

→ Sémantique et interopérabilité

Antoine ISAAC

Vrije Universiteit Amsterdam

Koninklijke Bibliotheek

aisaac@few.vu.nl

Journée d'étude BnF / AFNOR, 28 Mars 2008



Contexte

- CATCH (NWO)
 - Continuous Access To Cultural Heritage
 - Projets de recherche appliqués au domaine patrimonial
- STITCH
 - SemanTic Interoperability To access Cultural Heritage
 - Échange et intégration de méta-données

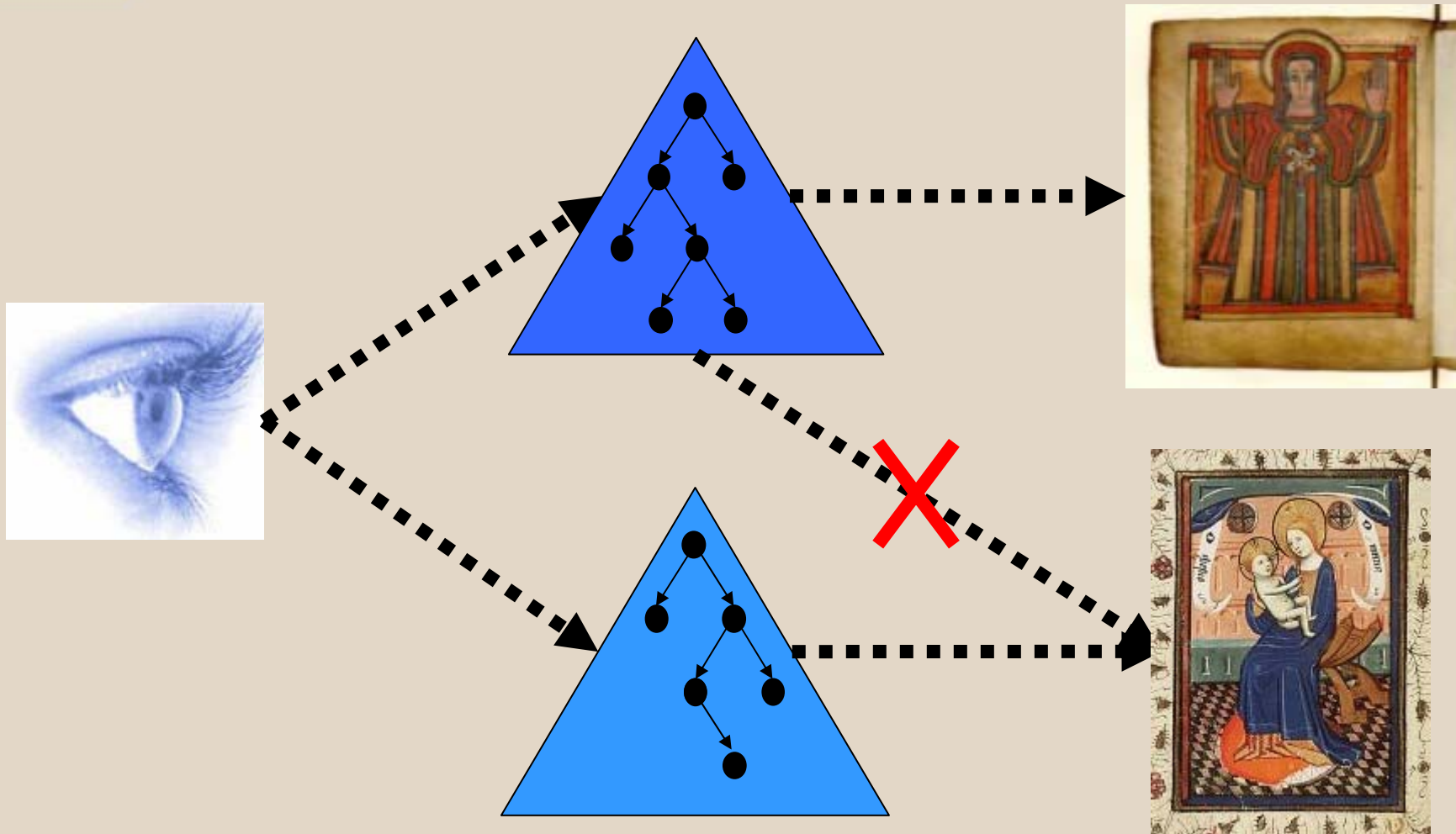


Plan

- Démo
- Intérêt du web sémantique
- Représenter des référentiels sur le web sémantique
- Aligner des référentiels

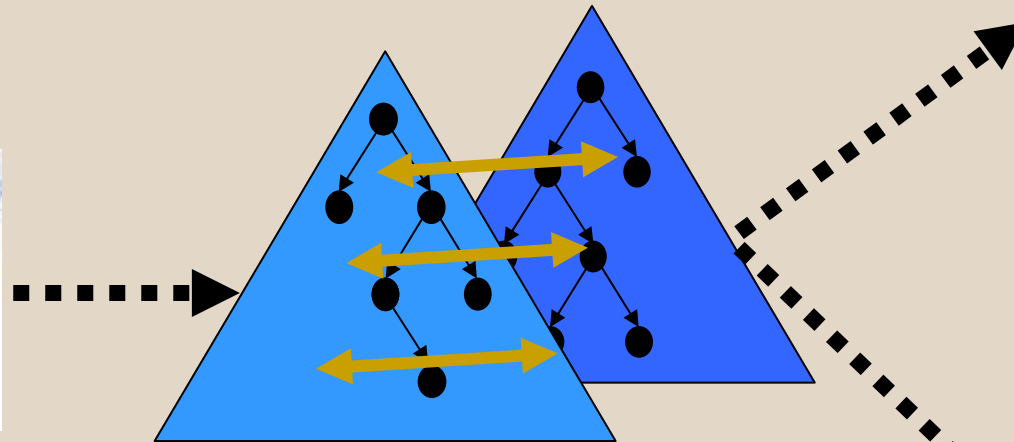


Objectifs de la démo





Objectifs de la démo





Démo

- Collaboration avec la BnF
- KB Illuminated Manuscripts
- Base Mandragore de la BnF
- <http://galjas.cs.vu.nl:33333/MANDRA-SV-ICE-mandraNewNONE> , amphibians
- [Blé](#)



Comment faire fonctionner une telle démo ?

Problème de l'interopérabilité au niveau sémantique

- Des référentiels conceptuels distincts
- Des modèles différents: *classes, concepts, termes*, etc.

Les bases de la solution employée :

- Une infrastructure web sémantique
- Un modèle standard pour les référentiels
 - Notion de Knowledge Organization System (KOS)



Plan

- Démo
- Intérêt du web sémantique
- Représenter des référentiels sur le web sémantique
- Aligner des référentiels



Qu'est-ce que le web sémantique ?

- Promu par le World Wide Web Consortium
Tim Berners-Lee
<http://www.w3.org/2001/sw/>
- “Le web sémantique est un web de données”
- “Il s’agit de proposer des formats communs pour intégrer des données provenant de sources différentes”



Le web sémantique (1/2)

- Pointer des **ressources**
 - **Éléments de connaissance**
 - Tout ce à quoi on peut vouloir faire référence
 - documents, personnes, sujets...
 - **Uniform Resource Identifiers**
 - URLs: <http://www.few.vu.nl/~aisaac/>
 - urn:isbn:0-395-36341-1



Un web de ressources

myVoc1:Article

<http://ex.org/files/file1>

myVoc2:Amsterdam

阿姆斯特丹，荷兰首都，人口约74万。

目录 [隐藏]

- 1 历史
 - 1.1 城市建立和初期发展
 - 1.2 黄金时代
 - 1.3 工业革命
 - 1.4 两次世界大战
 - 1.5 战后重建
 - 1.6 现代的城市
- 2 交通
- 3 文化
 - 3.1 體育活動
- 4 旅游景点
- 5 友好城市
- 6 外部链接

历史

阿姆斯特丹的历史最早可以追溯到13世纪时的渔村。人们曾在附近**阿姆斯特河**（Anstel）上建筑水坝，阿姆斯特丹就得名于此。原来的名字“Anstelredam”，意指“阿姆斯特水坝”。17世纪是阿姆斯特丹历史上的“黄金时代”。阿姆斯特丹是当时世界上最重要的港口和银行业中心。

城市建立和初期发展

在上个千年之初，一些冒险者垂青由挖空的

阿姆斯特丹 Amsterdam	
	
<i>别名: Anstelredan</i>	
國家	荷蘭
省份	北荷兰省
现任市長	Job Cohen
面積	219平方公里
—城市面積	166平方公里
—土地面積	53平方公里
—水域面積	(2006年1月1日)
人口	743,027 (2006年1月1日)
—總人口	4,459/平方公里
—人口密度	
時區	UTC +1
位置	
—經度	4° 54' E
—緯度	52° 22' N
	

<http://www.ned.nl/rep321>

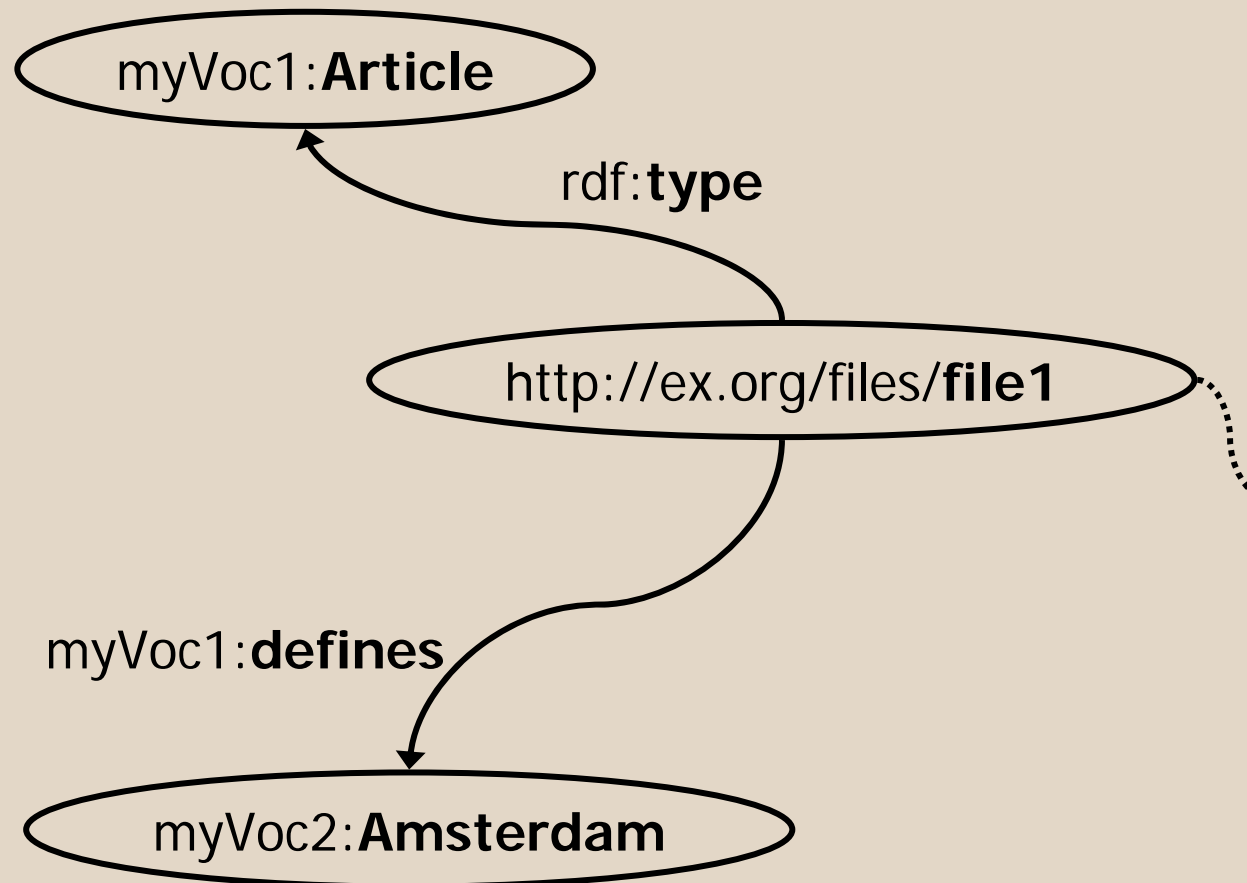


Le web sémantique (2/2)

- Pointer des ressources : URIs
- Créer des assertions structurées concernant ces ressources
 - **Liens typés** entre les ressources
 - **RDF (Resource Description Framework)**



Données sous forme de "graphe" RDF



阿姆斯特丹，荷兰首都，人口约74万。

目录 [隐藏]

- 1 历史
 - 1.1 城市建立和初期发展
 - 1.2 黄金时代
 - 1.3 工业革命
 - 1.4 两次世界大战
 - 1.5 战后重建
 - 1.6 现代的城市
- 2 交通
- 3 文化
 - 3.1 體育活動
- 4 旅游景点
- 5 友好城市
- 6 外部链接

历史

阿姆斯特丹的历史最早可以追溯到13世纪时的渔村。人们曾在附近阿姆斯特河（Amstel）上建筑水坝，阿姆斯特丹就得名于此。原来的名字“Amstelredan”，意指“阿姆斯特水坝”。17世纪是阿姆斯特丹历史上的“黄金时代”。阿姆斯特丹是当时世界上最重要的港口和银行业中心。

城市建立和初期发展

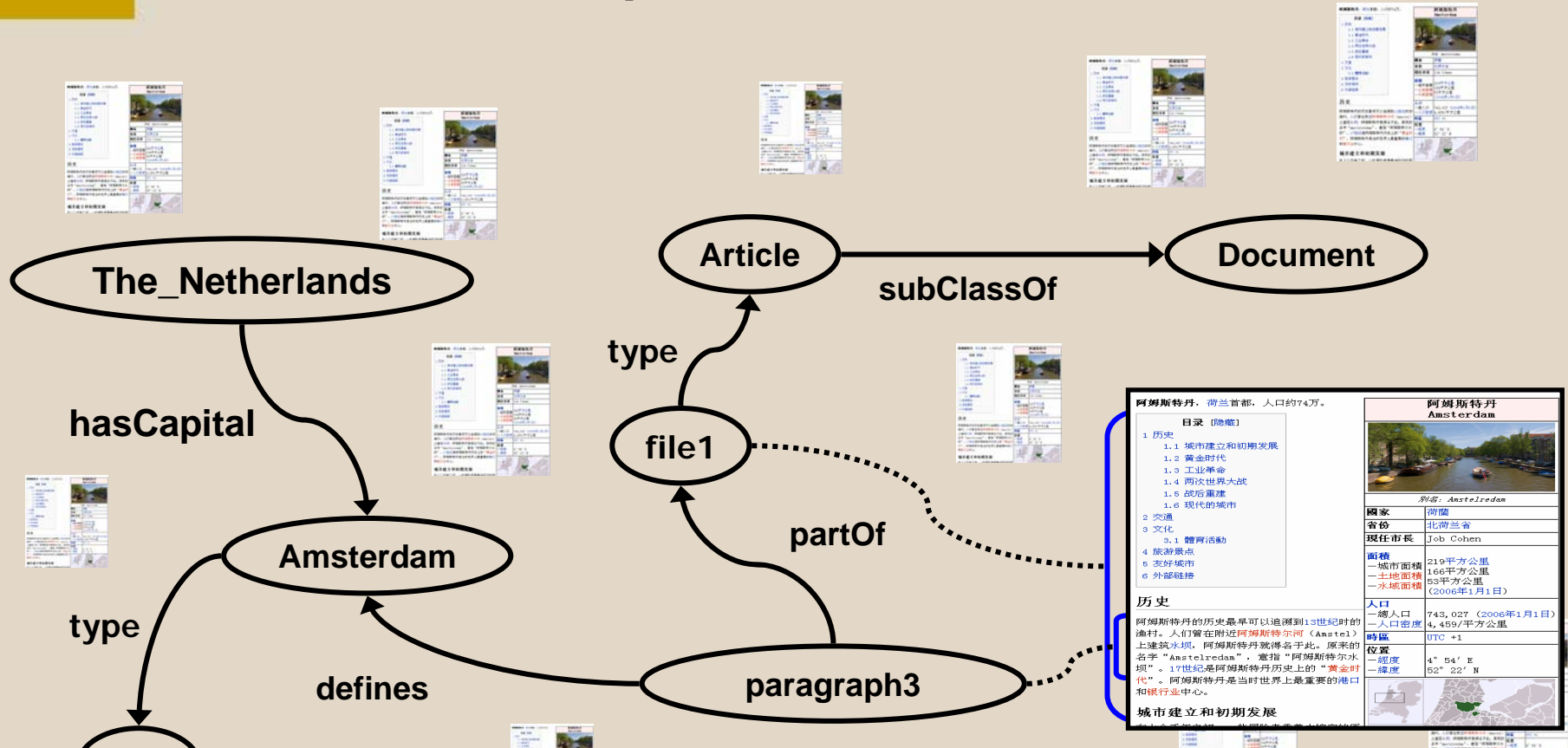
在上个千年之初，一些冒险者垂垂由控空的原

阿姆斯特丹 Amsterdam	
	
别名: <i>Amstelredan</i>	
國家	荷蘭
省份	北荷兰省
現任市長	Job Cohen
面積	
—城市面積	219平方公里
—土地面積	166平方公里
—水域面積	53平方公里 (2006年1月1日)
人口	
—總人口	743,027 (2006年1月1日)
—人口密度	4,459/平方公里
時區	UTC +1
位置	
—經度	4° 54' E
—緯度	52° 22' N
	

<http://www.ned.nl/rep321>



Le web sémantique : un web de (méta-)données



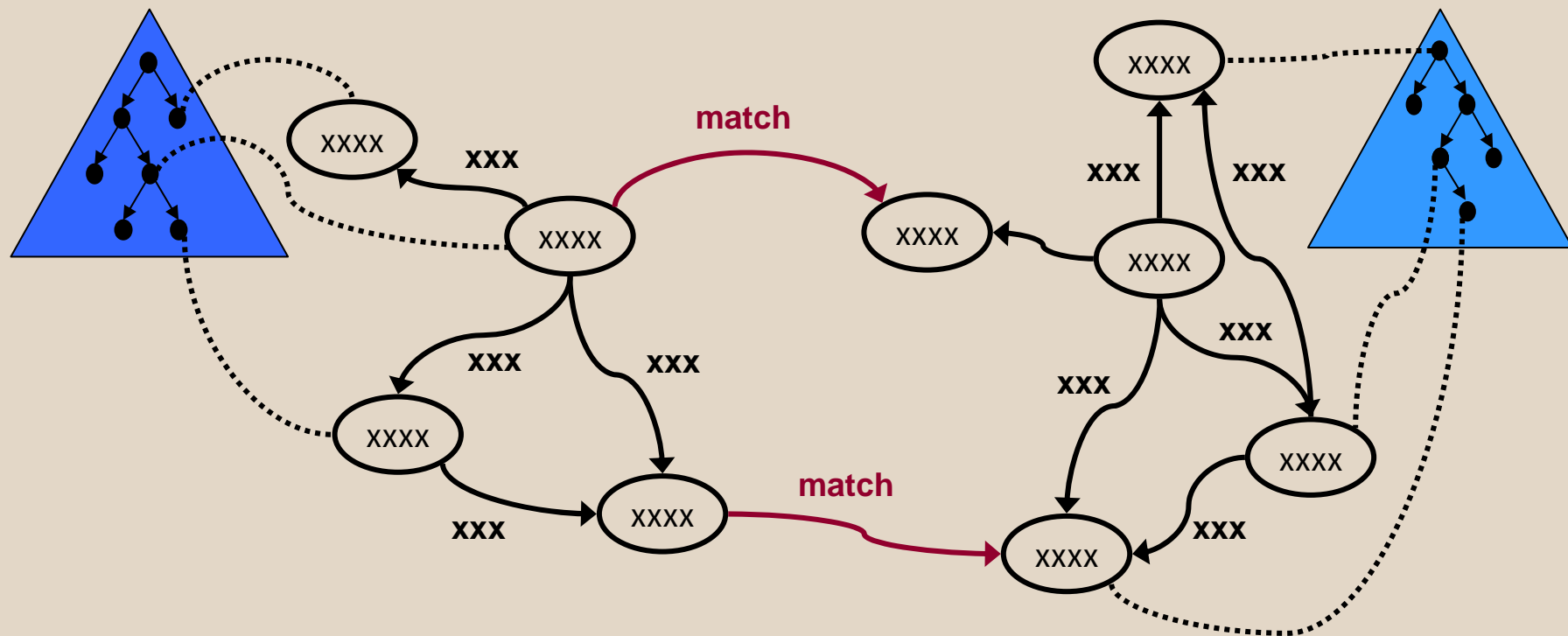
Utiliser des ressources web facilite la distribution et le partage

- des vocabulaires de description
- des (méta-)données



Quel est le rapport avec les référentiels?

- Représenter les référentiels sur le web sémantique
- Les aligner (mettre en correspondance)





Plan

- Démo
- Intérêt du web sémantique
- Représenter des référentiels sur le web sémantique
- Aligner des référentiels



SKOS (Simple Knowledge Organization System)

- Modèle simple et standard pour représenter des KOSs sur le web sémantique
 - Rôle semblable à celui de Dublin Core, mais pour des concepts
 - Inspiré par les thésaurus, mais convient à d'autres référentiels
- Attention
 - SKOS est simple: on peut perdre de l'information !
 - But: publier de l'information sur des référentiels
 - pas de remplacer des langages comme MARC
- En cours de finition au W3C

<http://www.w3.org/2004/02/skos/>

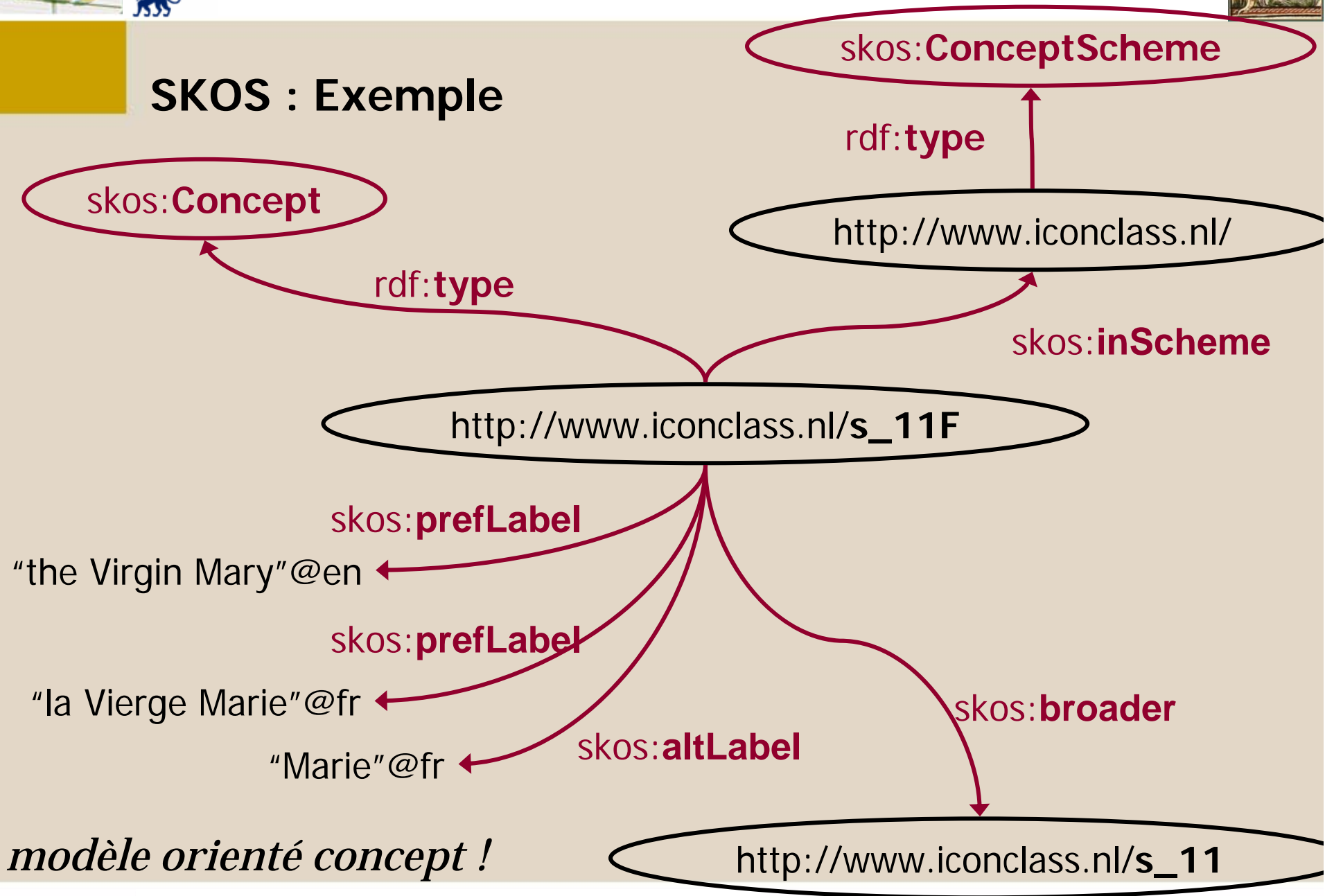


SKOS (Simple Knowledge Organization System)

- SKOS: un modèle pour représenter des KOSs en RDF
 - Types d'objets: **Concept** and **ConceptScheme**
 - Propriétés lexicales (multilingues)
 - `prefLabel`
 - `altLabel`
 - Liens sémantiques
 - `broader`, `narrower`
 - `related`
 - Notes
 - `scopeNote`
 - `definition`



SKOS : Exemple





Plan

- Démo
- Intérêt du web sémantique
- Représenter des référentiels sur le web sémantique
- Aligner des référentiels



Alignement de référentiels

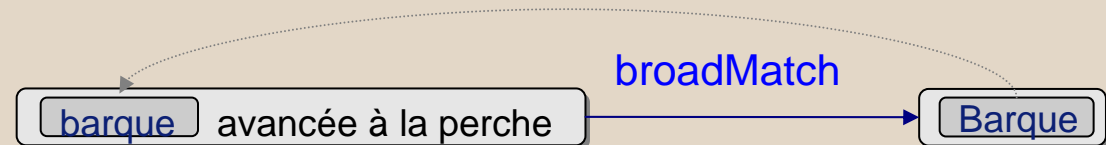
- But: trouver des correspondances sémantiques
 - “ruines antiques” \approx “paysage avec ruines”
 - “Sainte Marie” = “la Vierge Marie”
- Le faire (semi-)automatiquement
- S’inspirer des techniques d’alignement d’ontologies



Techniques d'alignement automatiques

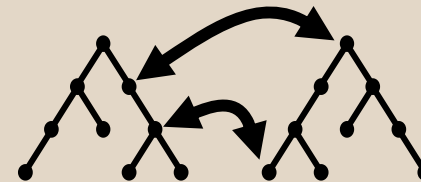
- Lexicales

Libellés des concepts, définitions



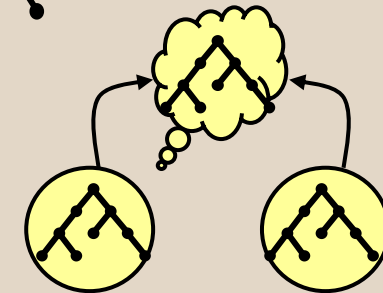
- Structurelles

Structure des vocabulaires



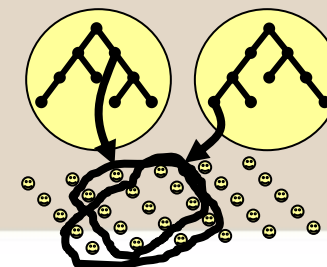
- Connaissances de contexte

Exploitation des liens d'une référence partagée



- Extensionnelles

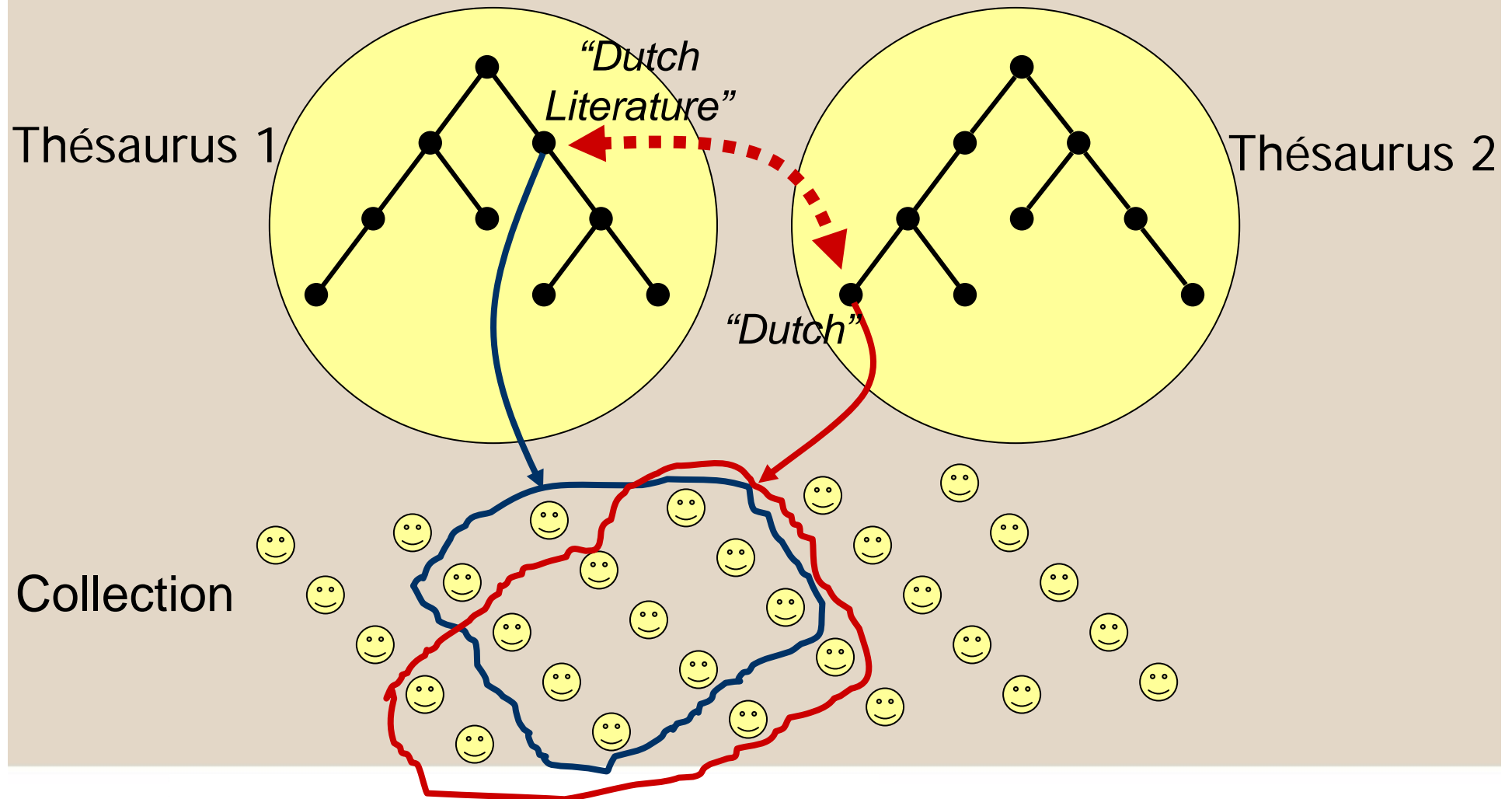
Information attachée aux objets (par. ex. livres)





Techniques d'alignement statistiques

- Utilisant l'indexation des objets





Alignement: pas de solution triviale

- Ces techniques permettent d'obtenir des résultats
- Mais seules, elles ne suffisent pas
 - Combinaison avec travail manuel (vérification, complétion)
 - Combinaison de plusieurs techniques
- Note: la qualité d'un alignement dépend aussi du type d'application qui l'emploie !
 - Reformulation de requêtes, fusion de référentiels...
 - Pose la question de l'utilisation d'un référentiel pour une application pour laquelle il n'a pas été conçu
- L'alignement est un problème de recherche difficile



Plan

- Démo
- Intérêt du web sémantique
- Représenter des référentiels sur le web sémantique
- Aligner des référentiels



Message

Les techniques du web sémantique

- **Representation** des référentiels
- **Alignment**

Peuvent aider à résoudre des problèmes cruciaux

- Publication et **accès**
- Mise en **correspondance** sémantique
 - Sans passer forcément par une fusion des vocabulaires ou une reformulation des descriptions

En fait, inter-fécondation : les référentiels peuvent contribuer à la vision du web sémantique



Merci !



Liens

- Portail web sémantique du W3C
 - <http://www.w3.org/2001/sw/>
- SKOS
 - <http://www.w3.org/2004/02/skos/>
- Projets articulants web sémantique et domaine patrimonial
 - MuseumFinland, <http://www.museosuomi.fi/>
 - eCulture, <http://e-culture.multimedien.nl/>
 - STITCH, <http://www.cs.vu.nl/STITCH/>
 - CATCH, <http://www.nwo.nl/catch>