

L'internet social (ou Web 2.0): opportunités, impacts et défis

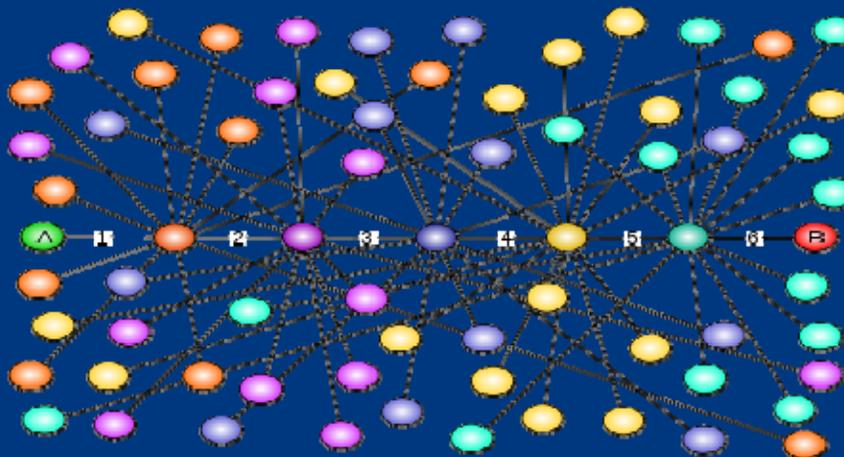
Patrick Valduriez

Séminaire HiPhis, 9 février 2010

INSTITUT NATIONAL
DE RECHERCHE
EN INFORMATIQUE
ET EN AUTOMATIQUE



centre de recherche
SOPHIA ANTIPOLIS - MÉDITERRANÉE



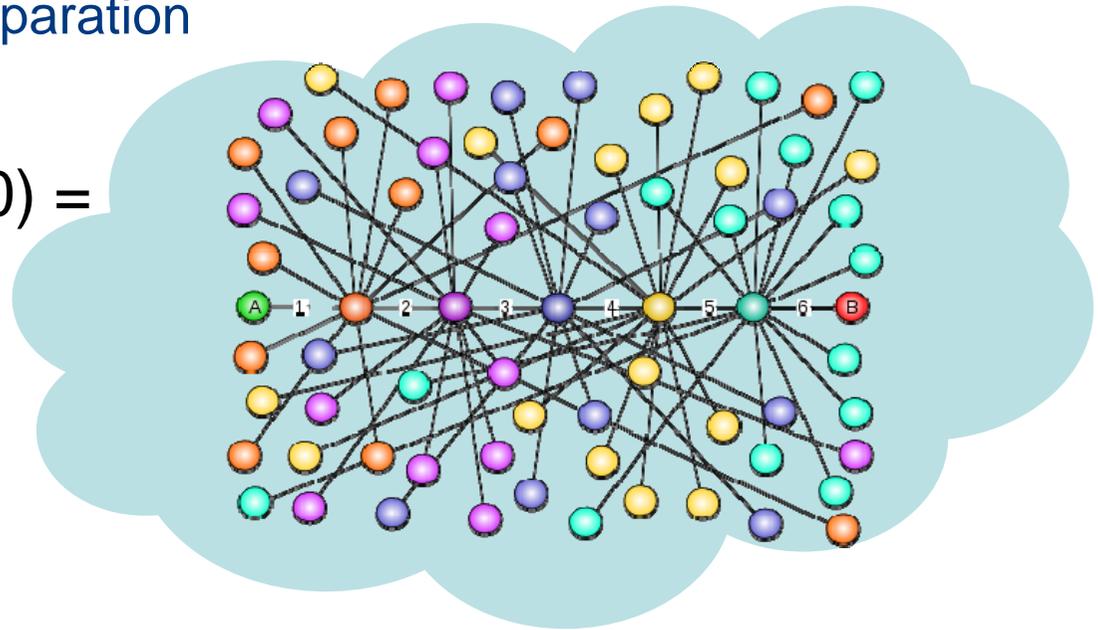
Web et réseau social

Le Web: réseau d'information sur Internet

Réseau social: ensemble d'entités sociales reliées par des relations et interactions sociales

- Le concept du petit monde (Stanley Milgram, Harvard 1967) ou les six degrés de séparation

Internet social (Web 2.0) =
Web + Réseau social



L'internet social en action

Obama 2008: omniprésence dans les réseaux sociaux durant la campagne présidentielle

- Présent dans 15 réseaux; 5 M supporteurs, en plus de son propre réseau
- Ex. sur Facebook, plus de 5,4 millions d'utilisateurs ont cliqué "J'ai voté", pour informer et inciter leurs amis

Iran 2009: transmission en direct sur Twitter des infos sur le mouvement de protestation

- Ex. sur Twitter, on pouvait envoyer des adresses de proxies pour permettre aux opposants de communiquer sans être bloqués par les autorités

Haiti 2010: mobilisation des réseaux sociaux pour informer en temps réel et collecter les dons

- Plus de 500 groupes créés, ex. Earthquake Haiti sur Facebook; 1 page dédiée sur Wikipedia



Plan de la présentation

Evolution des technologies de l'information et de la communication

Le Web 2.0

Le Web sémantique

Les réseaux sociaux en ligne

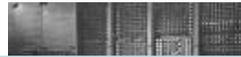
Mon analyse de l'internet social

Pistes de recherche

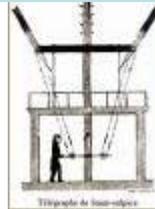


Evolution des TIC

Supports d'information



Communication instantanée
Toute information (numérique)
Pour tous!



Supports de communication



Technologies actuelles

Info

Opportunités d'innovation

Connaissance en TIC

Donnée

Description élémentaire d'un fait

- Ex. Observation, mesure, état d'un objet, lettre, photo, note d'examen, transaction, etc.

Informations

Obtenu par interprétation et traitement

- Ex. Prévision météo, classement d'un objet, etc.

Connaissance

Obtenu par structuration et interprétation

- Ex. Règle métier, règle d'inférence, etc.



Plan de la présentation

Evolution des technologies de l'information et de la communication

Le Web 2.0

Le Web sémantique

Les réseaux sociaux en ligne

Mon analyse de l'internet social

Pistes de recherche



Web 2.0: Définition

Terme inventé par O'Reilly Media

Pas de définition acceptée : concept technologique pour les uns, évolution fonctionnelle du Web pour les autres, vaporware marketing pour d'autres

Parfois opposé au concept du Web sémantique

Synthèse des points de vue

1. Réseau d'interaction sociale à l'échelle mondiale
 - Importance de l'utilisateur et de ses relations avec les autres
2. Basculement des techniques du Web vers des services
 - Cloud computing
3. Plate-forme d'innovation qui fait du Web un nouveau système d'exploitation
 - Ex. Google apps



Web 2.0 : Exemples de techniques

Ajax: communication asynchrone entre client Web et serveur via Jscript et XML

- Très grande interactivité (ex. Googlemaps)

REST: architecture de services légère avec appels HTTP

XForms: Saisie déclarative de formulaires contrôlée en XML

XUL/XAML: interfaces graphiques déclarées en XML

RDF: langage d'annotation de ressources en XML

- Pour donner du sens aux ressources (Web (sémantique))

RSS et ATOM: syndication de flux d'information XML



Plan de la présentation

Evolution des technologies de l'information et de la communication

Le Web 2.0

Le Web sémantique

Les réseaux sociaux en ligne

Mon analyse de l'internet social

Pistes de recherche

Questions ouvertes



Le Web Sémantique (ou Web 3.0 ?)

The Semantic Web, Tim Berners-Lee, James Hendler and Ora Lassila, Scientific American, May 2001

- Projet du W3C pour rendre l'information plus accessible et la traiter automatiquement

Vision d'un vaste espace de ressources (pages web, documents, etc.) avec une sémantique explicite

- Permet de comprendre et interpréter les ressources

Description des ressources avec des métadonnées dans un formalisme à base de connaissances (RDF)

Utilisation d'ontologies (RDF Schema, OWL) pour annoter les documents avec des méta-données



Ontologie

Ensemble structuré des termes et concepts représentant un domaine d'information

- Définition des concepts, de leurs propriétés et leurs relations

Exemples:

- Finance, Tourisme, Transport, Médecine, Immobilier ...

Utilisée par les gens, les bases de données et les applications pour classer et mieux partager l'information et son sens, et en déduire des connaissances



Limites des ontologies

Exige une très bonne classification à priori, donc une très bonne connaissance du domaine

- Bons exemples: tableau périodique des éléments chimiques, classification scientifique des espèces, référentiel d'un système d'information

Difficultés

- La connaissance évolue, très vite dans le domaine du Web, donc il faut faire évoluer la classification
 - Ex. la hiérarchie de catégories complexe de Yahoo (avec des catégories additionnelles) versus la classification automatique de Google par mots-clés
- Le producteur d'une information doit pouvoir la ranger dans une catégorie (ou plusieurs le cas échéant)



Wiki: l'outil de collaboration du Web

Permet à n'importe qui [autorisé] de créer des contenus et de les éditer sans restriction

- Simplification et passage à grande échelle du processus de création de pages Web

Wikipedia: l'encyclopédie universelle libre

- Source de référence: un des 10 sites les plus visités au monde, 267 langues, 13 millions d'articles en 2009 (au total)
- Libre: gérée par une fondation
- Auto-organisée, plutôt correcte
- Plusieurs niveaux de surveillance des informations
- Critiques concernant l'anonymat, la communauté des contributeurs et l'absence de comité d'édition



Plan de la présentation

Evolution des technologies de l'information et de la communication

Le Web 2.0

Le Web sémantique

Les réseaux sociaux en ligne

Mon analyse de l'internet social

Pistes de recherche



Réseaux sociaux en ligne (Web-Based Social Networks)

Partage d'informations dans un réseau d'utilisateurs ayant un intérêt en commun

- Gestion des profils utilisateurs
- Maintien des liens entre utilisateurs
- Annotation de contenus et de signets (*bookmarks*)
- Recommandations, avis
- Recherche intégrée par mots-clés
- Nombreux autres services



Quelques catégories

Général (social, divertissement, humanitaire, ...)



Blogs et microblogs



Relations professionnelles



Partage de photos, d'avis sur les films ...



Réseaux spécialisés



Pour créer son propre service de blogs



Folksonomie

De l'anglais *folksonomy* (contraction de folk et taxonomy)

Méthode de classification simple et collaborative de l'information par des *tags*

- L'utilisateur décrit lui-même l'information à partager (des photos, des signets, ...) par des *tags* (mots-clés), sans recourir à une terminologie prédéfinie
- La classification d'une information dans un réseau social devient l'union des tags associés et sert à l'indexation
- Améliore la recherche d'information
- Utilisé par Wikipedia, Flickr, delicious, Twitter, etc.

Limites ou forces?

- Manque de précision, pas de synonymie, explosion du nombre de tags



Web sémantique vs folksonomies

Web sémantique: top-down et organisé

- Besoin d'ontologies bien définies, grande précision sémantique
- Nécessite un grand effort de catégorisation de l'information

Folksonomies: bottom-up et désordonné

- Mettre des tags est facile, naturel et populaire
- Pas de précision sémantique, pas de hiérarchie des termes

En attendant le Web 3.0: les micro-formats

- Formats de données XHTML ou XML avec balises standardisées pour les métadonnées, ex. nom, adresse, tags
 - Ex. en xFN (xhtml Friends Network)
Jean



Plan de la présentation

Evolution des technologies de l'information et de la communication

Le Web 2.0

Le Web sémantique

Les réseaux sociaux en ligne

Mon analyse de l'internet social

Pistes de recherche



Utilisation

Réseaux sociaux

- Cohésion de groupe grâce à la poursuite d'un but commun fédérateur
- Auto-organisation, très efficace
- Maintien des relations de confiance entre personnes (égal-à-égal)
 - Puissance de la recommandation pour la recherche individuelle (aide, conseil, rencontre, emploi, contact, etc.)
- Nouvelles possibilités de partage d'opinion, d'apprentissage, etc.

Entreprises

- Outil de communication d'image, de fidélisation des clients
- Mine d'or pour le marketing
 - Profils de consommateurs fiables à partir de leur nuage de tags



Impact sur l'information et la communication

Facilité d'utilisation

- Interfaces multimédias intuitives
- Mime les comportements naturels

Qualité de l'information

- Plus l'information est partagée, plus elle devient de qualité!
- Contributions, avis et recommandations par des personnes motivées, voire expertes
- Confiance dans les recommandations (des amis, des collègues)

Communication efficace

- Accélération de la transmission d'information
 - Réactivité des membres du réseau
- Grande couverture des différents réseaux: l'effet petit monde
 - Une personne est membre de plusieurs groupes ou réseaux



Opportunités au delà du social et du marketing

Communautés professionnelles en ligne

- Partage d'information en ligne
- Relations entre spécialistes
- Production collective à grande échelle
 - Logiciel libre: Linux, Eclipse
- Collaboration en temps réel

Ozmosis: partage des connaissances médicales

- Clinique, pratique des soins, politique de santé
- Ex. à terme: les études épidémiologiques
 - Faisant intervenir patients, médecins, chercheurs, etc.



Atteinte à la vie privée



Problèmes éthiques et légaux

- Selon la CNIL, en France, l'utilisateur doit pouvoir contrôler l'utilisation de ses données personnelles
- Mais le modèle économique des opérateurs de services de réseaux sociaux est précisément l'exploitation de ces données
 - Marketing, publicité, etc.
- Conditions générales d'utilisation de ces données
 - Perpétuelle, transférable, pour le monde entier
 - Ex. Controverse suite à la modif. de Facebook le 15/2/2009
- Pas de cadre juridique international
 - La langue juridique reste l'anglais



Problèmes de sécurité et confidentialité

Usurpation d'identité

- Facile sur un site de réseau social
- Problème pour les rencontres (physiques)
- Pas un délit sur Internet
- Solution: les services d'identification numérique en ligne

Diffusion instantanée d'information

- Pas de 
- Pas de contrôle de la propagation
- Exposition maximale



Vers un standard d'échange de données personnelles?

Pour pouvoir faire interopérer différents réseaux sociaux

- Ex. importer une liste de contacts Myspace depuis Facebook

L'enjeu est le contrôle des échanges de données personnelles entre internautes

Les forces en présence

- Facebook Markup Language (FBML), le leader avec ses 400M membres
- OpenSocial de Google, interopérable avec les réseaux sociaux qui l'ont adopté: Orkut, Hi5, Friendster, Viadeo, MySpace, Netlog, LinkedIn, ...
- The Friend of a Friend (FOAF) project, ouvert, mais assez peu utilisé



Plan de la présentation

Evolution des technologies de l'information et de la communication

Le Web 2.0

Le Web sémantique

Les réseaux sociaux en ligne

Mon analyse de l'internet social

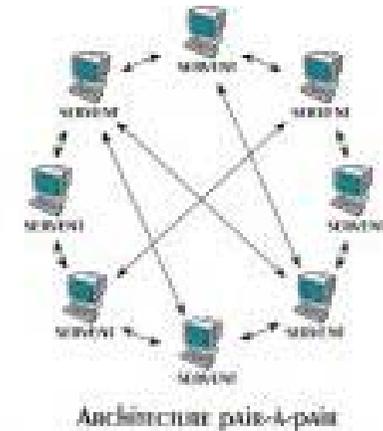
Pistes de recherche



Internet social et P2P

Internet social

- Architecture client-serveur

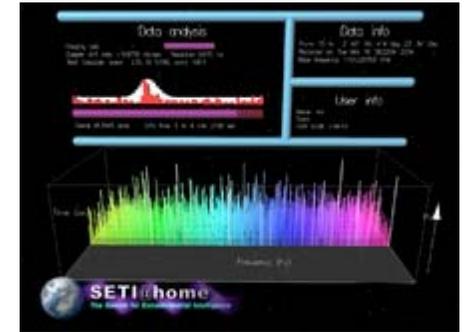


Le P2P

- Célèbre pour le téléchargement illégal de fichiers (Bittorrent, FastTrack, etc.)
- Architecture décentralisée, auto-organisée
 - Pas de contrôle centralisé des données
 - Distribution et parallélisme à très grande échelle
 - Haute disponibilité grâce à la réplication des données
- Protocoles de communication robustes
 - Ex. le *gossiping*: exploite le concept petit monde



Le projet SETI@home



Search for Extra-Terrestrial Intelligence

- Projet du Space Sciences Laboratory de UC Berkeley
- Recherche de transmission radio extraterrestre à partir des observations faites par le radiotélescope d'Arecibo, Puerto Rico

Principe

- Les données d'observation sont numérisées et divisées en petites unités (de fréquence et temps)
- L'analyse de ces unités est faite par des participants volontaires, à l'aide d'un logiciel gratuit téléchargé sur leur PC

Résultats

- 6 millions de participants depuis le lancement en 1999
- L'équivalent de 776.491 TeraFLOPS le 8/2/2010 (~300k PC)
- Mais aucun signal extraterrestre détecté à ce jour



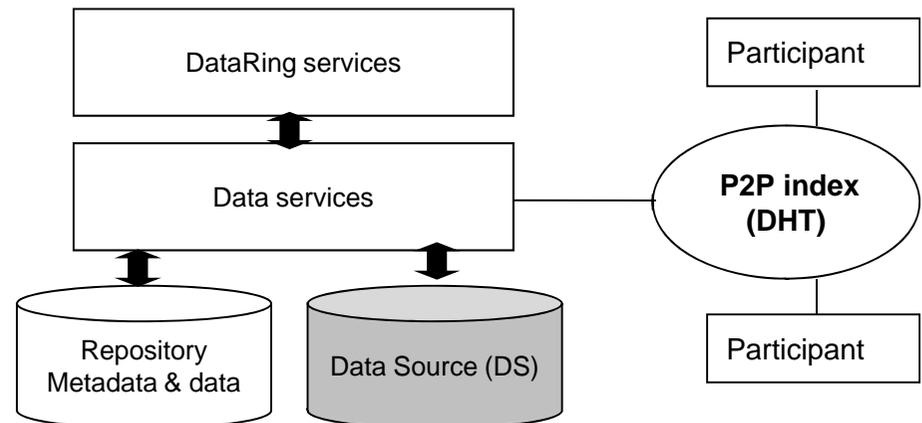
P2P pour communautés en lignes

Le projet ANR VERSO DataRing (2009-2011)

- **INRIA Sophia**, INRIA Saclay, LIG, LIRMM, Telecom Paristech

Approche

- Chaque participant garde le contrôle de ses données
- Services avancés
 - Intégration sémantique
 - Données incertaines
 - Réplication
 - Confidentialité, confiance



Conclusion

Valeurs universelles de l'internet social

- Partage, collaboration, implication, confiance

Cercle vertueux entre progrès technique et usage innovant

- Toujours plus de pouvoir au citoyen, consommateur, etc.

Impact majeur sur toutes les dimensions de la société

Possibilités immenses d'inventer de nouveaux usages

- Ex. pourrions nous gouverner ensemble?

