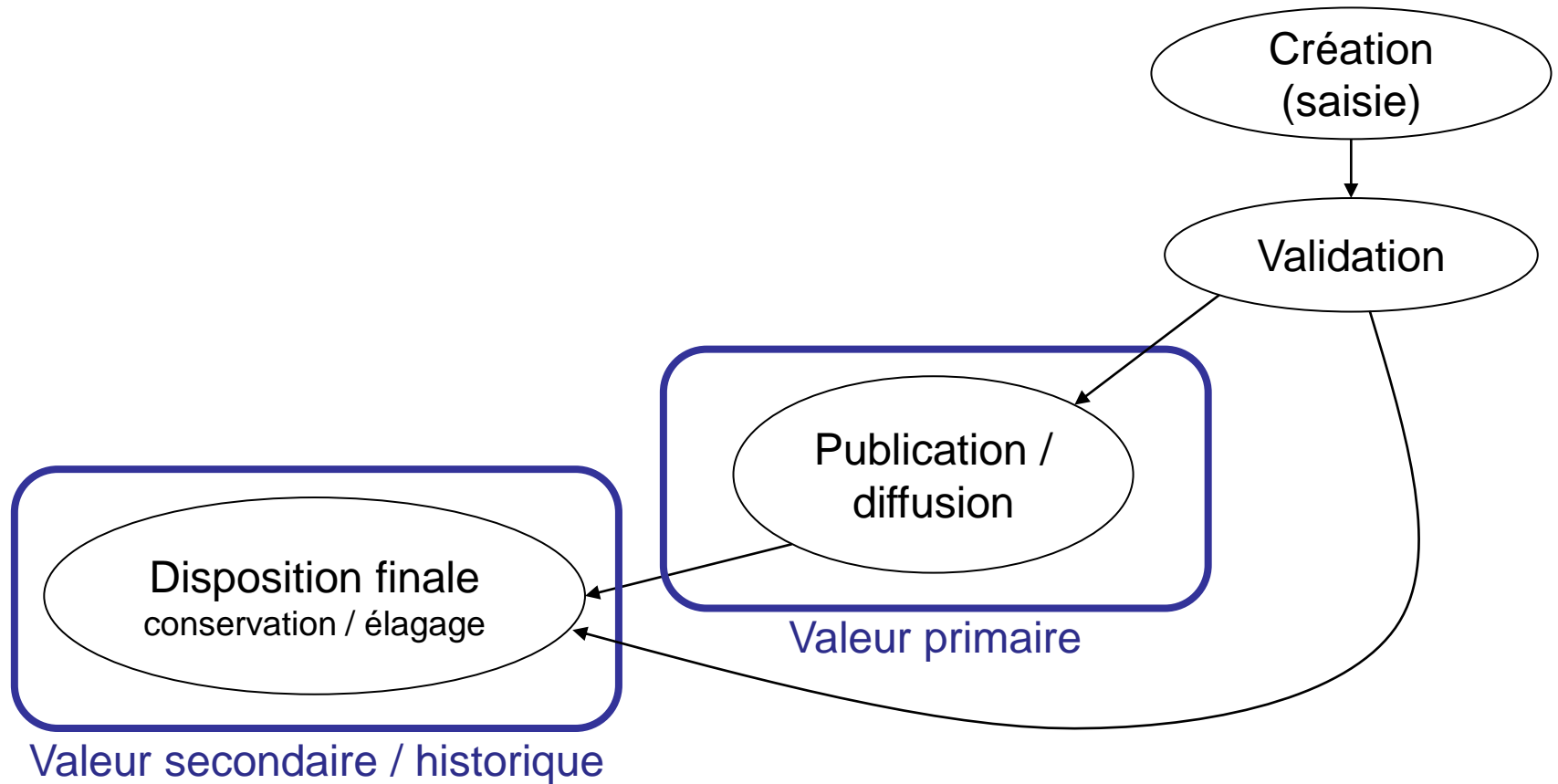
Préparation au Mini-quiz 3 (3/3)

- Cours 11 :
 - Utilité / objectif / avantages de :
 - Espaces de noms XML (*XML Namespaces*)
 - Modèles normalisés
 - Caractéristiques générales des langages de contraintes autres que DTD
 - Surtout Schémas XML du W3C
- [StudiUM](#)

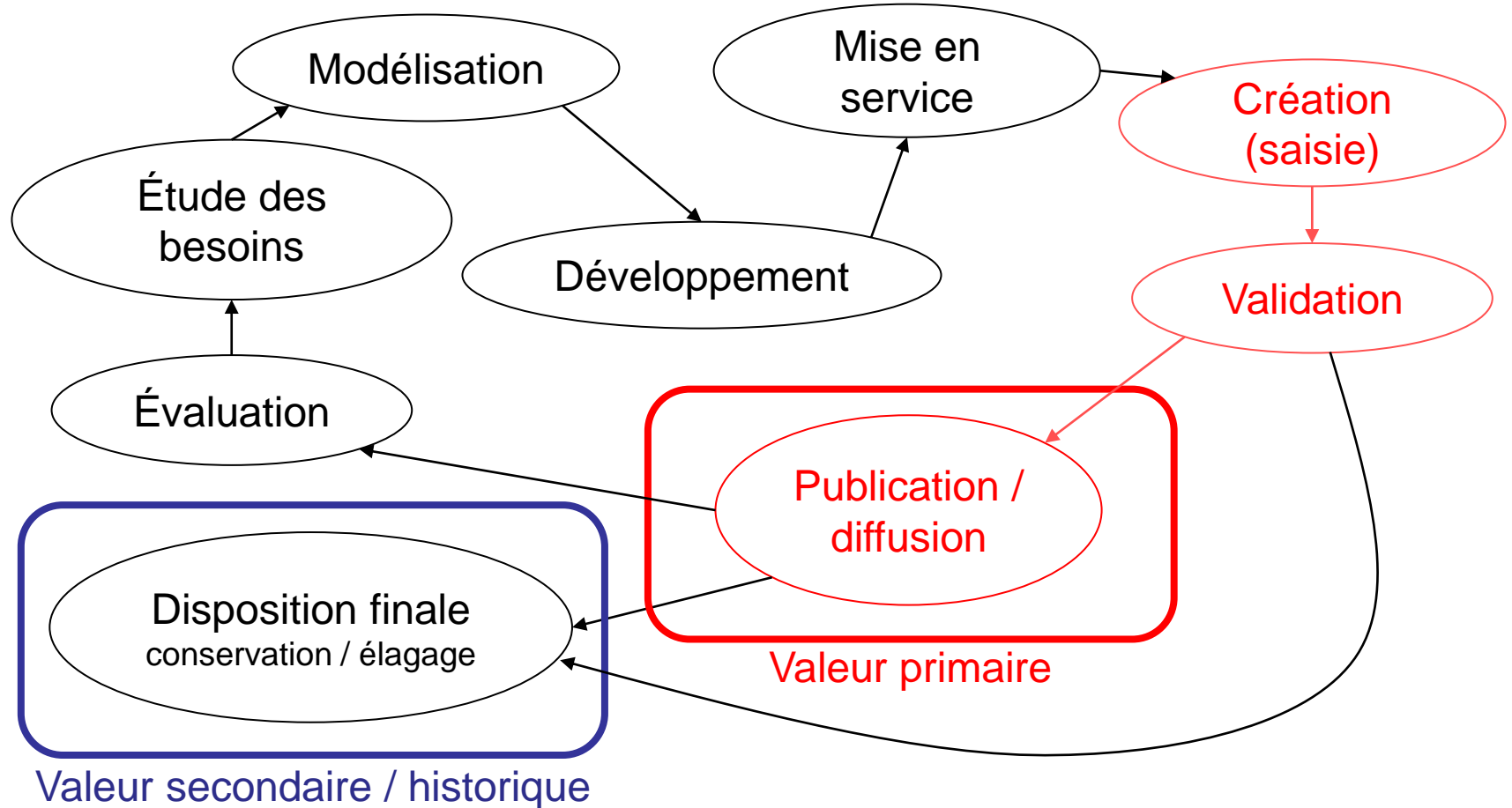
Chaînes de traitement XML

Que se passe-t-il en aval de la
modélisation ?

Cycle de vie "archivistique" des documents



Cycle de développement "archivistique intégrée"



Création des documents (1/3)

- Éditeur XML
- Conversions à partir de bases de données ou autres logiciels spécialisés
 - Bibliographique, catalographique
 - Instruments de recherche archivistiques

Création des documents (2/3)

- Saisie Web
 - Formulaire interactif
 - Exemple 600-Ex-saisie-Web-AutAbsence
 - XSLT + JavaScript
 - XForms 1.1
 - [W3C Recommendation 20 October 2009](#)
 - Peut s'intégrer au HTML, comme les formulaires actuels (`<form>`)

Création des documents (3/3)

- "Up-conversions"
 - Enrichissement du balisage basé sur la détection de patrons (*patterns*)
 - Souvent avec traitement de texte + styles
 - FLATXML (export à partir de Libre Office)
 - Au moins partiellement automatisable
 - Exemple [700-Ex-up-conversion](#)

[Up-conversion using XSLT 2.0 \(de Michael Kay\)](#)

Validation

- Particulièrement importante pour documents dérivés d'autres formats
- Intégrée dans la création si éditeur XML
- En général validation dupliquée sur réception des documents à l'entrée dans le système de traitement

Stockage

- Dossiers de fichiers distincts
- Bases de données natives XML
 - Ex.: BaseX, eXist, MarkLogic Server
- Bases de données "+ XML"
 - Ex.: Microsoft SQL Server, MySQL

Publication / diffusion quelques normes

- Onix (industrie du livre), basé XML
- ePub (livre électronique), basé XML
- Jats-NLM, basé XML

- Les transformations sont le plus souvent en XSLT

Single-source publishing

- Une seule source XML sert à alimenter plusieurs canaux de diffusion
- Chaque canal requiert un format de document particulier
 - Ou des documents XML conformes à un schéma particulier
- Exemple : projet du cours : source XML unique, diffusion en HTML et PDF

Conférence

- Conférencière : Madame Edith Cannet
 - Éditrice (CNRS - Institut des sciences humaines et sociales)
 - ingénierie éditoriale
 - humanités numériques
 - Pôle Document numérique – Maison de la Recherche en Sciences Humaines – Université de Caen (France)

