



# Cours 3 Architecture Web

Pr. Okba KAZAR

Professeur des universités  
Directeur du Laboratoire d'INFormatique Intelligente  
LINFI

Département d'informatique  
Université de biskra

# Contenu

**1- Du WWW au GGG**

**2- Architecture Web**

**3- Applications e-business..**

**4- Indexation dans le Web**

**5- La personnalisation de la recherche**

# Références

## ➤ Web:

- Cours en ligne :  
<http://ocw.mit.edu/OcwWeb/Civil-and-Environmental-Engineering/1-264JDatabase--Internet--and-Systems-Integration-TechnologiesFall2002/LectureNotes/index.htm>
- Web Services working group :  
<http://www.w3.org/2002/ws>
- Microsoft Lab :  
<http://www.microsoft.com/france/msdn/lab/default.asp>

## ➤ Livres :

- .NET Framework Wrox
- Php/MySQL Dreamweaver Enrolles

# 1. Le World Wide Web

- Système hypermédia distribué fonctionne sur TCP/IP
  - Simple à utiliser : I.E, Netscape, Mozilla
  - Portable
  - Faible coût
  - Mobilité
    - Sans frontière
  - Standard
    - HTTP, URL, HTML, XML
  - Distribué : ActiveX, COM, CORBA, EJB, WebServices

# du WWW au GGG

- Evolution du World Wide Web vers un *Giant Global Graph*, selon [Tim Berners-Lee](#)
- Distinction de trois niveaux :
  - Les infrastructures : le *Net*
    - Relie les machines
  - La plate-forme de contenus : le *Web*
    - Relie les documents
  - Le graphe social : *social networking* et *web sémantique*
    - Relie les hommes et leurs ressources

# Net – Web - Graph

**GRAPH = Relations sociales**

**Étage de dissémination**

Étage de dissémination

**Seuil de rétroaction / perméabilité**

**WEB = plateforme de contenus**

**Étage de diffusion**

Étage de diffusion

**Seuil d'interaction / de production**

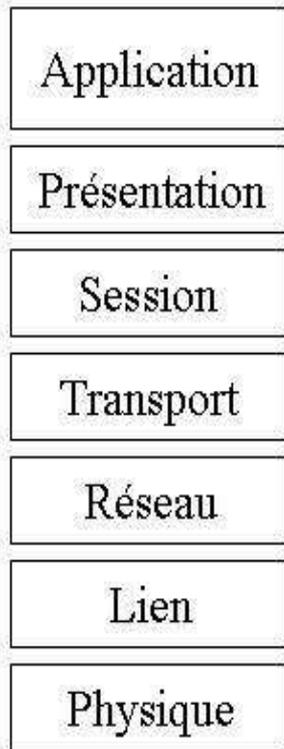
**NET = infrastructure**

**Étage de stockage**

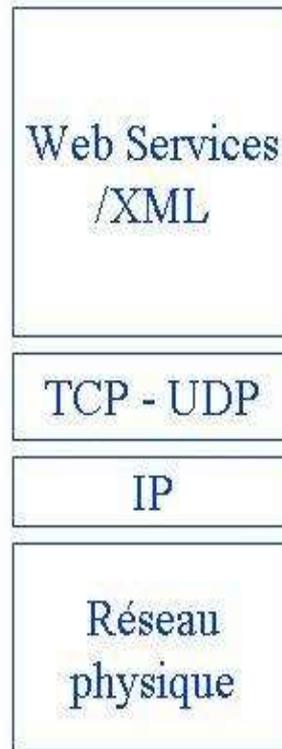
Étage de stockage

**Seuil « technologique »**

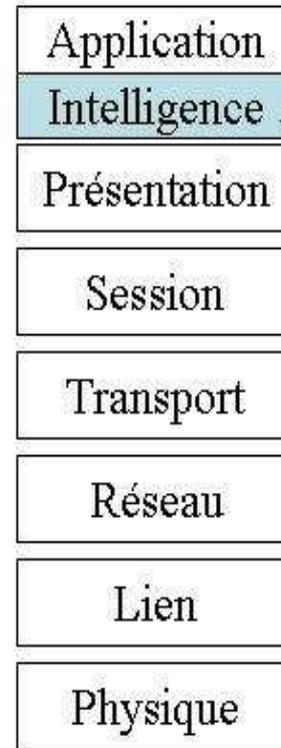
# Modèle du Web



Modèle OSI



Modèle Web  
TCP/IP



Modèle OSI+

*infomédiaire*

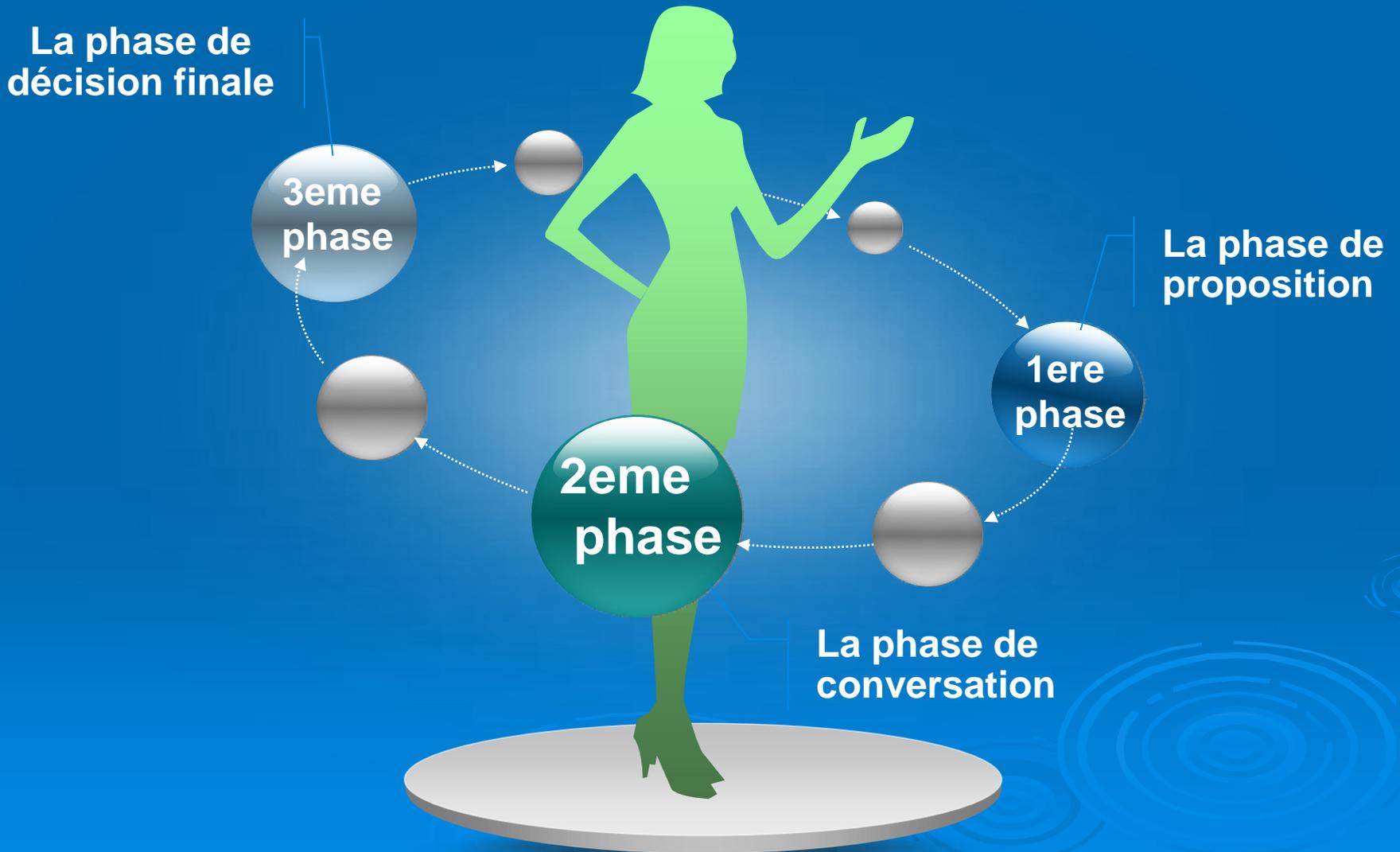
# Applications e-business

- **E-business = utilisation professionnelle du Web**
  - Au sein de l'entreprise : B2E
  - Avec les clients : B2C
  - Avec les partenaires : B2B
- **Applications majeures**
  - Intranets
  - Portails d'information d'entreprise
  - Commerce électronique
  - Publication en ligne

# E-commerce

- Elimination des procédures manuelles (papier) entre les partenaires d'un commerce
- Au minimum : B2C
  - Sites très connus, (Amazon.com, Dell Computers, etc.)
  - Gestion de la relation client
  - Intégration avec les bases de données
  - Transactions électroniques sécurisées pour commander
    - Cartes de crédit, monnaie électronique, E-check
  - Pas d'automatisation du côté client
- Pleine puissance : B2B
  - Intégration de la chaîne de distribution avec d'autres fournisseurs

# Les phases du protocole



# Etude de cas



NEGOCIATEUR



Accueil

Règles

Annonces

Location

Aide

## Présentation générale

*Le locataire définit sa recherche, son budget, ses souhaits et ses garanties  
Sa demande est mise en forme et envoyée au propriétaire concerné  
Le propriétaire prend contact directement avec les locataires qu'il sélectionne.  
Le propriétaire et le locataire louent rapidement sans aucun frais d'agence.  
Grâce à ce concept original, nous sommes les meilleurs*

*Soyez les biens venu Mr propriétaire , remplissez ce formulaire s'il vous plait pour vous enregistrer*

Nom  Prénom  Numéro telephone

E-mail

Mot de passe

Reentrez mot de passe

type de la maison à louer

Cochez et remplissez les attributs suivants:

Prix de la location

Paiement par mois ou semaine

Période de la location

bien meublé ou non

Avance de la location

paiement par

Attributs de l'offre:

*Sélectionnez les attributs et leurs marges et importances pour configurer la négociation s'il vous plaît.*

Attributs	Marge minimale	Marge maximale	Importance
<input checked="" type="radio"/> Prix location	85.000 DA	100.000 DA	70
<input checked="" type="radio"/> Période location	6 mois	18 mois	20
<input checked="" type="radio"/> Avance location	6 mois	12 mois	10
<input type="radio"/> Paiement par moi ou semaine			
<input type="radio"/> Bien meublé ou non			
<input type="radio"/> Paiement par			

Lancer négociation

Annuler

# Intranet

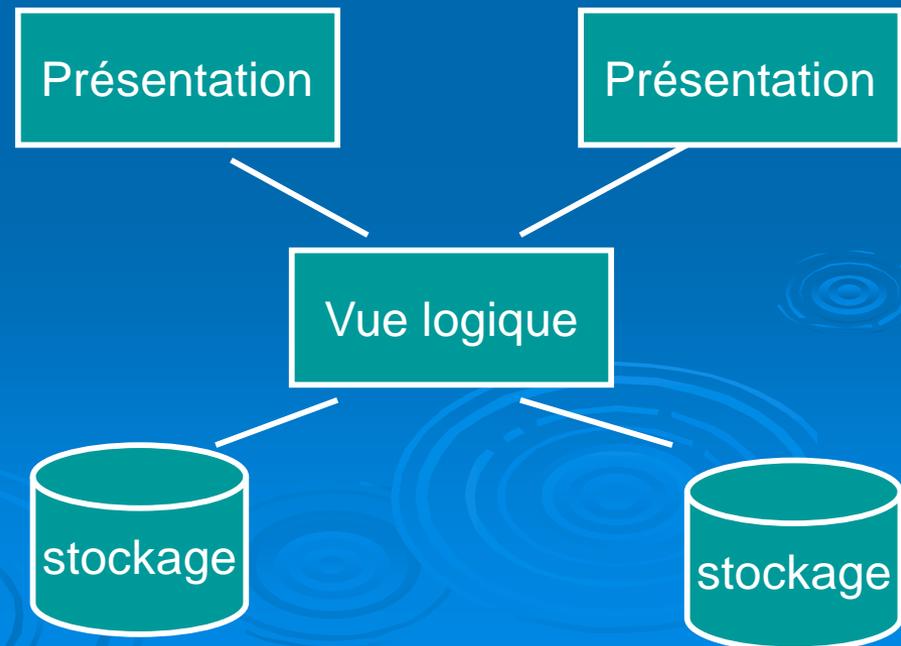
- Serveur Web pour les besoins internes de l'entreprise
  - Réseau privé avec accès sécurisé à Internet
    - firewalls
  - Utilisateurs connus
    - Services avancés : visio-conférence
  - Services internes
- Publication d'information et workflow
- Forum de discussion, email et groupware
- Réduit les coûts de communication et améliore l'efficacité
- Problème :
  - Intégration avec les sources de données de l'entreprise

# Portail d'information d'entreprise

- Porte d'accès à l'information de l'entreprise
  - Vue uniforme d'information agrégée à partir de sources de données hétérogènes
    - Applications existantes
    - Bases de données
    - Systèmes documentaires
  - Pour des utilisateurs spécifiques
    - Décideurs, employés, clients, fournisseurs
- Problème :
  - Difficile d'accéder à toute l'information de l'entreprise
  - Confiance

# Gestion d'information

- Gestion de l'information indépendamment de tout programme, y compris celui qui l'a créé.
- Principe = indépendance des données
  - du stockage
  - de la présentation



# Fonctionnement du Web

Pour consulter un document (ex: <http://serveur/page.html>)

1. Le navigateur demande à son DNS (Domain Name Server) de lui renvoyer l'adresse IP du serveur demandé dans l'URL (Uniform Resource Locator)
2. Le navigateur envoie une requête à l'adresse IP pour demander la page.html précisée dans l'URL.
3. le serveur envoie la page au client.
4. le navigateur interprète la page et l'affiche.

HTTP : protocole de transfert permettant de demander un fichier au serveur ainsi qu'une description de types des données associées.

IP : Internet Protocol

# 2. Couplage BD et Web

## Pourquoi coupler ?

- Deux domaines qui se sont développés indépendamment mais qui ont de nombreux points communs :
  - Ensemble important de données (modélisation, stockage, indexation)
  - Interrogation (langage déclaratif)
  - Multi-utilisateurs
  - Performances
  - Fiabilité

# Données du Web

- Données hétérogènes, fortement corrélées, évoluant souvent
- La gestion de ces données avec un SGF et un système d'exploitation pose des problèmes :
  - gestion des liens difficile
  - gestion de l'évolution (pas d'indépendance logique-physique)
  - performances moyennes (pas ou peu d'index)

# Apports des SGBD

- Gestion de gros volumes d'information
- Bonnes performances d'accès (index, stockage, gestion du disque)
- Fiabilité des données (cohérence, sûreté de fonctionnement, sûreté d'accès)
- Partage et accès concurrents (transactions)
- Productivité (indépendance physique-logique, réutilisation, outils)
- Evolutivité (évolution du schéma)
- Langage de requêtes et optimisation

# Histoire de la technologie de programmation

▲ Low-level programming ( machine code, Assembler )

▲ High-level programming ( Fortran, Cobol, Basic, ...)

▲ Structured programming ( Pascal, C, C++, ...)

▲ Object-Oriented Programming ( C++, Java, Delphi ...)

▲ Component Oriented Programming  
( COM, JavaBeans, ...)

▲ Component Oriented Web Programming  
( C#, Forte Java, ...)

▲ Service-Assemblage/Creation  
Programming  
( Web Services...)

1950 1960 1970 1980 1990 2000 2010 2020 2030 2040 2050 2060

# Primary programming - modèles OS

▲ Low-level programming ( IBM Mainframe OS )

▲ High-level programming ( IBM Mainframe OS... )

▲ Structured programming ( Unix ... )

▲ Object-Oriented Programming ( Unix, MS DOS, ... )

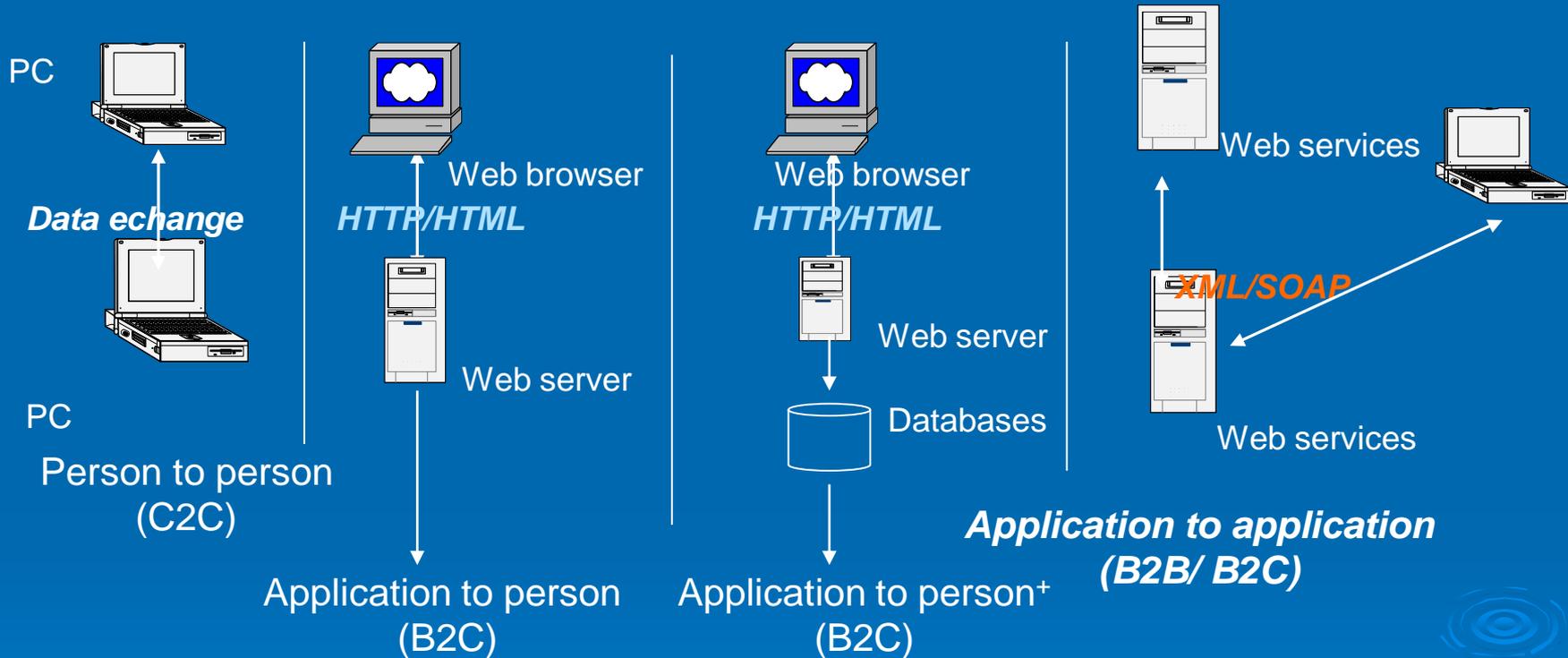
▲ Component Oriented Programming  
( Windows, Unix, Linux, ... )

▲ Component Oriented Web Programming  
( DotNET, JEEE, ... )

▲ Service-Assemblage/Creation  
Programming  
(...)

1950 1960 1970 1980 1990 2000 2010 2020 2030 2040 2050 2060

# Intelligence dans l'architecture x-tiers



Architecture 1-tier → Architecture 2-tiers → Architecture 3-tiers → Architecture "n-tiers"

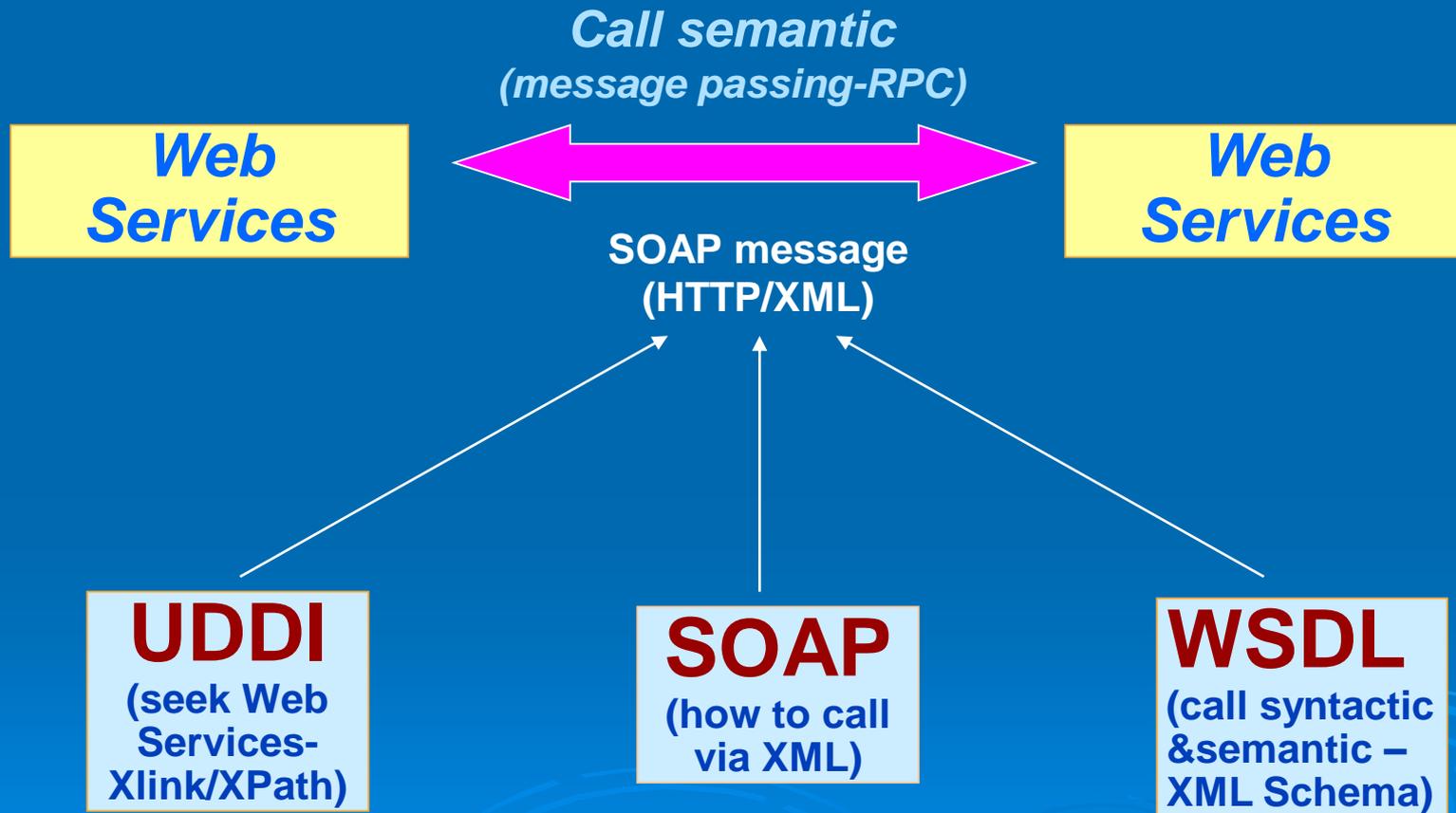
Evolution de l'architecture

# XML pour l'échange de données universel

- Le langage XML (Extensible Markup Language)
  - Standard du W3C
  - Décrit le contenu, pas la présentation
    - Structure, type, schéma, requêtes, etc.
  - Une base forte : XML est un sous-ensemble de SGML
  - Fournit l'indépendance des données au stockage et à la présentation .
- Facilite l'échange de données entre applications

# Web Services

Un Web Services est une "unité logique applicative" accessible en utilisant les protocoles standard d'Internet, indépendamment de plate-forme et peut être réutilisé.

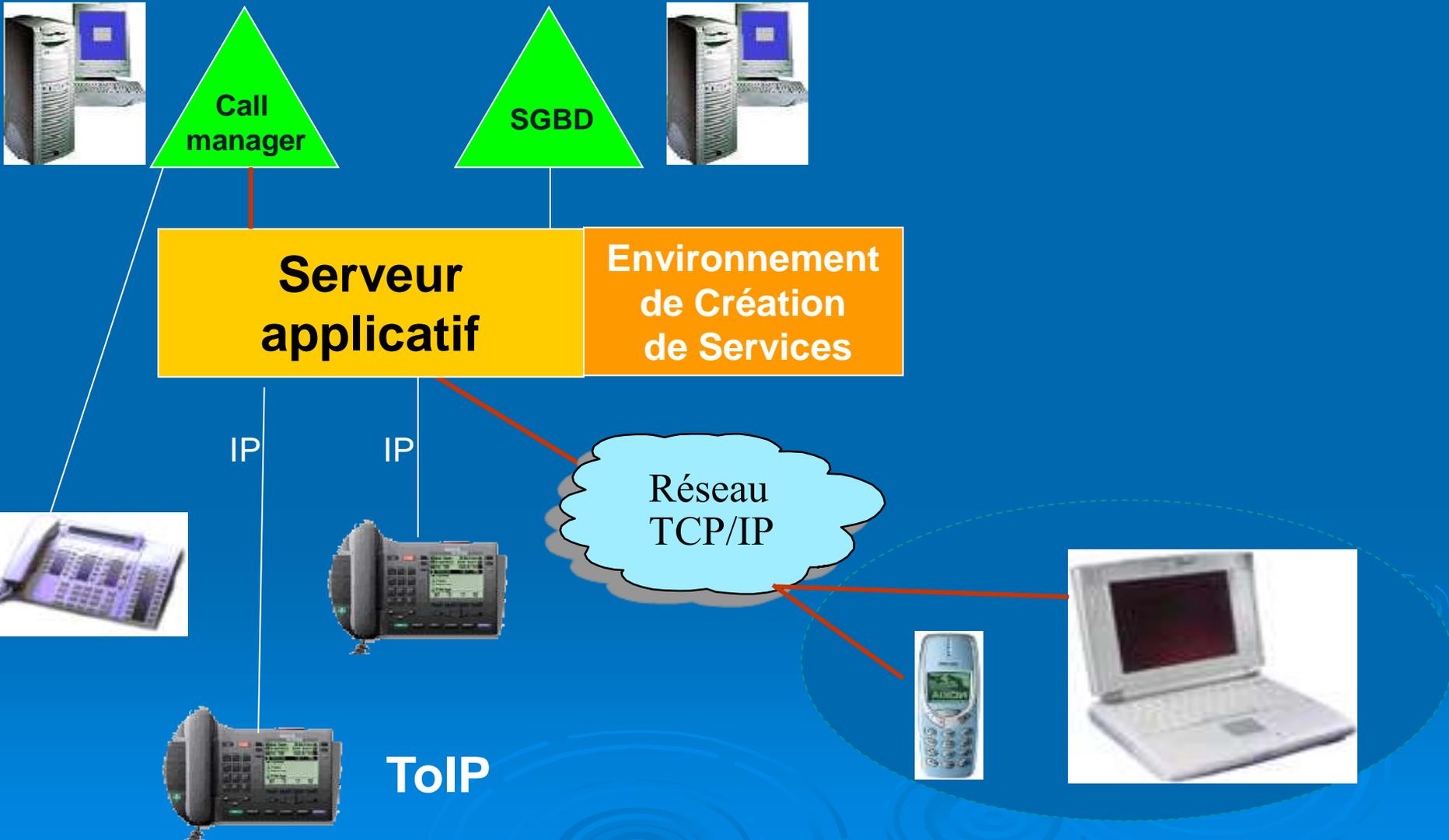


# 4. Web dans l'entreprise

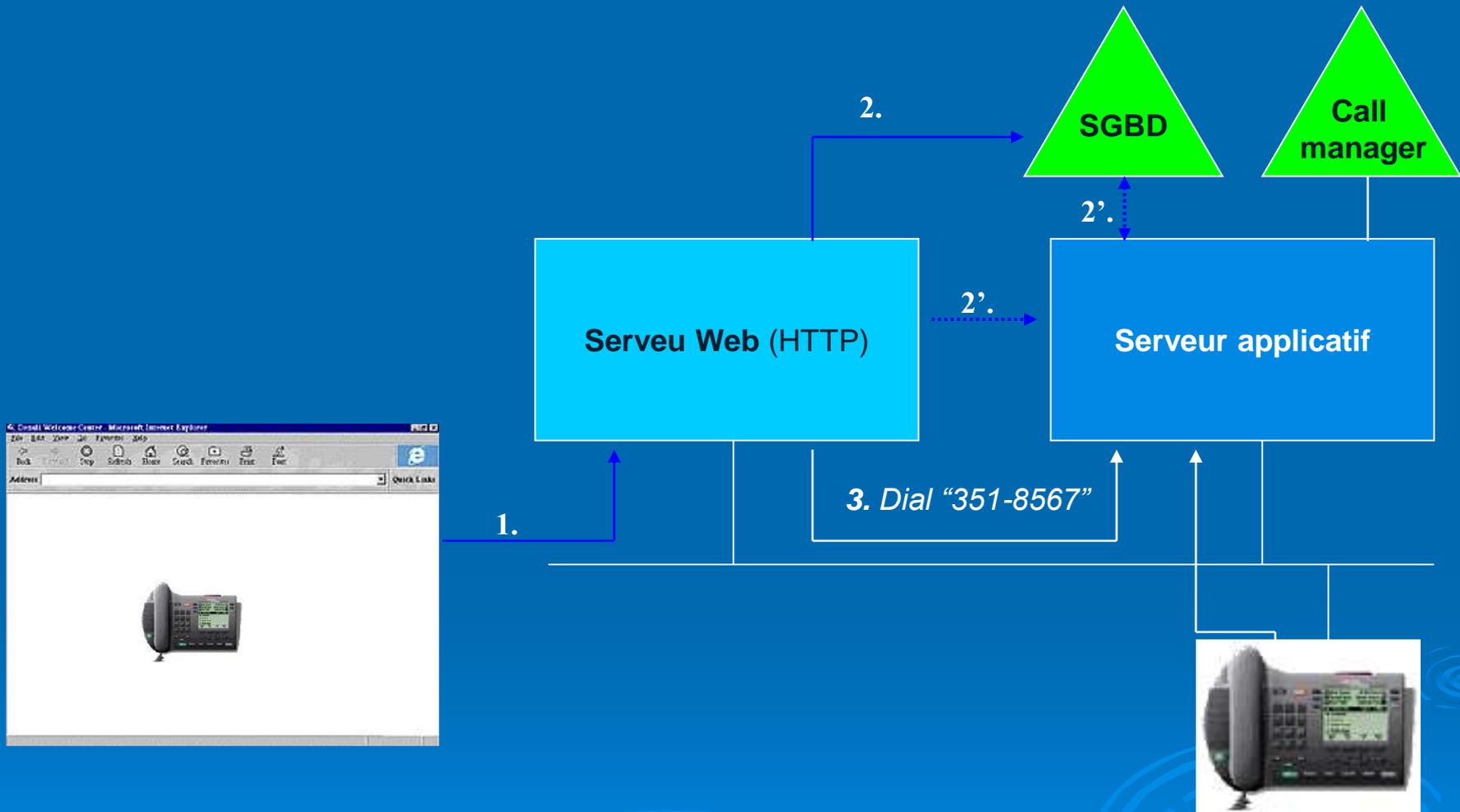
## Serveur applicatif dans l'entreprise

- Simple = serveur Web
  - Petites applications, pas de transactions
- Serveur d'entreprise = serveur Web +
  - Support des grandes applications, distribuées standard
  - Support des transactions
  - Support des composants
  - Support de XML
  - Outils de développement

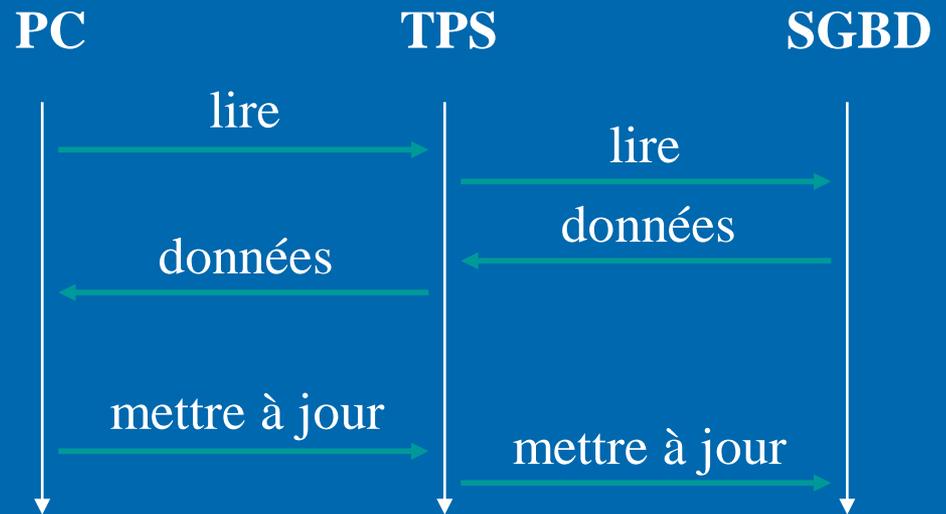
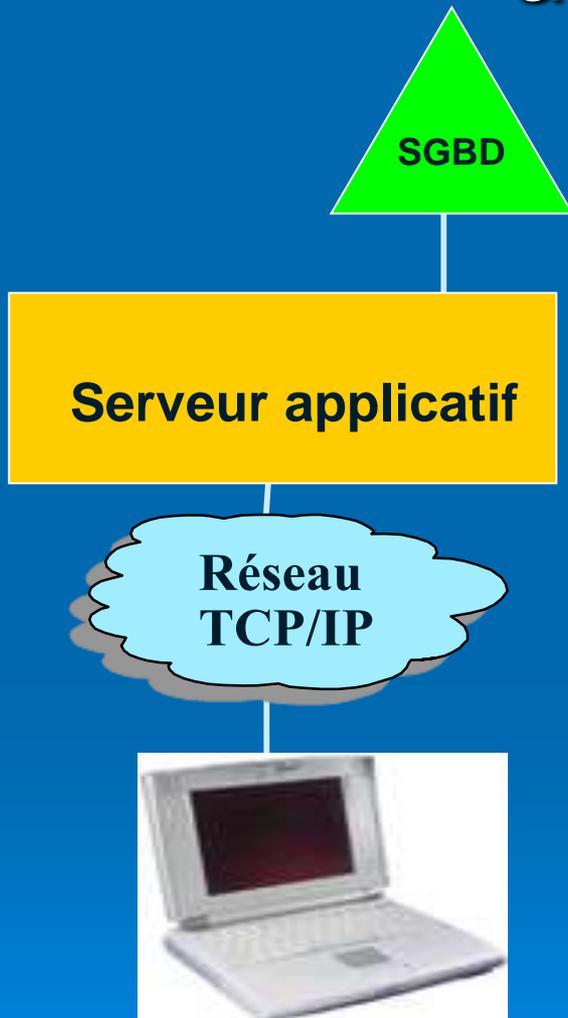
# Exemple : Web dans l'entreprise



# Appel à partir du Web



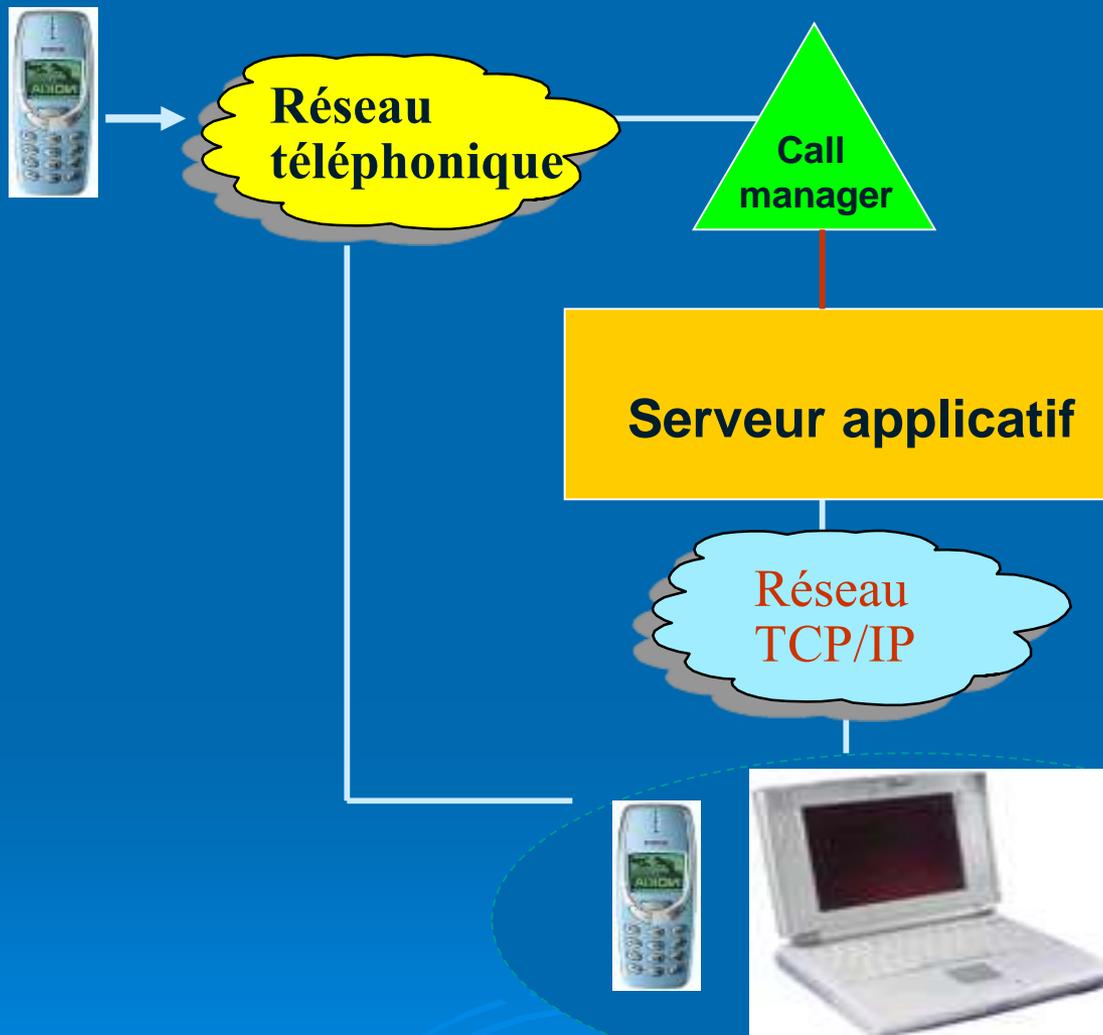
# Accès aux données à distance via Web



## Données :

- Répertoire personnel : lire et MAJ
- Profil de renvoi : lire et MAJ
- Configuration de touches (une par une):  
lire et MAJ
- Journal : lire

# Appel Sortant / Appel Entrant



# 5/ Evolutions dans l'indexation

- **Rappel des finalités de l'indexation :**
  - Recherche de documents
  - Accès au contenu des ressources
  - Analyse, filtrage automatique, cartographie de l'information

# Evolution dans l'indexation

- **Imbrication des différentes modalités d'indexation sur le web :**
  - **Indexation documentaire** : thésaurus, description des ressources (annuaires)
  - **Indexation automatisée** : le TAL (Traitement Automatique des Langues)
  - **Indexation sociale** : tags du web 2.0, recommandations
  - **Indexation sémantique** : les métadonnées et les ontologies

# 5.1 Evolutions dans l'indexation : L'indexation automatisée « full-text »

- **Deux grandes méthodes d'analyse dans l'indexation automatisée :**
  - **analyse statistique** : fondée sur la **fréquence** des mots
  - **analyse linguistique** : fondée sur la **reconnaissance** des mots
- **4 niveaux d'analyse du texte intégral :**
  - **niveau morphologique** : reconnaissance du mot
  - **niveau lexical** : réduction du mot à sa forme canonique > **lemmatisation**
  - **niveau syntaxique** : utilisation de la grammaire
  - **niveau sémantique** : reconnaissance des concepts

## 5.2 L'indexation automatisée : les moteurs de recherche sémantique

- Emergence de nouvelles catégories de moteurs :
  - Hakia :
    - Lancé en 2006
    - moteur de requêtes en langage naturel
    - mélange d'analyse sémantique, d'ontologie, de logique floue et d'intelligence artificielle
  - Powerset :
    - Mai 2008 ; racheté par Microsoft en juillet 08
    - Recherche sémantique sur Wikipedia
    - Analyse des phrases contenant les mots d'une requête
    - Propositions de nombreux mots-clés



# 5.4 Evolutions dans l'indexation :

## L'indexation sociale : tags et folksonomies

➤ **Le principe de la folksonomie** (indexation personnelle) , est un:

- système de classification de « classification collaborative décentralisée spontanée », s'appuyant sur les termes choisis par les utilisateurs
- Objectif : faciliter l'indexation des contenus et la recherche d'information
- Tags peuvent s'appliquer à des signets web, à des photos, à des projets, des vidéos, ou encore des billets de blogs (nuages de tags)
- Constitution d'une communauté de « spécialistes » parmi les internautes

## 5.5 Indexation : quelles conséquences pour la formation des usagers ?

- Sensibiliser les étudiants à l'importance de l'indexation :
  - Dans la recherche sur les moteurs
  - Dans les plates-formes web 2.0
- Montrer :
  - l'ancienneté de l'indexation
  - la diversité des modes d'indexation et d'accès aux contenus
  - les enjeux pour la recherche d'information
- Expliquer les notions essentielles :
  - indexation, classification, traitement linguistique,, métadonnée, indexation automatisée, full-text, indexation sociale, tag, folksonomie, ontologie...

# 6. La personnalisation de la recherche

- **Evolution des moteurs vers des services personnalisés pour fidéliser les internautes**
  - Une tendance lourde de la RI, va en s'amplifiant
  - Remonte au moins à 2000
  - Innombrables aspects :
    - De l'historique des recherches au moteur personnalisable
  - Innovations permanentes
- **Possibilité de personnalisation par :**
  - le moteur de recherche
  - un Plug in, ex : SurfCanyon, plug in firefox et IE
  - par l'internaute

# La personnalisation de la recherche

- **9 Modèles de personnalisation par le moteur (Jeff Quipp)**
  - **Personnalisation basée**
    - sur la localisation
    - Sur l'interface
    - Sur l'historique des requêtes
    - Sur le mode entrée / sortie
    - Sur le temps
    - Sur le comportement individuel de l'utilisateur
    - Sur le comportement de groupe
    - Sur la recherche sociale / graphe social de l'utilisateur
    - Personnages virtuels / sur les Assistants

# La personnalisation de la recherche

- **Les différentes facettes de la personnalisation par l'internaute :**
  - Personnalisation des requêtes
  - Personnalisation des résultats
  - Personnalisation des interfaces
  - Personnalisation des outils
  - La création de communautés

# La personnalisation de la requête

## ➤ Les bases de la personnalisation de la requête :

- L'historique des requêtes
- La définition d'un profil et d'une alerte
- La géolocalisation
- Le choix des sources d'information

## FIN du cours3

