

## Sciences de l'information - Rappel de quelques dates

**1940** – Johannes Von Neuman (Université de Princeton) pose les bases des architectures informatiques actuelles. Les premiers outils de RI vont naître dans les universités et les industries militaires américaines.

**1957** - Suite au lancement du premier Spoutnik par les Soviétiques, le président Dwight D. Eisenhower crée l'ARPA (Advanced Research Project Agency) au sein du DoD (Department of Defense) pour piloter un certain nombre de projets dans le but d'assurer aux USA la supériorité scientifique et technique sur leur ennemi russe.

**1972** - L'ARPA est renommée DARPA (Defense Advanced Research Projects Agency).

**1974** - Vinton Cerf (Université de Stanford) et Robert Kahn (DARPA), mettent au point le protocole TCP (Transmission Control Protocol) et ce qui deviendra le protocole IP (Internet Protocol). Cerf et Kahn sont considérés comme les créateurs de l'expression Internet. Ces deux protocoles représentent la fondation d'Internet.

**1980** - Vinton Cerf crée une passerelle entre ARPAnet et CSnet (protocole TCP/IP), puis DARPA lève le secret militaire sur TCP/IP et le met à disposition gratuite du public.

**1989** – Genève (Suisse) - Timothy Berners Lee et et Robert Cailliau inventent le lien hypertexte, qui marque la naissance du world wide web (web, www ou w3).

**1991** - Le World Wide Web est mis à la disposition du public.

La recherche d'information, est au carrefour de nombreuses disciplines scientifiques : linguistique, intelligence artificielle, architecture informatique...

Détenir l'information, la trouver plus vite que les autres, est le nouveau défi stratégique.

Technologiquement, c'est l'ère de la recherche d'informations en texte intégral (full text) ; les machines fouillent des bases de données à l'aide de « mots clés » enregistrés dans des « index ».

**Le NIST (National Institut of Standards Technology) et DARPA/ITO (Defense Advanced Research Projects Agency - Information Technology Office) lancent la conférence annuelle TREC.** TREC est un système d'évaluation de RI, il encourage les travaux de recherche en informatique documentaire en réunissant chercheurs et industriels qui confrontent leurs savoirs sur les techniques et les systèmes développés, favorisant les discussions et les échanges entre groupes participant. Les résultats obtenus dans TREC sont chaque année résumés dans des actes édités par le NIST.

**1994 – Naissance du W3C**

**En Europe, l'INRIA (France), en accord avec le CERN, prend en charge l'organisation du Web. Son interlocuteur outre Atlantique est l'Institut de Technologie du Massachusetts (MIT). Le Consortium du World Wide Web (W3C) naît en octobre.**

Création de Yahoo.

**1995, décembre** - Les chercheurs du laboratoire de la société DEC (Digital Western Research Lab, Pao Alto, Californie), à l'initiative d'un français, Louis Monier, lancent **AltaVista** qui fouille le web en texte intégral.

D'autres inventent un nouveau concept, celui des **metadatas (ou métadonnées)** pour désigner les données secondaires décrivant des données primaires. C'est une avancée scientifique. Les metadonnées vont permettre aux « agents » informatiques de parler entre eux. C'est l'essor d'un nouveau langage de communication entre machines.

**1998 - Larry Page et Sergey Brin (Université de Stanford) créent Google**, un moteur de recherche basé sur le concept de « popularité » des pages web.

**2000 – DARPA (Defense Advanced Research Projects Agency) lance le projet « Darpa Agent Mark up Language » (DAML).** L'agence du département de la Défense américaine est à l'origine du développement d'un langage et d'outils spécifiques qui donneront naissance au « **Web Sémantique** ».

**2001 – Le gourou, Tim Berners Lee, à la tête du W3C, a une vision : le «Web Sémantique** », un web où les machines commenceraient à comprendre le « sens » des instructions de l'homme, grâce aux progrès de la science dans le domaine des métadonnées, de la linguistique, de la sémantique et des langages de communication entre « agents ».

**2003 - La R&D aux Etats-Unis est dans une situation florissante**

*Sources : Ambassade de France aux Etats Unis, Mission pour la Science et la Technologie.*

En 2003, les Etats-Unis auront disposé de près de \$300 milliards pour le financement de leur R&D. Le seuil des 100 milliards (dollars constants 2002) avait été atteint vers la fin des années 60, celui des 200 milliards au tout début des années 90. La loi de programmation 2004 pour la seule R&D fédérale américaine aura, encore une fois, atteint des records : 127 milliards de dollars, soit 9,5 milliards d'augmentation par rapport à 2003 (+8,1%).

**2004, 10 février – Le W3C annonce que le RDF (Ressource Description Framework) et l'OWL (Web Ontology Language) ont acquis leur statut de « recommandation » officielle.**

L'annonce de cette « recommandation » marque l'émergence du **Web sémantique comme plate-forme commerciale générale pour les données sur le Web. Le déploiement de ces standards dans des produits et services à caractère commercial marque un tournant pour cette technologie.**

Jusqu'à présent considéré comme un projet de recherche et de développement avancé dans les cinq dernières années, le Web sémantique est devenu une technologie pratique, à utiliser massivement dans des outils de production permettant l'accès à des données structurées sur le Web.

**Le Web sémantique est l'aboutissement d'un progrès scientifique majeur**, aux frontières de l'Intelligence Artificielle, essentiellement développé et financé par la première Agence du Département de la Défense américaine : DARPA (US Defense Advanced Research Projects Agency), qui a doté ses entreprises et universités d'une connaissance technologique d'avant-garde, devenue la nouvelle « norme » de l'Internet mondial.

La naissance de l'**OWL** marque la fin, notamment sur le web, des systèmes de recherche d'informations classiques (Google, Yahoo,...), sans les condamner, en leur permettant d'évoluer vers une nouvelle dimension de la gestion des connaissances.

Elle célèbre aussi la réussite impériale de la R&D américaine, qui, en finançant massivement les technologies de l'information depuis 10 ans, dispose du plus grand portefeuille de brevets dans le domaine et de la plus grande réserve de chercheurs épanouis !

*Auteur : Boris Debove, e-mium SARL ©*

*Extraits - Dossier Stratégique sur la Recherche d'Informations.*