



L'ingénierie des documents d'affaires dans le cadre du Web Sémantique

Jamel Eddine Jridi

Supervisé par Guy Lapalme

11 Juillet 2012



- Échange électronique de documents.
- Grande quantité de documents en B2B:
 - Des relevés des transactions,
 - Des bons de commandes,
 - Des contrats, etc.
- Différentes branches d'affaires.
- Plusieurs définitions des documents.

Introduction





Problématique

- « *Reference* » :
 - (1) Transaction de paiement
 - (2) Numéro de la commande
- « *Total* » :
 - (1) Montant TTC
 - (2) Montant sans taxes

BON DE COMMANDE 1

PROMO. Inc.
 Address: Montréal, QC

Order no 445882 Order date Dec 10, 2010 Sub-total \$198.00 Taxes \$25.75 Shipping \$11.95 Credits \$10.00 **Total \$225.70**



Deisha Sunglasses in Gray
 Size: One Size
 Price: \$99.00
 SKU: GIOGARMANI560SQMB



Deena Sunglasses in Silver and Blue
 Size: One Size
 Price: \$99.00
 SKU: GIOGARMANI599SVHV

Name: Alain
 Address: 2222, Vigi Avenue,
 Montréal, QC
 Paid via: Master Card
Reference: 0102986473622

Reference: 445882
 Order Placed: Dec 10, 2010

Item	Item Code	Description	Price
	GIOGARMANI560SQMB	Deisha Sunglasses in Gray Quantity: 1	\$99
	GIOGARMANI599SVHV	Deena Sunglasses in Silver and Blue Quantity: 1	\$99

Total:	\$198.00	Client: Alain
Postage and Handling:	\$11.95	Shipping Adress:
Taxes:	\$25.75	2222, Vigi Avenue, Montréal, QC
Order Payment:	\$225.70	
Payment Method:	Master Card	
Transaction Number:	0102986473622	2



Contribution

- Utilisation de la même définition d'informations.
- Comment les informations et les données sont utilisées et définies ?
- Description: structure + sémantique + relation entre les concepts.
- Rendre possible à la machine d'interpréter les données.
- Création d'une ontologie partagée entre les applications d'affaire.

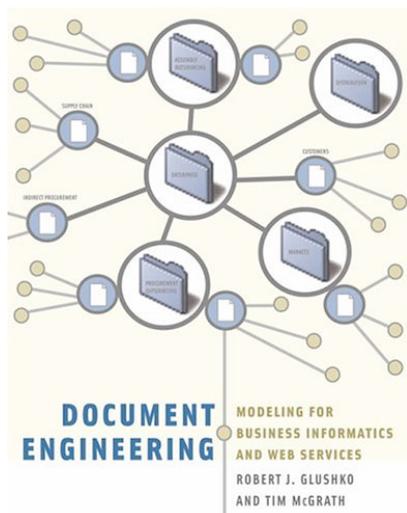
Ingénierie de documents



Définition de l'ingénierie de documents

“ L'ingénierie de documents est une nouvelle discipline pour la spécification, la conception et la mise en œuvre des documents échangés électroniquement. Cette discipline est appliquée dans plusieurs domaines afin de faciliter la communication et gérer la grande masse d'informations qui circulent dans/et en dehors d'un organisme. ”

[Glushko et al. 2005]



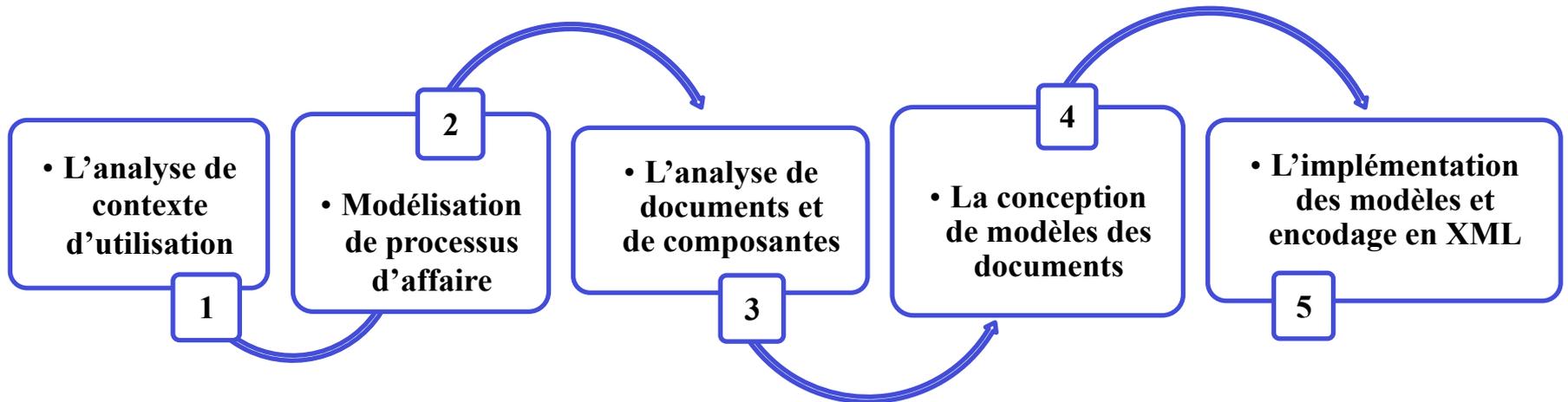


Processus d'ingénierie de documents

- Basé sur des méthodes d'analyse et de conception:
 - Analyse de processus d'affaire.
 - Analyse des tâches.
 - Analyse de documents.
 - Analyse de données.



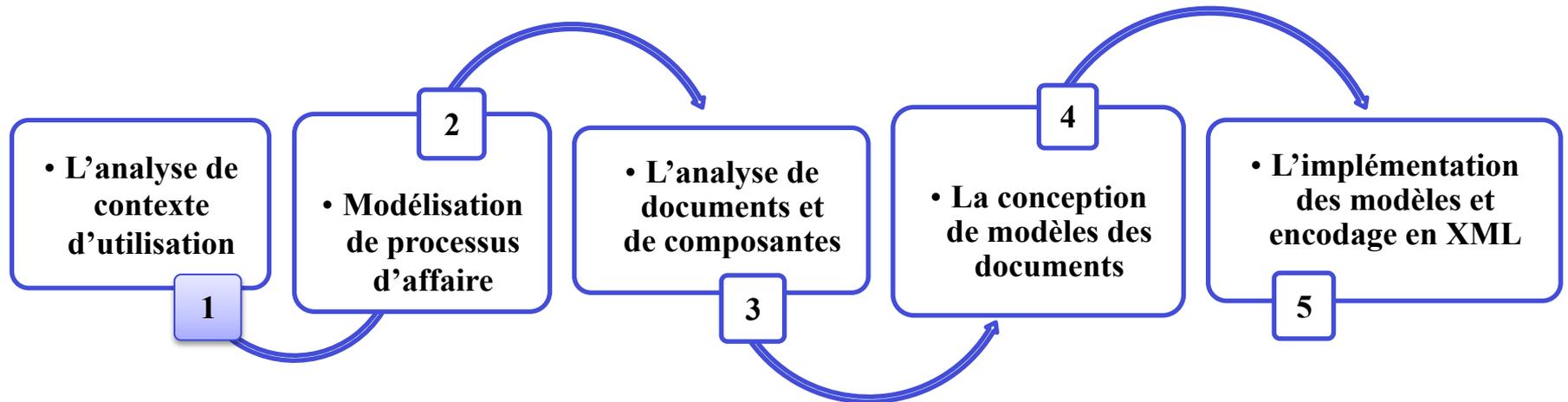
Processus d'ingénierie de documents



[Glushko et al. 2005]

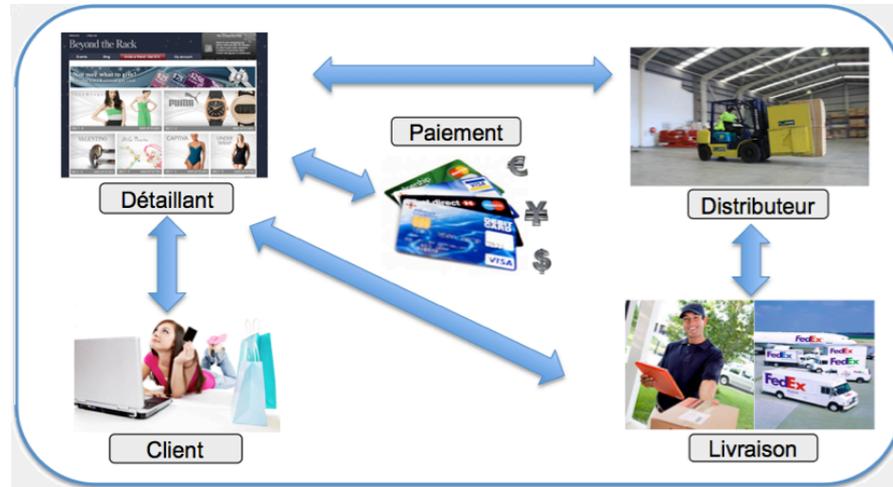


Processus d'ingénierie de documents



[Glushko et al. 2005]

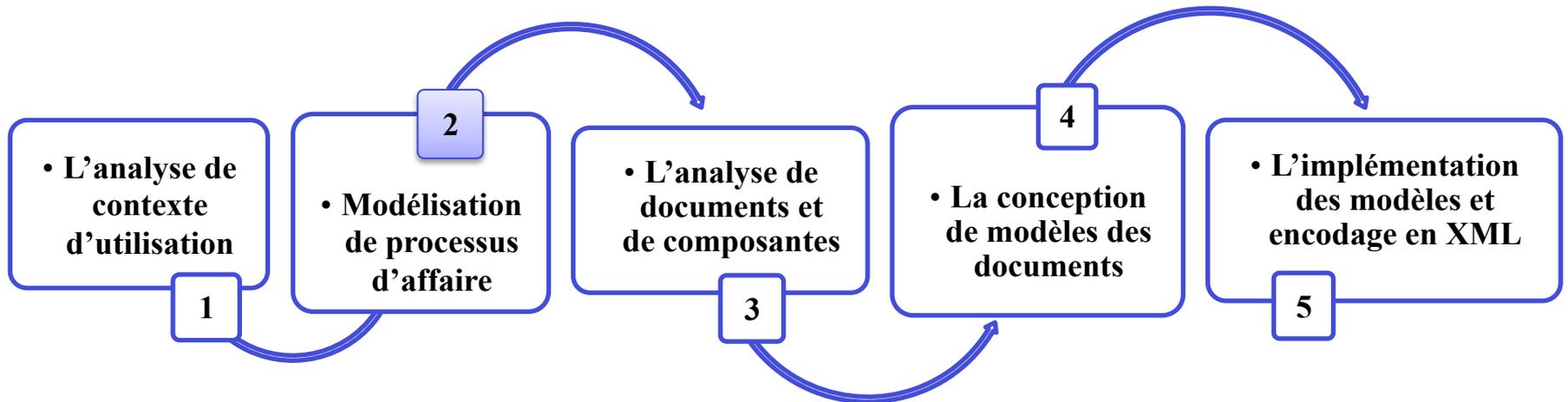
Étape 1: Analyse de contexte d'utilisation



- Contexte général: processus de vente en ligne.
- Étapes: réalisation de la commande + paiement + traitement de la commande + livraison.
- Participants: client + détaillant + livreur + distributeur (l'entreprise).



Processus d'ingénierie de documents



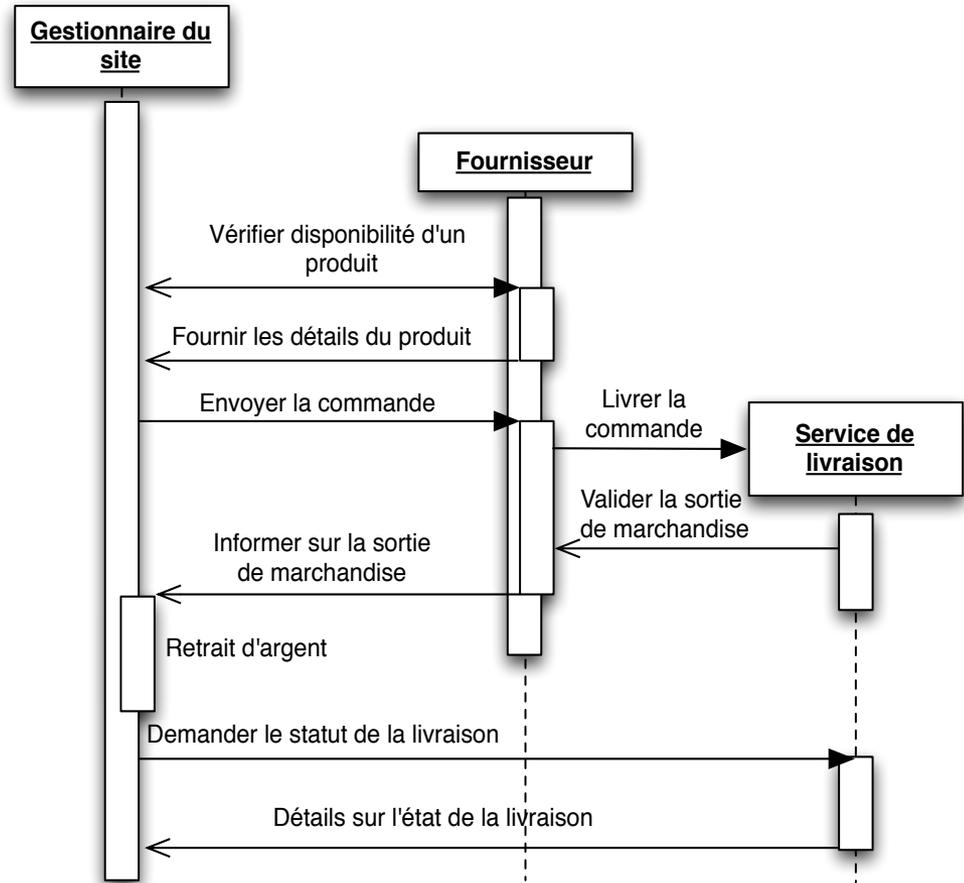
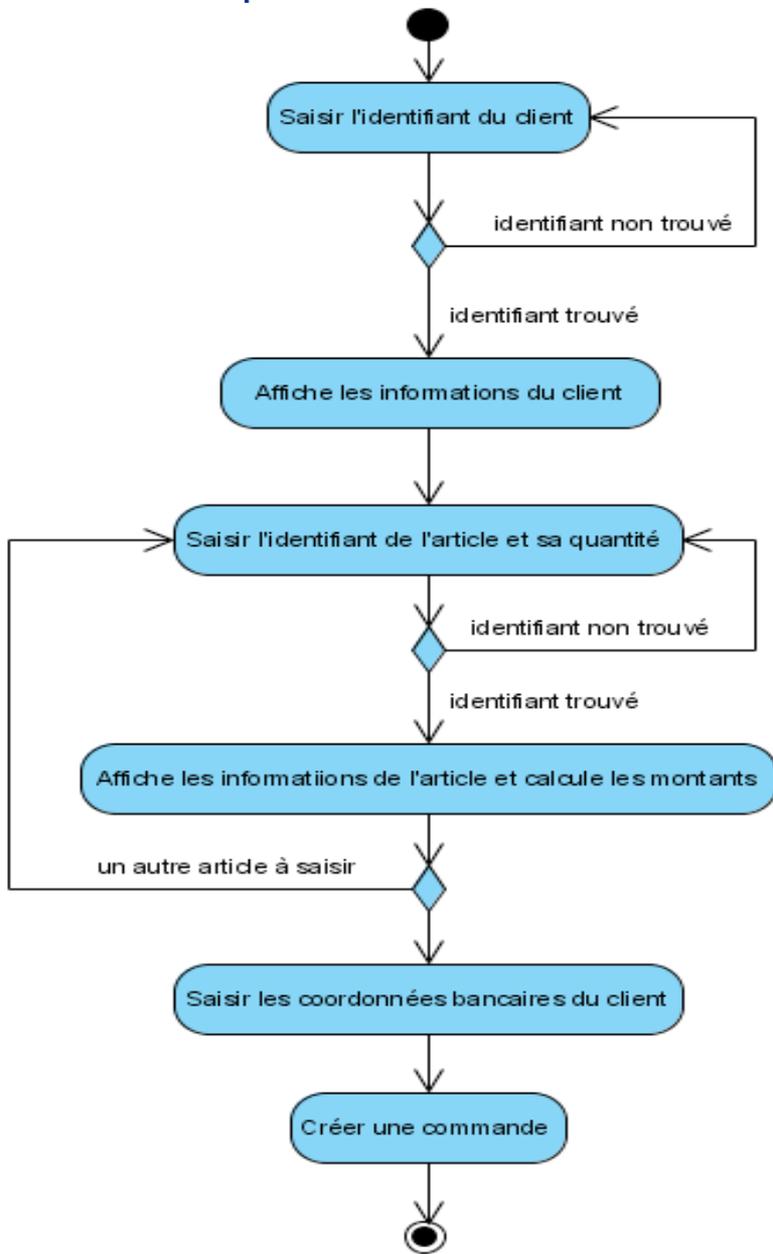
[Glushko et al. 2005]



Étape 2: Modélisation du processus d'affaire (1/2)

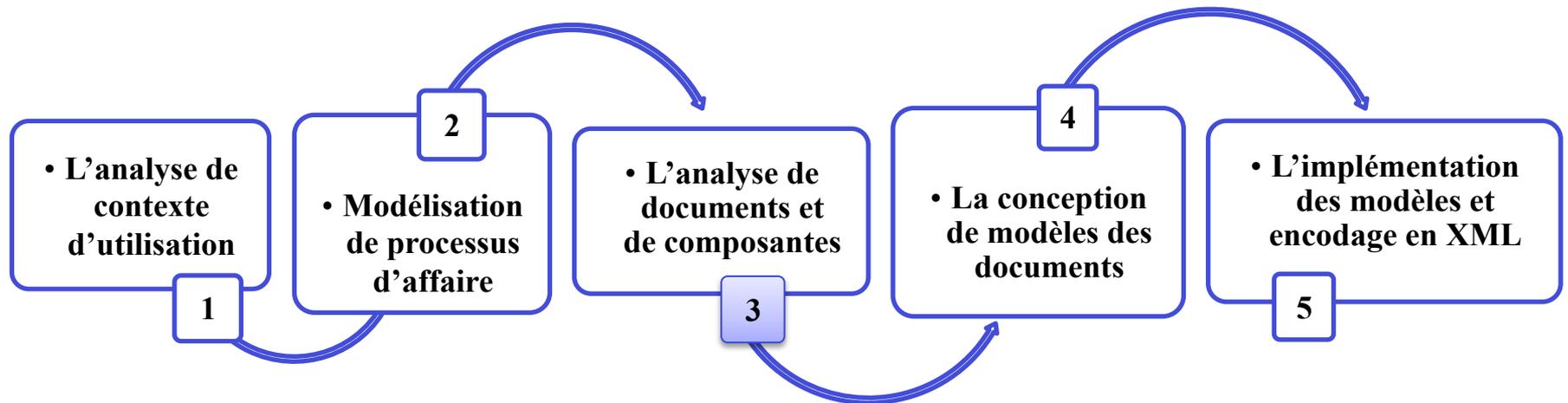
Questions	Réponses
Quel est le nom du processus?	Processus de vente en ligne.
Quel est l'objectif de ce processus?	But: gestion de vente en ligne. <ul style="list-style-type: none">• Choix des produits et paiements.• Traitement de la commande.• Livraison de la commande.
Quelles sont les industries, les organisations impliquées dans le processus?	<ul style="list-style-type: none">• Services de cartes de crédit.• Fournisseur.• Sociétés de livraison.
Quelles sont les intervenants ou les participants au processus?	<ul style="list-style-type: none">• Client (acteur principal).• Gestionnaire du site.• Distributeur (le fournisseur)• Livreur (FedEx, UPS, etc.)

Étape 2: Modélisation du processus d'affaire (2/2)





Processus d'ingénierie de documents



[Glushko et al. 2005]

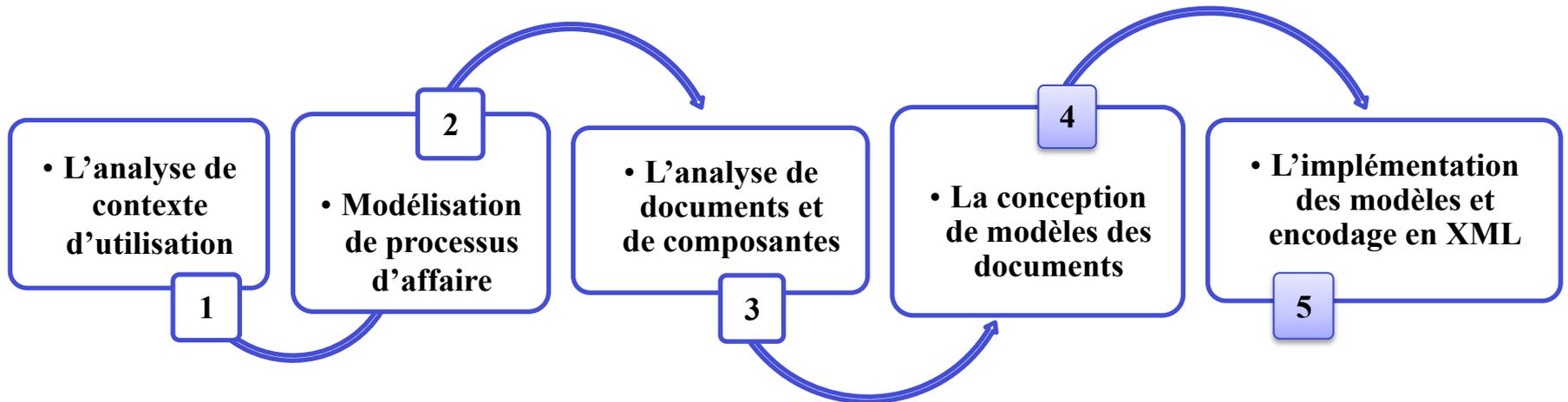


Étape 3: L'analyse de documents et de composantes

- Analyse des documents basée sur:
 - Le contenu,
 - La structure,
 - La présentation et la sémantique des informations.
- Création d'un modèle conceptuel.
 - Organiser les composantes (dépendance fonctionnelle).
 - Optimiser la réutilisation.
 - Minimiser la redondance.



Processus d'ingénierie de documents



[Glushko et al. 2005]



Étape 4 & 5: La conception et l'implémentation des modèles

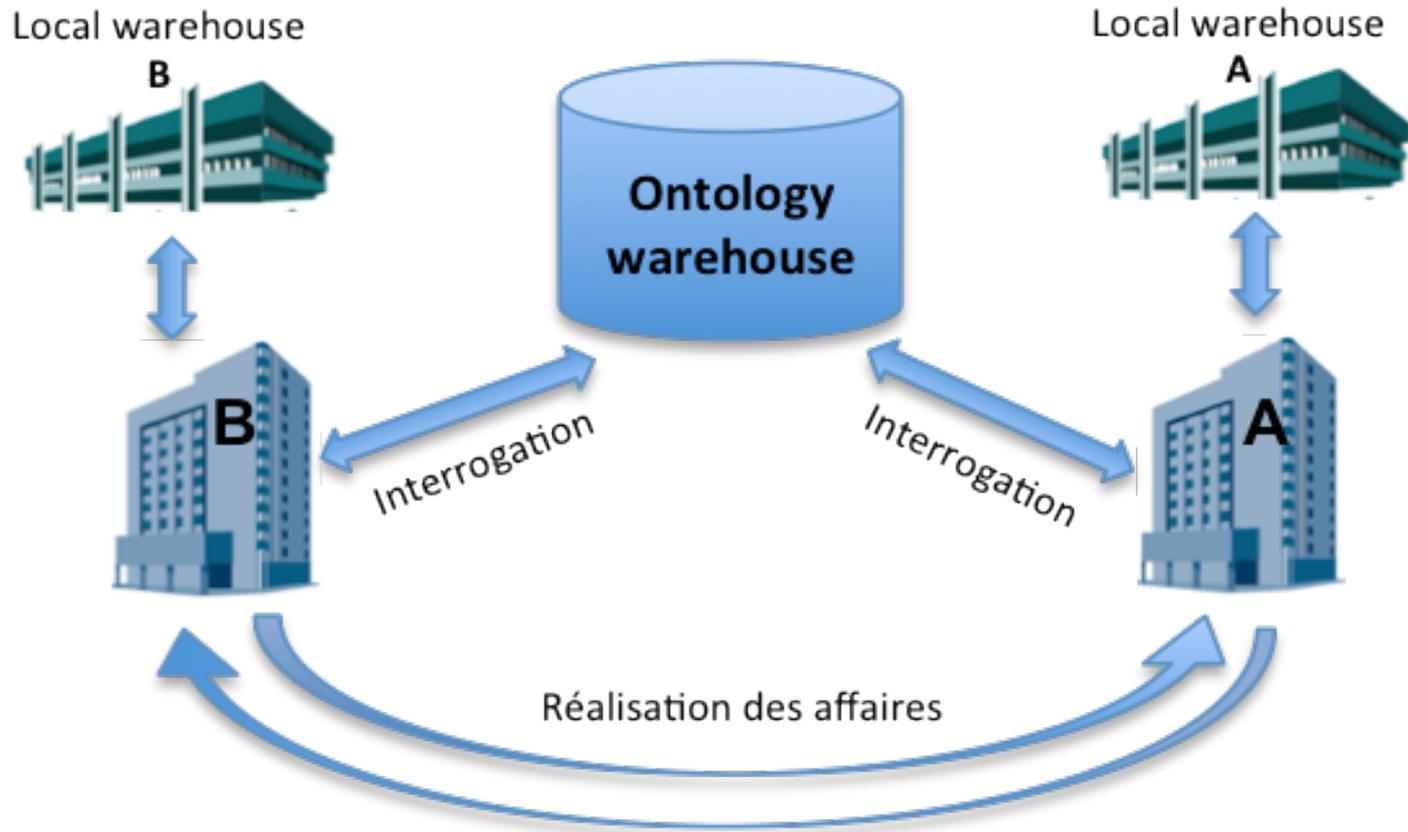
- Étape d'analyse \implies Étape de conception = Modèle conceptuel
- Modèle conceptuel (liens entre les patterns) = Diagramme de classes
- Modèle conceptuel \implies **encodage en XML = DTD, Schéma XML**



Standards en e-business

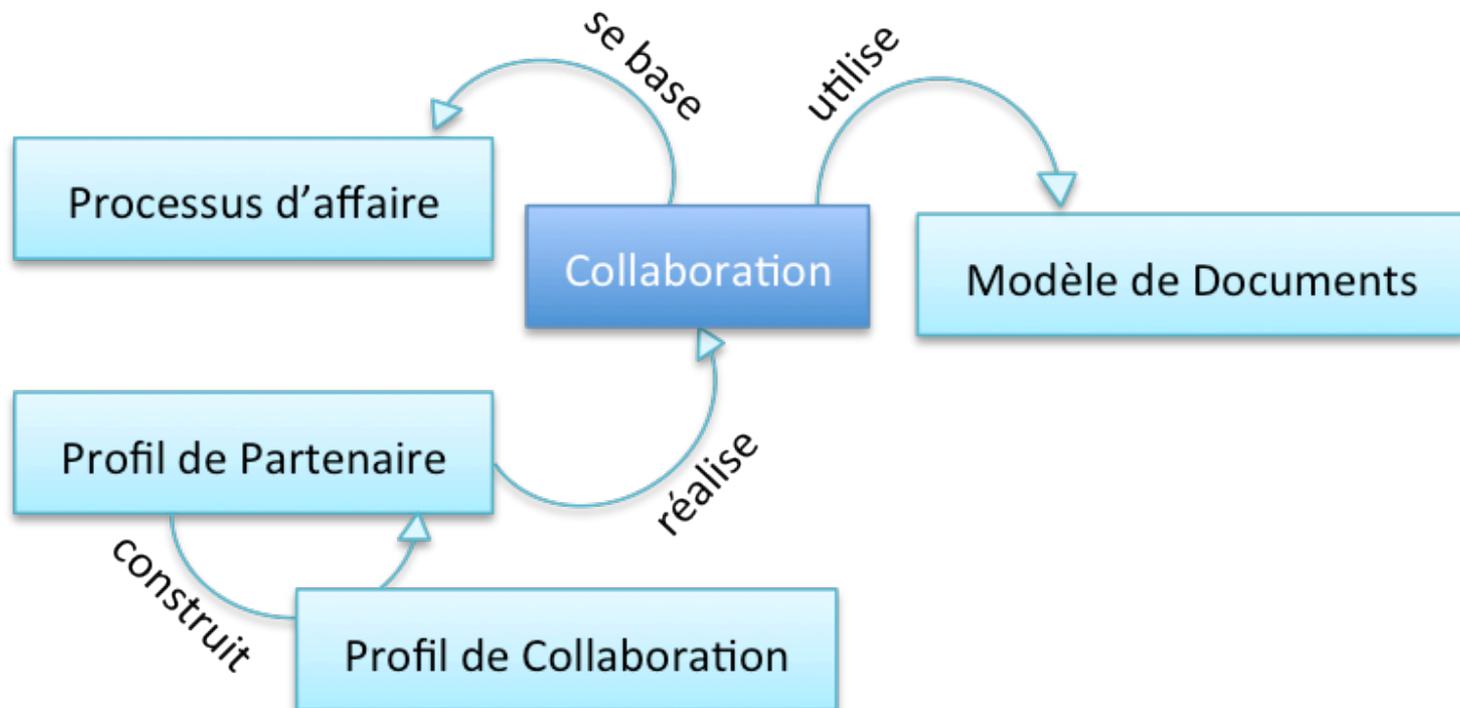
- XML based standards:
 - xCBL
 - RosettaNet [Kantola et al., 2002]
 - ebXML [Huemer et al., 2002]
- Languages formels basés sur la sémantique (WSML, OWL-S)
- Intégration des ontologies
 - ebXML [Dogac et al., 2004]
 - RosettaNet [Kotinurmi et al., 2009]

Architecture ebOWL





Modèle des concepts ebOWL





Pourquoi OWL ?



- Recommandé en 2004.
- Formaliser un domaine en définissant des classes et propriétés de ces classes.
- Définition des combinaisons logiques (union, intersection, etc).
- Définir des individus et affirmer des propriétés les concernant.
- Reasonner sur ces classes et individus.

```

1 <!-- RosettaNet XML Message Schema
2 3A4_MS_R02_00_PurchaseOrderRequest.dtd (19-Apr-2001 10:
3 This document has been prepared by Edifecs (http://www.
4 based On the Business Collaboration Framework from requi
5 in conformance with the RosettaNet methodology. -->
6 <!ENTITY % common-attributes "id CDATA #IMPLIED" >
7 <!ELEMENT Pip3A4PurchaseOrderRequest (
8 fromRole ,
9 GlobalDocumentFunctionCode ,
10 PurchaseOrder ,
11 thisDocumentGenerationDateTime ,
12 thisDocumentIdentifier ,
13 toRole ) >
14 <!ELEMENT fromRole ( PartnerRoleDescription ) >
15 <!ELEMENT PartnerRoleDescription (
16 ContactInformation? ,
17 GlobalPartnerRoleClassificationCode
18 , PartnerDescription ) >
19 <!ELEMENT ContactInformation (
20 contactName?, EmailAddress?,
21 facsimileNumber?, telephoneNumber?,
22 PhysicalAddress? ) >
23 <!ELEMENT contactName ( FreeFormText ) >
24 <!ELEMENT FreeFormText ( #PCDATA ) >
25 <!ATTLIST FreeFormText xml:lang CDATA #IMPLIED >
26 <!ELEMENT EmailAddress ( #PCDATA ) >
27 <!ELEMENT facsimileNumber ( CommunicationsNumber ) >
28 <!ELEMENT CommunicationsNumber ( #PCDATA ) >
29 <!ELEMENT telephoneNumber ( CommunicationsNumber ) >
30 <!ELEMENT GlobalPartnerRoleClassificationCode ( #PCDATA ) >
31 <!ELEMENT PartnerDescription (
32 BusinessDescription ,
33 GlobalPartnerClassificationCode? ,
34 PhysicalAddress? ,
35 ContactInformation? ) >
36 <!ELEMENT BusinessDescription (
37 GlobalBusinessIdentifier? ,
38 GlobalSupplyChainCode? ,
39 businessName? ,
40 NationalBusinessTaxIdentifier? ) >
41 <!ELEMENT GlobalBusinessIdentifier ( #PCDATA ) >
42 <!ELEMENT GlobalSupplyChainCode ( #PCDATA ) >
43 <!ELEMENT GlobalPartnerClassificationCode ( #PCDATA ) >
44 <!ELEMENT GlobalDocumentFunctionCode ( #PCDATA ) >
45 <!ELEMENT PurchaseOrder (

```

46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89

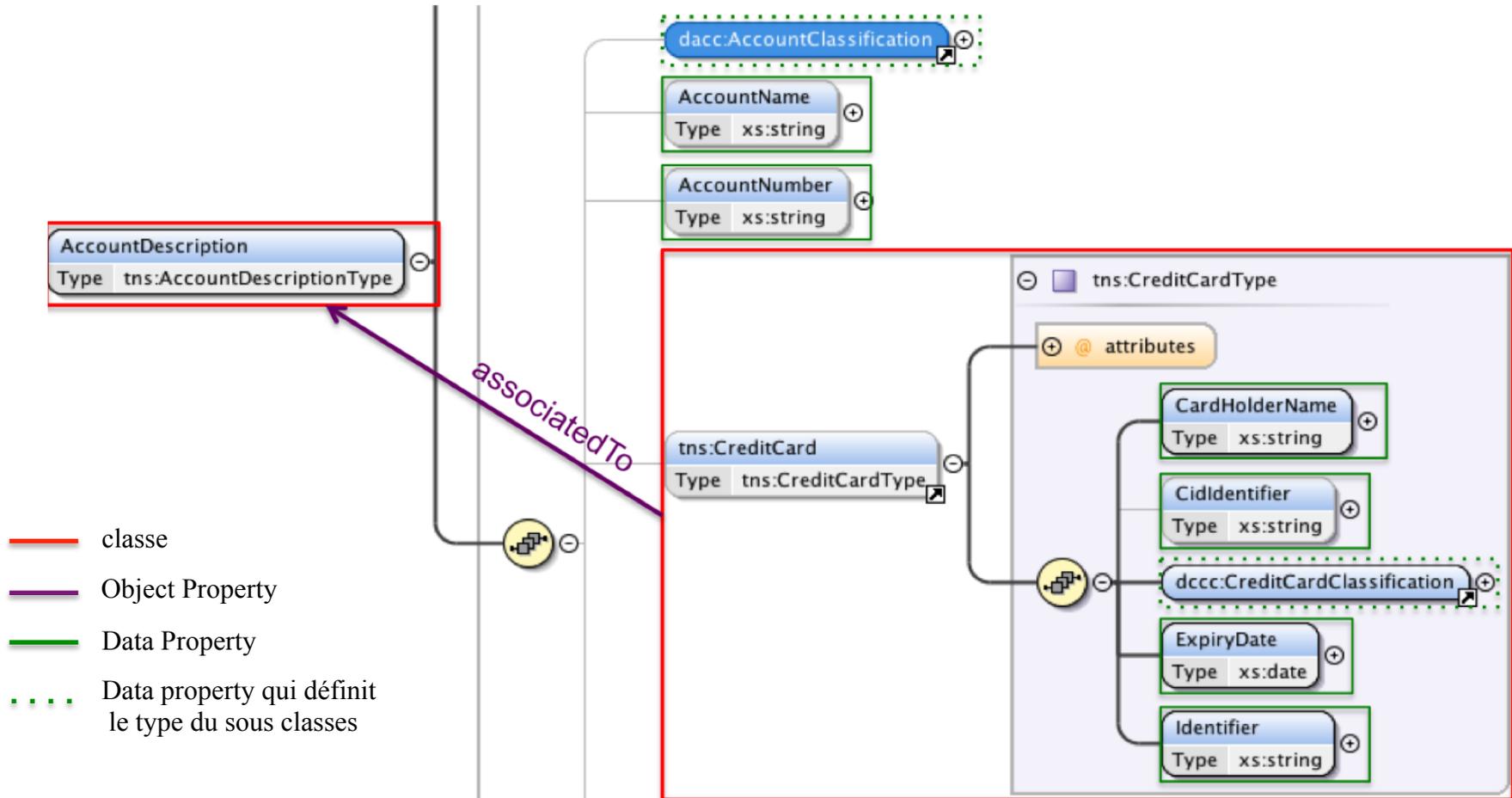
```

AccountDescription? , comments? ,
ContractInformation* , DocumentReference* ,
FinancingTerms* ,
generalServicesAdministrationNumber? ,
GlobalGovernmentPriorityRatingCode? ,
GlobalPurchaseOrderFillPriorityCode? ,
GlobalPurchaseOrderTypeCode+ ,
governmentContractIdentifier? ,
installAt? ,
OrderShippingInformation? ,
ProductLineItem+ ,
proprietaryInformation* ,
purchaseOrderIdentifier ,
requestedEvent* , requestedShipFrom* ,
SecondaryBuyer? , shipTo ,
TaxExemptStatus? , totalAmount? ) >
<!ELEMENT AccountDescription (
accountName , AccountNumber? ,
billTo , CreditCard? ,
financedBy? ,
GlobalAccountClassificationCode? ,
prePaymentCheckNumber? ,
WireTransferIdentifier? ) >
<!ELEMENT accountName ( FreeFormText ) >
<!ELEMENT AccountNumber ( #PCDATA ) >
<!ELEMENT billTo ( PartnerRoleDescription ) >
<!ELEMENT PhysicalAddress (
addressLine1? , addressLine2? ,
addressLine3? , cityName? ,
GlobalCountryCode? ,
GlobalLocationIdentifier? ,
NationalPostalCode? ,
postOfficeBoxIdentifier? , regionName? ) >
<!ELEMENT addressLine1 ( FreeFormText ) >
<!ELEMENT addressLine2 ( FreeFormText ) >
<!ELEMENT addressLine3 ( FreeFormText ) >
<!ELEMENT CreditCard (
cardHolderName , creditCardIdentifier ,
expiryDate , GlobalCreditCardClassificationCode ,
ProprietaryCIDIdentifier? ) >
<!ELEMENT cityName ( FreeFormText ) >
<!ELEMENT GlobalCountryCode ( #PCDATA ) >
<!ELEMENT GlobalLocationIdentifier ( #PCDATA ) >
<!ELEMENT NationalPostalCode ( #PCDATA ) >

```



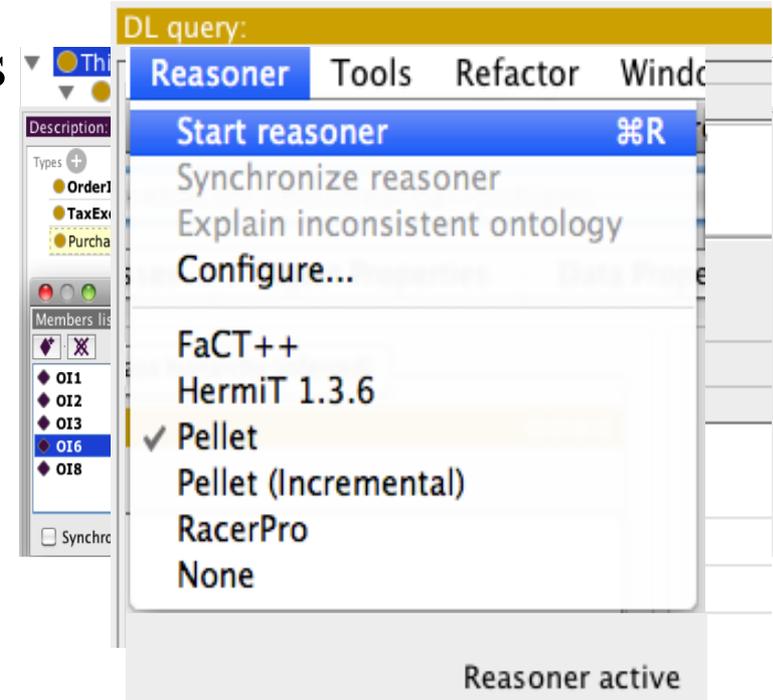
Règles de transformation RosettaNet PIP3A4: XML à OWL 2





Étapes de construction

- ① Définition des concepts (classes, propriétés)
- ② Définition des propriétés complexes
 - Data Property: entre classe et littéral
 - Object Property: relie deux classes.
- ③ Création des individus.
- ④ Vérification de la consistance.
- ⑤ Validation avec des requêtes en LD.





Conclusion

- Problème d'ambiguïté sémantique.
- Partage et échanges de documents efficaces.
- Ontologie partagée.
- OWL based approach (ebOWL).
- e-Government, medicine, e-Learning



Échéancier

Hiver 2011: Début de la thèse et prédoc écrit.

Été 2011: Définition de la problématique.

Automne 2011: Cours.

Hiver 2012: Cours + prédoc spécialité.

Été 2012: Transformation automatique des PIPs.

Automne 2012: Réalisation du prototype.

Hiver 2013: Conception d'un système d'annotation.

Été 2013: Validation par une étude empirique.

Automne 2013 et Hiver 2014: Rédaction.

