

Kim H. Veltman

Le nommage et l'indexage comme vision panoptique du reel

Unpublished: Originally planned for: *Les Cahiers Du Numérique*, Paris, No. 8, Mars 2002, pp. 1-10.

-
1. Introduction
 2. Développements de L'Internet
 3. Origines du Domain Name Spaces (DNS)
 4. Internet et Les Institutions de Mémoire Collective
 5. Défis pour l'Organisation des Connaissances et pour l'Accès
 - Global, National, Régional et Local
 - Dimension Temporelle
 - Cartes Dynamiques
 - Les Connaissances Dynamiques
 6. Culture par Rapport au Réseau Sémantique et L'Intelligence Machinelle
 7. Conclusions
-

1. Introduction

Dans les beaux temps d'antan la question des noms de domaines semblait un problème simple. Il y'avait une petite communauté et le remarquable Dr Jon Postel avait tout dans ses mains. Dans les contes de fées on vit toujours. Malheureusement Jon Postel est mort (le 16 novembre 1999) peu après la constitution de ICANN (Internet Corporation on Assigned Names and Numbers)¹ et les problèmes ont fleuri.

Cette article passe en revue l'histoire plutôt pragmatique des espaces des noms de domaines (anglais: Domain Name Spaces ou DNS). Aujourd'hui les DNS procèdent du spécifique au générale, sont présentés du droit à gauche et manquent une valeur sémantique. On propose une nouvelle approche qui renverse la direction (comme dans le plan de nombrage E.164), et se divise en trois parties: une partie qui détermine le type de server (i.e. www, ftp); une partie géographique, suivie par une partie conceptuelle, laquelle est lié avec les systèmes de classifications dans le monde des bibliothèques, musées et archives. Très simplement, on propose que les expériences en tant que l'organisation du savoir et des connaissances de nos insitutions de mémoire collective peuvent nous aider pour répondre aux défis de l'Internet. Une deuxième section de l'article traite le besoin de nouvelles méthodes de connaissances dynamiques pour répondre aux défis de la complexité culturelle. Pour y arriver on propose un projet à longue terme, MEMECs, et un nouveau réseau à haut débit pour lier les universités et instituts de recherche au niveau européen et ensuite globalement.

Au début l'Internet était pragmatique sans méthodes bien définies pour le nommage et l'indexage. Lorsqu'il s'agissait d'une petite groupe de chercheurs ça ne posait guère de problèmes. Mais jusq'aout 2001, cette groupe de l'Internet avait plus de 513 million de

membres. Si on continue sans système, nous serons lancé dans un chaos prodigieux, lequel va détruire les rêves d'un accès universel qui a inspiré l'Internet au début. On a besoin d'une structure de nommage logique, systématique et sémantique.

Tout d'abord on propose une révision des noms des domaines et leurs abréviations pour leur rendre plus cohérentes. En anglais les anciens noms sont Top Level Domains (TLDs), Second Level Domains (SLDs) et Third Level Domains (3LDs). Il serait plus clair de les nommer First Level Domains (1LDs); Second Level Domains (2LDs) et Third Level Domains (3LDs). On propose aussi de leur traiter comme des classes, sous classes et sous-sous classes respectivement écrites de gauche à droit.

Cette révision a des conséquences pour trois secteurs:

- 1) les systèmes qui dépendent de la lecture machinelle des données, comme les moteurs de recherche (search engines), des browsers et des agents intelligents. Il sera plus simple à déterminer les adresses, ce qui leur rendra plus efficaces;
- 2) les professionnels peuvent mieux s'orienter et trouver ce qu'ils cherchent;
- 3) l'utilisateur/utilisatrice quotidien(ne) peut mieux s'orienter et mieux déterminer l'authenticité des données ce qui rendra plus fiable l'Internet, chose essentielle si on vise au commerce électronique.

Presque tout le monde est d'accord que nous avons besoin d'un réseau sémantique. Mais il y a deux interprétations fort différentes de sémantique, ce qui mène a deux voies opposées pour achever cette vision. D'un part, il y en a ceux qui veulent un réseau uniquement lisible aux machines; un réseau débarrassé des problèmes de langage, culture et d'histoire, accessible aux agents robotiques "intelligents," lesquels un jour pourraient remplacer totalement l'intervention humaine. Cette vision est particulièrement liée avec les aspirations du militaire.

D'autre part il y a une vision, que nous partageons, d'un réseau sémantique dans le sens grammatical avec signifiante, symbolisme, interprétation, nuance, multi-lingue, multi-culturel, avec une dimension historique; un réseau qui reflète les besoins de la condition humaine plutôt que les exigences machinelles. Pour y arriver nous devons ajouter les connaissances culturelles et historiques à nos banques de données. Nous devons passer des modèles statiques aux modèles dynamiques du savoir. Donc on propose un projet à long terme, Metadonnées et Mémoire Collective Systématique (MEMECS), ce qui pourrait être achevé à travers d'un réseau de centres d'excellence selon la vision d'un Espace Européen de Recherche (European Research Area, ERA).

2. Développements de L'Internet

L'Internet a commencé par lier une petite groupe de chercheurs en physique et le militaire. Avec le début du World Wide Web en 1990, l'Internet a commencé à croître énormément. De mars à septembre en 1998 il a plus que doublé de 66.6 à 147 millions. Dans les derniers trois années il a épanoui jusqu'à 513 millions² (figure 1).

Septembre	1998	147 million
	1999	201
	2000	368
	2001	513

Figure 1. L'expansion de l'Internet de 1998-2001.

Le rôle de la langue anglaise a changé aussi. Au début l'Internet était presque exclusivement anglais. En 1996 il y avait vers 40 million utilisateurs en anglais (80%) et 10 millions en autres langues (20%). En septembre 2001 il y avait 220 million en anglais (43%, un pourcentage qui va se réduire à 25 % en 2005 selon les prédictions, quand il y'aura 320 million en anglais d'un total de 1120 millions).³ Le chinois qui représente pour le moment 9.2 %, donc la deuxième langue de l'Internet, sera la langue la plus fréquente en 2007. Déjà on constate que plus de 70 des 6,500 langues se trouvent sur l'Internet. En tant que l'expression "Internet pour tous" évolue d'un slogan à réalité, on a besoin d'un cadre pour répondre au multilinguisme de l'Internet. Heureusement, il y a un nouveau Multilingual Internet Name Consortium (MINC), établi à l'occasion de INET 2000 (Yokohama) qui s'occupe avec ces problèmes.

La grandeur de l'Internet est le sujet des forts débats. On parle d'un contenu de plus de 2.1 milliard de pages au surface avec plus de 550 milliards de pages si on veut inclure tous les banques de données et autres matériaux qui ne sont pas accessibles au grand public.⁴ On parle de 7 millions nouvelles pages par jour.⁵ Ces dimensions et tels changements posent des grands défis en ce qui concerne l'organisation des connaissances et les problèmes d'accès.

Autres dimensions de change concernent la vitesse des processeurs, la capacité des disques, et la vitesse de transmission. Dans les années 1980, les modems étaient typiquement limités à 200-400 Bauds par seconde. En 1995, l'exhibition du G7 sur la Société de l'Information a réalisé des vitesses réelles (throughputs) de 27 MB par seconde dans une démonstration entre le Canada (Vancouver, Ottawa, St. John) et Belgique (Bruxelles à travers Berlin). En 1998, on a eu des démonstrations d'un terabit par seconde; en 1999 de 3 terabits par seconde. On nous dit que dans cette décennie le pouvoir computationnel deviendra une million fois plus potente et que la connectivité sera dans les petabits. Donc les limites technologiques qui ont posé des contraintes fortes même dix ans auparavant, n'y seront plus dans le proche avenir: à moins qu'on parle de l'internet interplanétaire qui est au delà de la portée de cet essai.

3. Origines du Domain Name Spaces (DNS)

Quand Jon Postel a commencé à utiliser les Top Level Domains (TLDs) en 1984, ils ont servi comme moyen simple de classer les catégories principales d'utilisateurs. Cet approche a été formalisé en mars 1994 (RFC 1591) quand les TLDs génériques ont établis sept classes élémentaires, c'est à dire: gouvernement (.gov); éducation (.edu); commerce (.com); militaire (.mil); organisations (.org); réseau (.net) et international (.int).⁶ Cette liste de sept catégories de base sera agrandi pour inclure autre sept

catégories de base: .aero, .biz, .coop., .info, .museum, .name, .pro.⁷ Effectivement toutes ces catégories sont des différentes organisations. Donc on pourrait dire que le Top ou First Level Domain (1LD) .org, sert comme une classe avec 13 sous-classes.

Ces First Level Domains (1LDs) font partie d'une hiérarchie et ils contiennent des Second Level Domains (2LDs) qui à leur tour contiennent des Third Level Domains (3LDs) et sont positionnées du droit à gauche. Donc l'adresse Internet www.icann.org, la partie "org" est un First Level Domain (1LD), et "icann" est un Second Level Domain (2LD), tandis que www indique que le serveur utilise le world wide web (en contraste de ftp). Notre proposition est de renverser l'ordre: c'est à dire d'aller du gauche à droit du 1LD au 2LD etc.

Avec le développement de l'Internet autour du monde, ces noms de domaines hiérarchiques étaient combinés avec les codes de pays de deux lettres pris du ISO-3166 comme fr. pour France et de. pour Deutschland (i.e. Allemagne). Récemment, il y'avait en plus une proposition d'ajouter une code régionale comme eu. pour europe. Bien que la code .us existait pour les Etats Unies, elle n'a pas été utilisé dans leurs sites web jusqu'au novembre 2001 et ensuite ne que pour des espaces de noms de racines alternatives.

Au début de l'Internet quand il n'y'avait que quelques sites dans l'ouest des Etats Unies, traiter leur pays dans une manière spéciale avait du sens. Mais dans un monde multilingue et multiculturel ou les Etats Unies ne représentent que 4% de la population mondiale, il est mieux trouver une solution cohérente qui s'applique à tous.

Une solution plus logique serait de diviser l'adresse en trois parties distinctes: 1) le type de server (e.g. www, ftp); 2) une partie géographique; 3) une partie avec les noms de domaines lesquelles peuvent se rejoindre avec les systèmes de classification. Donc au lieu d'un adresse www.louvre.museum.fr.eu on aurait www.eu.fr.louvre.museum. On pourrait donner au www le sens de global. Donc on va (du gauche à droit) du global, au région (dans le sens de continent); au pays, à l'institution. Si, en plus, chaque institution reçoit une système globale de positionnement (GPS) et une système d'information géographique (GIS), l'adresse internet pourrait faciliter des recherches géographiques. En effet on pourrait lier ces adresses avec des gazetiers de sorte que l'adresse www.eu.fr.louvre.museum, puisse révéler la province, ville et section de ville comme si l'adresse aurait été [www.eu.fr.filedeFrance.Paris..6arrondissement\].louvre.museum](http://www.eu.fr.filedeFrance.Paris..6arrondissement].louvre.museum). Pour ceux qui ne sont point fort en géographie, une moteur de recherche pourrait leur indiquer ou le musée, monument ou lieu se trouve sur une carte.

4. L'Internet et Les Institutions de Mémoire Collective

Au début, l'Internet était en premier lieu une expérience par rapport à nouvelles méthodes de communication: notamment les connaissances personnelles (e-mail, MUDs, MOOs, MUSHs) et la connaissance collaborative. Dès l'essor du World Wide Web (1990), l'Internet s'occupe de plus en plus avec le savoir et les connaissances à longue durée des institutions de mémoire collective (les bibliothèques, musées et archives). Quelques visionnaires nous prédisent que dans le proche avenir toutes ce savoir et les

connaissances à longue durée sera tout à fait accessible sur ligne. Un défi sera comment intégrer ces nouvelles formes de connaissances personnelles et collaboratives avec les connaissances à longue durée. Étant donné que les conservateurs de ces connaissances à longue durée (en forme de bibliothécaires, curateurs, experts du catalogage et indexage) ont plus de deux millénaires d'expérience dans ce champ, il serait opportun d'en profiter.

Dès le début du World Wide Web on a été bien conscient que les utilisateurs de l'Internet ne pourraient guère disposer de la discipline et les habiletés professionnelles des bibliothécaires. Un projet initié (1995) par le (Dublin,) Ohio Computer Library Center, le Dublin Core (Metadata Initiative) a identifié un numéro minimal de 15 catégories lesquelles peuvent être utilisées à décrire les nouveaux contenus électroniques sur le web. Les efforts du Dublin Core étaient au début un projet américain. Dans les dernières années, le Dublin Core est devenu un projet international. On veut utiliser les catégories du Dublin Core et pour décrire nouveaux contenus et pour faire des liaisons (en anglais mappings ou bridges) avec les systèmes de classification et de catalogage des institutions du mémoire collective, comme les fiches MARC).

Nos propositions sont complémentaires à cette vision. Si on ajoute des noms clés aux noms de domaines, lesquels sont des vocabulaires contrôlés des systèmes de classification, ils peuvent servir comme première orientation en tant que l'ubication d'une site et la nature de ces contenus. L'initiative Dublin Core peut offrir des descriptions plus détaillées (bien que assez limitées quand même) pour les informations originalement digitales.

Dans le monde des bibliothèques et dans les institutions de mémoire collective en générale il y'a une tendance à lier les domaines des systèmes de classification (par exemple Dewey, Library of Congress et Göttingen) avec les domaines et titres des fiches de catalogage afin d'avoir des nouvelles stratégies d'accès. On voit la même tendance avec les systèmes de classification en art comme Iconclass et le Thesaurus pour l'Art et l'Architecture de la Fondation Getty. Ces systèmes de classification ont aussi des noms de sujets systématiques, des soi-disantes noms ou listes d'autorités (authority lists). Si on utilise ces méthodes dans les adresses Internet, on aura un vocabulaire contrôlé qui peut aider à accéder soit les informations nées sur Internet soit les connaissances dans les institutions du mémoire collective.

Dans le cas d'une collection spécialisée cela ne demandera que l'addition d'un seul mot à l'adresse. Dans le cas des collections complexes et compréhensives telles que le Louvre les diverses sections auront des suffixes différentes. Par exemple, en anglais la section sur la peinture Italienne serait www.eu.fr.louvre.museum.italianpainting, laquelle pourrait être encore plus spécifiée avec des descriptions spatio-temporelles comme www.eu.fr.louvre.museum.italianpainting.florentine.renaissance. Si toutes les termes dans ces descriptions viennent des vocabulaires contrôlés, les recherches sur ligne peuvent être co-ordonnées automatiquement avec les recherches de littérature dans les bibliothèques, musées et archives.

Pour être plus efficace, on peut introduire divers niveaux de recherche. Pour celui qui fait une première recherche générale le système pourrait se limiter à présenter les premières cinq catégories de l'adresse Internet, c'est à dire: www.eu.fr.louvre.museum. Au même temps, les adresses détaillées peuvent être indexés afin que les experts intéressés à la peinture Florentine de la Renaissance puissent accéder non seulement le Louvre mais en plus les collections des Offices, Brera, Galérie nationale de Londres etc.

Connecter les noms des domaines de l'Internet avec a) les vocabulaires contrôlés des noms de sujets et b) les systèmes de classification des bibliothèques, musées et archives a au moins trois avantages en plus:

1) Souvent ces systèmes sont traduites en divers langues. Donc l'adresse française du site Louvre serait www.eu.fr.louvre.musée, www.eu.fr.louvre.musée.peintureitalienne.florentin.renaissance et il serait potentiellement égale si une personne introduisait l'adresse en français, en anglais ou une langue quelconque.

2) Les grands efforts des institutions de mémoire collective de créer des liens (en anglais mappings, bridges) entre leurs systèmes pourrait être appliqués aux collections électroniques sur l'Internet avec un minimum de travail.

3) Ces différents systèmes de classification peuvent nous accorder aperçus en tant que la diversité avec laquelle les cultures organisent leur savoir et connaissances.

5. Défis pour l'Organisation des Connaissances et pour l'Accès

Global, National, Régional, et Local

Dans le passé, le lieu où l'information se trouvait a souvent fourni des indices significatives en tant que leur provenance et en plus des indices en tant que leur status. Des informations dans une bibliothèque de village, à part de quelques oeuvres standards de référence (comme des atlas, dictionnaires et encyclopédies) étaient typiquement des informations appartenantes à cette village. Des informations dans une bibliothèque régionale étaient en premier lieu sur cette région. Pareillement, les bibliothèques nationales ont réfléchi les collections de leur propre pays. Enfin il y'avait un petit nombre de bibliothèques universelles comme le Vatican, celle du Duc August à Wolfenbüttel et plus récemment le British, La Bibliothèque Nationale de la France et le Library of Congress. Leur bât était (et reste) celui de collectionner tout possible. Un problème profond était que personnes à chaque niveau ont choisi leurs propres règles pour le catalogage et l'indexage avec la conséquence qu'il n'existait point une système standard pour gagner accès au savoir et connaissances dans tous ces niveaux différents.

Dès la deuxième moitié du 19ème siècle il y'avait de plus en plus des efforts d'arriver aux standards internationaux à travers des organisations telles que l'Organisation de Standardisation Internationale (OSI), l'Union Internationale de Télécommunications (ITU), la Fédération International des Associations de Bibliothèques (IFLA), la Comité Internationale des Musées (ICOM) et UNESCO. Il y'avait une supposition tacite que les standards locaux et régionaux avaient une valeur limitée; que les standards nationaux étaient mieux, et que les standards internationaux étaient les meilleurs. Il y'avait l'espoir

que si on pourrait réussir à atteindre un accord international tous les problèmes seraient résolus.

Dans les sciences, la technologie, et divers domaines de la médecine cette vision est vraie. En ce qui concerne les définitions de zinc, d'une formule chimique ou l'aorte du cœur il faut qu'il y'eu un standard international si on veut que la science et la médecine puissent fonctionner à l'échelle globale.

Par contre, en ce qui concerne les arts et la culture, la situation est fort différente. Pour citer un exemple prosaïque de la bière, il est bien possible qu'il existe une définition internationale surtout en ce qui concerne les normes pour assurer que le processus de fermentation soit sain. Mais ce qui rend la bière si importante est qu'il y'a des variations nationales, régionales et locales. Dans les sciences c'est le consensus global qui rend important un loi ou une règle une telle. Par contre, dans la domaine de la culture, c'est précisément dans la richesse des variations nationales, régionales et locales, qu'on détermine l'importance. Si la science cherche à trouver les aspects non variants d'une loi ou d'une règle, la culture cherche à noter les exceptions changeantes d'une règle parce que c'est là où se trouve l'unicité et l'originalité que nous associons avec l'art et la culture.

Donc tandis que les organisations globales puissent répondre aux besoins de la science et la technologie, des organisations globales comme UNESCO et ICOM ne peuvent guère répondre aux besoins de l'art et de la culture s'ils visent uniquement aux standards globaux de leurs bureaux centrales à Paris ou Genève. On a besoin d'une nouvelle coordination du savoir et des connaissances aux niveaux locaux, régionaux, nationaux et globaux: constatation qu'a fait récemment aussi la Commission Ruffolo.⁸ Autrement on n'aura que les visions souvent impérialistes des grands centres. Ce qu'il nous faut pareillement c'est accès à toute la richesse et diversité de l'expression culturelle: toutes les variations et alternatives de divers pays, régions, villes et villages. C'est là qu'on trouve un vrai espoir à dépasser les dichotomies entre (jihad) régional et (McWorld) global que nous a proposé Barber.⁹

Le World Wide Web (WWW) représente une avance notable en tant qu'il contient des informations de tous les niveaux. Pour la première fois nous avons un instrument de communication qui possède du savoir et des connaissances locales, régionales, nationales et internationales. Hélas, le WWW tel qu'il est aujourd'hui, manque un système pour indiquer de quel niveau vient l'information: il n'y a point pour déterminer si les informations sont locales, régionales ou nationales.

Dans cet'égard la connection proposée entre les institutions et a) les systèmes de positionnement globales (GPS); b) les Systèmes d'Informations Géographiques (GIS) et c) les gazetiens nous permettrait à établir des liens nouveaux entre local, régional, national et international, ce qui est essentiel pour l'évolution de moteurs de recherche significatives au domaine de la culture. Actuellement, si je cherche des exemples de *l'Annonciation* je peux soumettre cette parole dans les moteurs de Yahoo ou Hotbot. Mieux, je peux me servir de la fonction recherche d'image (image search) dans Google.

Quand même le résultat est une liste énorme sans aucune ordre logique ou spatiale. Par contraste, s'il pouvait être organisée par nation, région et localité, il serait possible à déterminer comment les *Annonciations* du Euregio sont différentes de celles de la Toscane ou de Catalogne. Un projet IST, IMASS, s'occupe avec des tels problèmes dans le contexte d'une salle de consultation virtuelle (anglais: Virtual Reference Room).¹⁰

La Dimension Temporelle

Une autre différence fondamentale entre les bûts de la science et ceux de la culture concerne la dimension temporelle. Dans la science on cherche à découvrir les lois et règles qui sont au delà du temps, qui ne changent point. Par contre, dans les arts, bien qu'on dise qu'une chose de beauté est une joie pour toujours (Keats) et bien que nous parlames de la beauté hors du temps d'une statue grecque, le temps joue un rôle fondamentale. Dans les sciences si une loi change avec le temps il faut le rejeter comme une hypothèse qui ne soit point une vraie loi. Par contraste si une statue grecque ou un temple romain a besoin de restauration, personne ne rejetterait pour être moins vrai ou moins réel à moins qu'il soit tellement réduit en ruine qu'on ne réussit point à le reconnaître. Donc, malgré les dictons d'un art hors du temps, en réalité le temps joue un rôle essentiel dans l'art et la culture: Quand un objet fut construit; quand il a été altéré; quand il a été restauré; parfois quand il a été perdu et quand il a été retrouvé, et surtout quand qui a pensé quoi sur l'objet d'art ou une personne de la culture.

Nos sites web ont besoin d'une dimension temporelle et donc d'une dimension historique. On a besoin d'une indication quand un site a été lancé, avec une section pour indiquer toutes les fois que le site a été modifié et non seulement la dernière fois, afin que nous puissions à l'avenir tracer l'évolution des sites.

Cartes Dynamiques

Cette dimension temporelle doit être incluse dans toutes nos connaissances. Par exemple, dans le cas des cartes géographiques, des sites web comme Map Quest me donnent des cartes contemporaines. Mais, en tant que historien, si je cherche pour la Pologne aux alentours de l'an 1000 je cherche pour un petit pays. Si je cherche pour Pologne en 1440, je cherche pour le plus grand pays de l'Europe de l'époque. Donc il faut des cartes dynamiques lesquelles puissent s'adapter aux changements temporels.

Les Connaissances Dynamiques

Comme on a déjà signalé il y'a un grand contraste entre les domaines de la science et de la culture. Dans la science on cherche une définition fixe laquelle ne va pas changer dans divers lieux ou divers temps. Dans la culture les définitions qui changent dans divers lieux au travers du temps jouent un rôle fondamentale. Les opinions changeantes des érudit(e)s concernant le nombre de peintures faites par Rembrandt ou la valeur relative de Shakespeare sont les bases de la vraie culture. Les listes statiques du savoir telles qu'elles existent aujourd'hui ne sont pas suffisantes. On a besoin de banques de données avec des listes dynamiques, qui nous révèlent les différences historiques et culturelles à travers

diverses interprétations et diverses théories. Dans cette diversité se trouve la richesse de nos traditions culturelles et pour y arriver on a besoin d'un projet à long terme, Metadonnées et Mémoire Collective Systématique (MEMECS), lequel prend comme point de départ la vision de Vannevar Bush avec son système Memex et y ajoute les dimensions multi-lingues et multiculturelles centrales à l'approche européenne.

Pour y arriver on a besoin d'une nouvelle coopération entre les universités et les instituts de recherche, à travers un réseau à haut débit de Centres d'Excellence pour le patrimoine digitale, suivant la vision d'une Espace Européenne de Recherche (ERA). Un tel réseau peut s'épanouir à inclure le réseau japonais à haut débit, le projet Japonnais/UNESCO pour les routes de soie digitales, et le portail global de l'UNESCO pour la culture et e-enseignement (anglais: e-learning).

Au début des années 1990s, il y'avait une grande enthousiasme pour la réalité virtuelle. Les dernières années nous ont introduit aux potentielles de la réalité augmentée: ici la réalité virtuelle sert comme une couche transparente super-imposée sur le monde physique. Donc on peut ajouter aux ruines de l'ancienne temple de Hera (Olympia) une reconstruction de comment elle a été dans l'Antiquité, comme on fait dans le projet Archeoguide.¹¹ Cette approche peut être étendue afin d'inclure diverses interprétations concernant le meme temple, ou pour visualiser comment les constellations grecques sont différentes de celles des Persans, Indiens, Chinois, et des Mayas. Donc la culture augmentée peut littéralement nous aider à voir comment les différentes cultures apperçoivent le monde dans différentes manières.

6. Culture par Rapport au Réseau Sémantique et L'Intelligence Machinelle

Le World Wide Web cherche à créer une toile d'araignée globale de raisonnement ou un réseau sémantique. Ici un des buts est de produire de l'information "lisible" aux machines, laquelle n'a plus besoin de l'intervention humaine. Une version extrême de cette quête cherche à aller beaucoup plus loin: à créer des robots autonomes, intelligents capable à faire des commandes et décisions sans aucune intervention humaine. Ce qui explique l'essor d'intérêt pour la langage naturelle, les soi-disantes ontologies de sens-commun (comme enCYClopædia). Bien que l'attraction militaire de telles solutions, lesquelles n'ont plus besoin de la présence humaine et même pas de l'existence humaine, soit évidente, une telle vision pour un avenir robotique ne peut guère répondre aux besoins de l'humanité et de la culture.

7. Conclusions

Nous avons paasé en revue la croissance de l'Internet à travers les années 1980s et 1990s. On a vu que l'usage un peu à l'hasard des Top level Domains (TLDs), Second Level Domains (SLDs) et les Third Level Domains (3LDs) restait acceptable lorsque l'Internet était surtout une phénomène américaine avec l'anglais comme la langue dominante. Dès ce temps on a constaté trois développements d'une nature fondamentale:

1) il y a maintenant plus de 513 utilsateurs/utilisatrices de l'Internet;

- 2) l'usage est multi-lingue en plus de 70 langues différentes où anglais ne représente que 43% du total; un pourcentage qui va descendre vers 25% entre trois ans;
- 3) les contenus de l'Internet qui étaient surtout des nouvelles méthodes de communication et connaissances personnelles et collaboratives, deviennent de plus en plus liées avec le savoir et les connaissances traditionnelles des institutions de mémoire collective (les bibliothèques, musées et les archives).

En face de cette évolution on propose une réforme dans les noms des domaines de l'Internet. Tout d'abord, on propose de les nommer First Level Domains (1LDs), Second Level Domains (2LDs) et Third Level Domains (3LDs) pour être plus cohérents en tant que noms et leurs abbréviations. Ensuite on propose que l'adresse Internet aura effectivement trois parties: 1) partie pour identifier le moyen de communication (c'est à dire, www, ftp, etc.); 2) partie pour établir la dimension géographique de l'adresse, laquelle peut être liée: a) aux systèmes de positionnement globales (GPS); b) aux Systèmes d'Informations Géographiques (GIS) et c) les gazetiers; 3) une partie conceptuelle commençant avec les classes, suivies par des sous classes et des sous-sous classes, liées à leur tour aux systèmes de classification des institutions de mémoire collective. Tandis que les vieux TLDs, SLDs et 3LDs étaient écrites du droit à gauche, on propose qu'à l'avenir leur ordre serait de gauche à droite.

Ces réformes peuvent assister la lecture machinelle pour rendre plus efficace les moteurs de recherche. En plus ils peuvent aider et les professionnels et l'utilisateur/utilisatrice quotidien(ne) à s'en servir de l'Internet et, en le rendant plus transparente et fiable, leur donner plus de confiance au système.

Ces réformes sont conformes avec autres initiatives comme le Dublin Core et le réseau sémantique à travers laquelle les expériences des institutions de mémoire collective servent de base pour une description plus systématique des ressources et pour rendre plus accessible leurs contenus. D'autre part, tandis que ces projets ont un modèle statique qui reflète en premier lieu les exigences de la science et la technologie, notre vision prévoit l'évolution des connaissances dynamiques avec une version de culture augmentée lesquelles reflètent la richesse de la culture, par établir des liens nouveaux entre le savoir local, régional, national et international.

Pour réaliser cette vision multi-culturelle et multi-lingue on propose un projet à long terme, Metadonnées et Mémoire Collective Systématique (MEMECS), ce qui entamera une nouvelle coopération entre les universités et les instituts de recherche, à travers un réseau à haut débit de Centres d'Excellence pour le patrimoine digitale, suivant la vision du Commissaire Busquin d'un Espace Européenne de Recherche (ERA), laquelle peut s'élargir vers un Espace de Recherche Globale en se servant d'une grille (anglais: Grid) pour la culture. En utilisant les méthodes de nommage et indexage du passé, nous pouvons arriver dans l'avenir à une nouvelle vision panoptique du réel.

Remerciements

Je veux bien remercier Richard Delmas, qui m'a invité à écrire cet article et qui a suggéré des corrections; mon collègue, John Beckers qui a proposé des améliorations et Monsieur Bob Shaw (ITU), pour des corrections au domaine technique.

Notes

¹ Voir: <http://www.icann.org/>.

² Voir: http://www.nua.net/surveys/how_many_online/world.html.

³ Voir: <http://www.gtreach.com/globstats/evol.html>

⁴ "Web Is Bigger Than We Thought, *San Jose Mercury News*, 28 July, 2000,

Voir: http://www.nua.ie/surveys/?f=VS&art_id=905355941&rel=true

Mark Frauenfelder, "Deep-Net Fishing," *Industry Standard Magazine*, June 18, 2001.

Voir: <http://www.techinformer.com/go.cgi?id=490878>.

Mike Nelson (IBM) at the INET 2001 Global Summit claimed that in the next 9 years there will be a 1 million fold increase in information. This amounts to the equivalent of 550,000,000,000,000,000 pages.

⁵ Cyveillance report, 10 July 2000.

Voir: <http://www.cyveillance.com/us/newsroom/pressr/000710.asp>

⁶ Voir: <http://www.iana.org/reports/museum-report-30oct01.htm>

⁷ Voir: <http://www.icann.org/tlds/>.

⁸ *The Unity of Diversities. Cultural Co-operation in the European Union*, ed. by The Parliamentary Group of the PSE, European Parliament, Florence: Angelo Pontecorboli Editore, 2001.

⁹ Benjamin R. Barber, *Jihad vs. McWorld*, New York: Times Books, 1995.

¹⁰ Voir: <http://www.i-massweb.org/>.

¹¹ Voir: <http://archeoguide.intranet.gr/project.htm>.