Simposietto - 9 mai 2009

Quel est le grain de la connaissance ? Questions et outils potentiels

Jean-Daniel ZELLER
Archiviste principal
Hôpitaux universitaires de Genève
Enseignant vacataire
Haute école de gestion de Genève

Le titre de ma présentation illustre bien la dimensions schizophrénique qu'elle implique.

D'un part la formulation « grain de la connaissance » renvoie à une dimension philosophique et épistémologique assez abstraite,

D'autre part les questions et les outils potentiels répondent eux à des préoccupations très concrètes.

Voyons le programme…

Grain de connaissance ? - Genèse

Interrogations sur la structuration de la connaissance en train de se faire, en vue de sa transmission et de son assimilation.

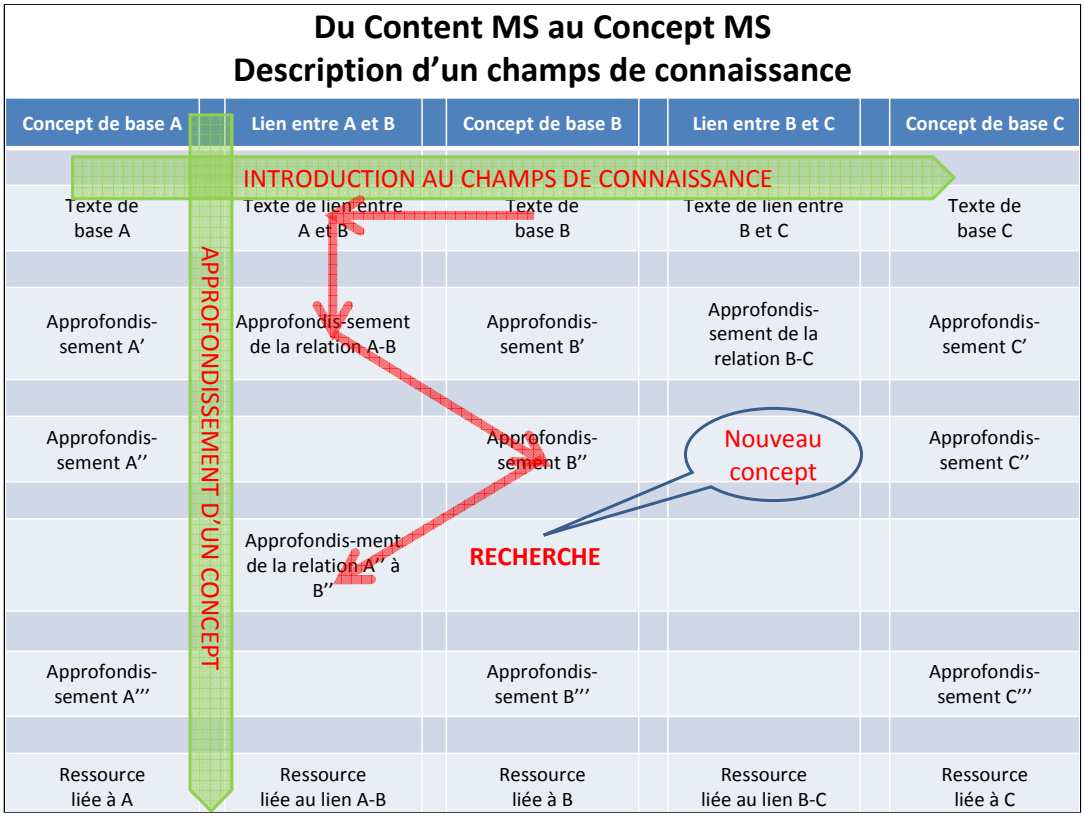
- A partir d'une pratique de l'enseignement d'un savoir en cours de constitution (en l'occurrence l'archivage électronique) et d'une pragmatique pédagogique (comment enseigner des incertitudes ?), je me suis interrogé sur la granularité des éléments de connaissance à transmettre.
- La difficulté réside dans le fait que cette granularité doit être congruente entre le plan conceptuel (les idées manipulées par la mémoire immédiate et à long terme) et le plan physique des information transmises (paragraphe, page, chapitre, …objet descriptif en général).
- Répondre à cette condition permettrait de constituer des briques (atomes ?, grain ?) de savoir librement associables pour construire une connaissance utile en fonction de besoins particuliers et évolutifs.
- Je présente différents concepts et outils qui seraient possiblement utiles à cette construction.

Cette réflexion ressort plutôt du plan moléculaire (= chimique) plutôt que du plan de la physique. J'ai néanmoins préféré garder le terme grain plutôt qu'atome ou molécule dans le sens ou je ne sais pas encore si l'analogie moléculaire est pertinente.

Il me semble néanmoins que ce plan intermédiaire est justement l'enjeu entre la physique (micro) et la psychologie (macro).

La forme même de cette présentation est une sorte de prototype de la **construction de liaisons à travers un corpus de concepts** qui est l'objet de mon exploration, à la distinction près que ce corpus n'est pas pré-constitué mais en devenir.

Par ailleurs mon intérêt pédagogique personnel est de pouvoir utiliser les mêmes éléments pour enseigner à des nouveaux différents (CfC, HES, Uni), seul l'arrangement, la succession, ou l'exhaustivité différenciant ces niveau d'enseignement, comme va le montrer le tableau suivant….



Ce tableau représente un champ de connaissances constitué d'items conceptuels relié entre eux. Il illustre mon hypothèse qui est que nous devons passer d'un Content management system (une gestion technique de contenus qui sont plutôt de contenants), à un Concept management system (qui relie des contenus en fonction même des contenus conceptuels et non de leur forme technique).

La forme carrée est une contrainte de présentation et elle est/sera dans la réalité beaucoup moins régulière (voir dias 6 et 18).

Dans l'hypothèse où ce champ est constitué, il peut être lu horizontalement (clic) et la présentation des concepts de base et de leurs textes associés constitue une **introduction au champ de la connaissance** (dans ma version initiale j'avais fait le lapsus textae significatif en écrivant contextes à la place de concepts).

Une lecture verticale (clic) correspond à l'**approfondissement d'un concept** ou de la relation entre une paire de de concepts.

Enfin une lecture « chaotique » (4 clic) correspond à un **processus de recherche**, susceptible d'aboutir à la création d'un (clic) nouveau concept.

Source :
Du Content management system au Concept management system
Une lecture critique de l'article de C. Lagoze et al. « Qu'est-ce qu'une bibliothèque numérique au juste ? »
Jean-Daniel Zeller
http://archivesic.ccsd.cnrs.fr/docs/00/15/43/52/DOC/Du_CMS_au_CMS.doc

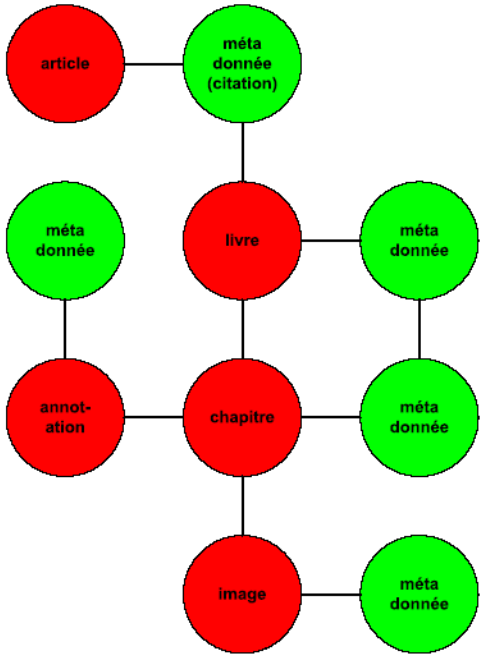
Du Content MS au Concept MS
Description d'un champ de connaissance

- Les flèches INTRODUTION et APPROFONDISSEMENT correspondent à une vision orthogonale « classique » de la connaissance.
- La flèche RECHERCHE étant indéterminée dans sa (ses) direction(s) ne peut être guidé que par des « balises » (tags), qui ont des positions relatives.
- L'ontologie sous-jacente du champ doit elle être préétablie ou peut-elle être généré par le réseau de balises de manière dynamique ?

La dernière question renvoie aussi à une question pédagogique qui est la suivante :

Est-il plus efficace de présenter un corpus théorique constitué, puis de l'illustrer par des exemples; ou bien de présenter des exemples, puis de déduire une théorie élaborée à partir des questions que posent les exemples ?

Relations conceptuelles entre les différents objets
de la base de connaissance (selon C. Lagoze)



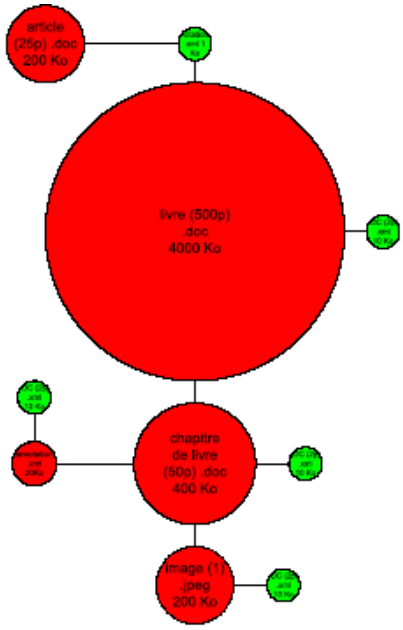
Dans ce schéma les points rouges représentent du contenu et les points verts (métadonnées) du liant entre les contenus.

La grande nouveauté de cette présentation, élaborée par C. Lagoze et se collègues de la bibliothèque de l'Université de Cornell, est l'importance attachée aux métadonnées qui « lient » les différents éléments (textuels ou autre) qui forment le champ de connaissances.

Cependant, cette représentation est à mes yeux critiquable … (clic)

Source:
Qu'est-ce qu'une bibliothèque numérique au juste ?
Carl Lagoze et al.
http://dilib.inist.fr/~ducloy/AMETIST/public/0/LAGOZE/

Relations conceptuelles «pondérées»
entre les différents objets de la base de connaissances
(compte tenu de la disproportion effective, l'échelle n'est pas exactement respectée)

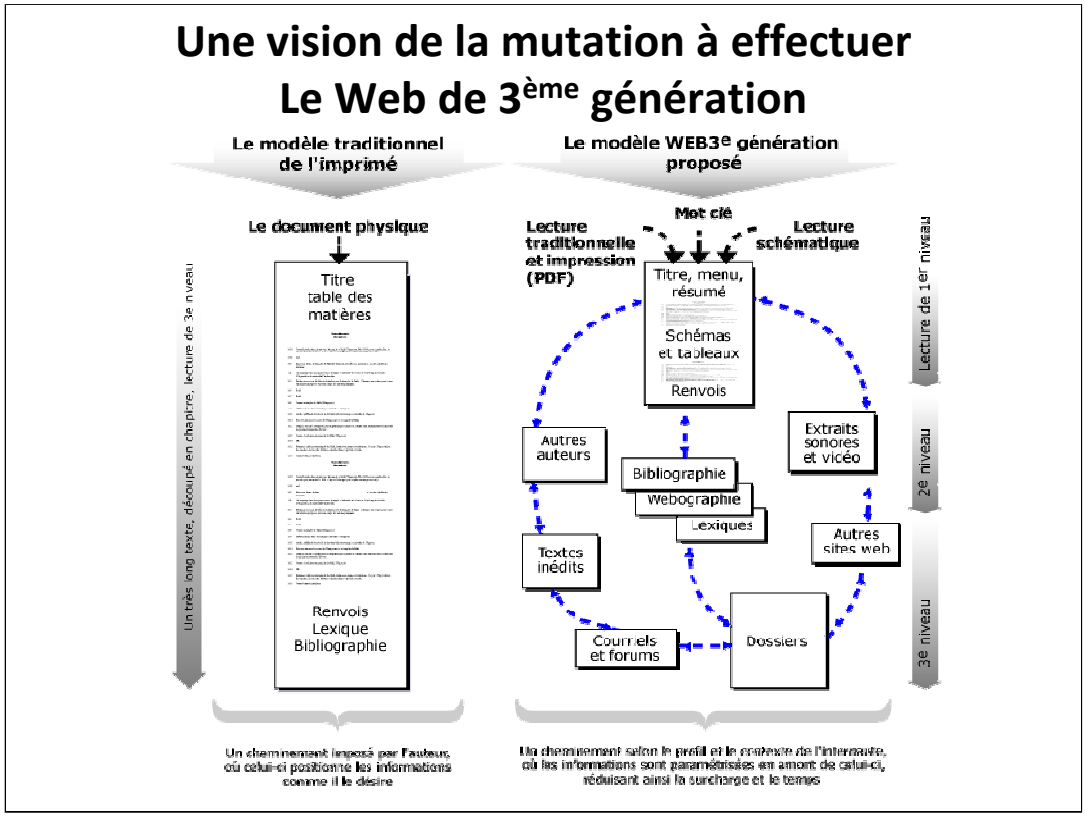


Voici la même image pondérée par les volumes effectifs des éléments.

Il saute aux yeux que le contenu livre est disproportionné par rapport aux autres éléments du système conceptuel.

Ceci m'a amené à proposer des compléments aux proposition de C. Lagoze.

Source :
Du Content management system au Concept management system
Une lecture critique de l'article de C. Lagoze et al. « Qu'est-ce qu'une bibliothèque numérique au juste ? »
Jean-Daniel Zeller
http://archivesic.ccsd.cnrs.fr/docs/00/15/43/52/DOC/Du_CMS_au_CMS.doc



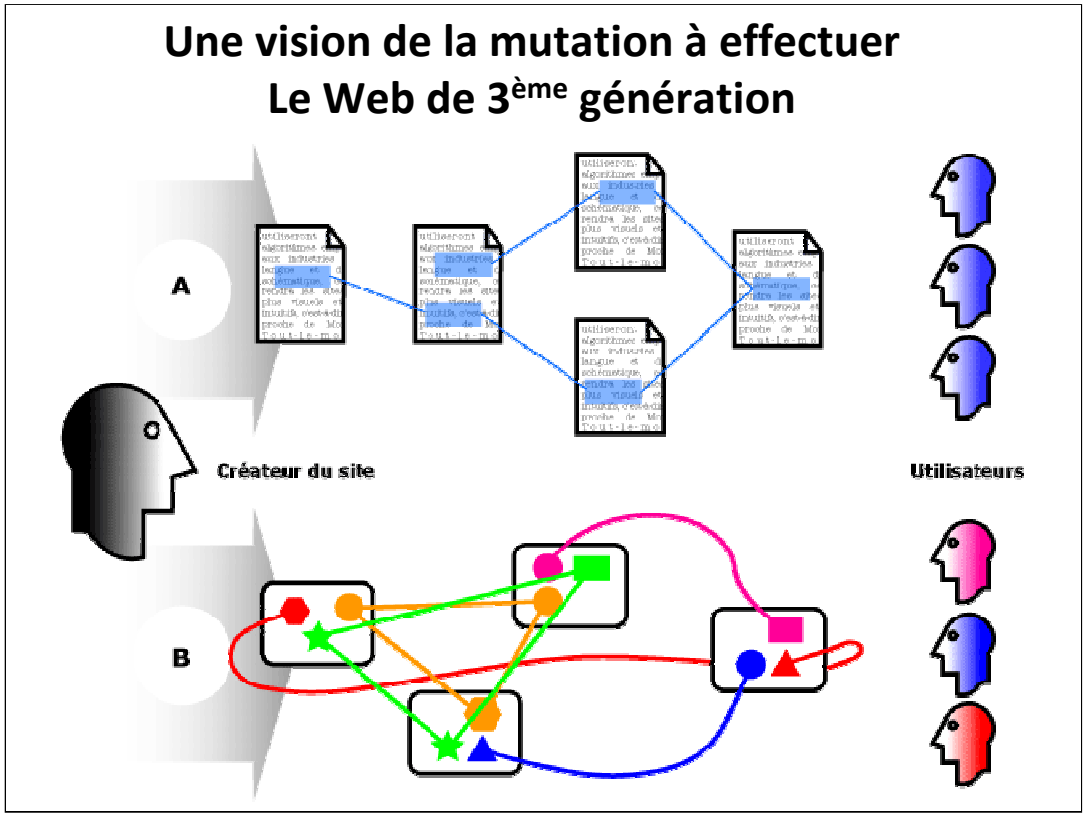
Cette analyse est celle de Michel Cartier, auteur canadien qui réfléchi sur le web de troisième génération.

Mais les éléments proposés sont décrits par leur caractéristiques techniques, bien que la dimension de recherche tel que je l'ai exposée dans le premier tableau est aussi présente dans ces recherche. (clic)

Citation M. Cartier (mise en gras par JDZ):
Contrairement au long texte imprimé, découpé en chapitres qui sert de modèle aux Web 1 et 2, le Web de troisième génération doit offrir ***plusieurs chapitres courts et beaucoup de liens*** *vers des documents destinés à une lecture de niveau 1, 2 ou 3*

Dans un premier temps, ***on peut établir un certain parallèle entre le réseau des réseaux et les réseaux neuroniques du cerveau****. Chez l'être humain, qui possède un cerveau moins gros que certains animaux, ce n'est pas tant les milliards de neurones qui importent mais la complexité de leurs connexions synaptiques. Créer, aimer ou penser devient possible par l'action intégrée des aires de spécialisation faites de ces connexions synaptiques. Il en va de même pour l'Internet, où* ***ces connexions entre les contenus, via les utilisateurs (par comparaison, ceux-ci deviennent les synapses), commencent déjà à développer une forme d'intelligence distribuée****, dont les effets dans plusieurs milieux deviendront majeurs vers 2004 ou 2005. C'est la granulation des connaissances et l'augmentation du nombre de connexions qui rendent possible l'intelligence distribuée ; mais celle-ci prend différentes formes selon les pressions exercées par l'environnement : le milieu, le facteur temps, les circonstances, les NTIC utilisées, le style narrowcasting, etc.*

Source :
MICHEL CARTIER
Le Web de 3e génération
Comment aider l'internaute à naviguer
http://www.michelcartier.com/McArticleB.php3?id_article=276



La partie A de l'image représente la navigation aujourd'hui habituelle sur l'Internet.

La partie B représente une navigation multi-dimensionnelle.

Citation de M. Cartier:
Schéma 20. La navigation via les textes et celle via les schémas utilisent des modes de pensée différents

Un utilisateur qui circule d'un texte, ou d'un mot, à un autre, navigue de façon assez linéaire *parce qu'il va d'un détail à un autre. De plus, parce que cette approche est de type broadcast, c'est-à-dire préprogrammée par le créateur du site pour être lu par tous les utilisateurs quel que soient leur profil, elle donne des résultats assez traditionnels. Cette lecture de troisième niveau, caractéristique des sites de type W1 et W2 et de leurs portails généralistes, donne des résultats sans surprise.*

Un utilisateur qui va d'un schéma à un autre connaît un autre type de cheminement*. Les résultats sont très différentes car cette navigation est plus personnalisée c'est-à-dire réagissant au vécu et aux besoins du visiteur. Cette approche narrowcast convient aux portails thématiques et personnalisés, donc aux niches. Le style W3 est plus visuel ; il utilise un degré différent de synthèse et un niveau d'interactivité plus riche. Cette lecture combinant les 1e 2e et 3e niveaux est plus intuitive et offre des découvertes par court-circuit (insight)*

Ce que l'auteur classe dans W1 et W2 ressorte de l'axe **Introduction** et **Approfondissement** de mon premier tableau (dia 3).
Le schéma B correspondrait à l'axe **Recherche**.

Source :
http://www.michelcartier.com/McArticleB.php3?id_article=278

Exemple d'un objet concept (selon J-D Zeller)

- Lien sur
  Qu'est-ce qu'un document.ppt

Pour illustrer ce que je considère comme un objet concept j'ai joint un annexe une autre présentation powerpoint que j'utilise dans mon cours à propos du concept de document dans un environnement électronique.

Techniquement, il s'agit à minima d'une image/tableau, lisible soit dynamiquement avec un commentaire sonore chronologique, soit diachroniquement en tant qu'image unique et son commentaire textuel.

Je présente brièvement son fonctionnement sans m'attarder sur le contenu même car c'est la forme qui nous intéresse ici.

Quelques pistes…

- Le **TRUC**
  Testable Reusable Units of Cognition
- **Little BANG**

  La nanotechnologie comme interface psychique ?
- La **Map of Science**

  Les liens réels versus un modèle théorique ?

Suite à mes premières réflexions, j'ai tenté de trouver des pistes de solution dans les recherches existantes qui pouvaient avoir un lien avec ma recherche.

Le TRUC

Testable
Reusable
Unit of
Cognition

Le TRUC est l'acronyme de Testable Reusable Unit of Cognition que l'on peut traduire par Unité de cognition testable et réutilisable.

Unité de cognition testable et réutilisable

- Concept développé par Bertrand Meyer (EPFZ)
- Pour des éléments de formation structurés en programmation informatique
- Défini comme une « collection de concepts, de compétences opérationnelles et critères d'évaluation, avec différentes propriétés »
- Défini comme étant autosuffisant pour enseigner un concept et évaluer l'assimilation de ce concept par les étudiants

C'est (voir le texte)

Il correspond assez bien aux besoins exprimés puisqu'il porte :
- sur la formation
- vise des compétences
- avec des critères d'évaluation (j'y reviendrais)
- autosuffisant

-Regardons cela plus en détail …

Propriétés du TRUC

- **Les composants sont fortement cohérents**
  (ils procèdent d'une idée centrale clairement identifiée)
- **Les composants sont clairement définis (délimités)**
  (ce qui permet leur utilisation par différents enseignants pour différents étudiants)
- **Le TRUC inclus des critères d'évaluation**
  (ce qui permet de s'assurer que l'étudiant maîtrise les concepts et les compétences associées)
- **L'objectif des différent thèmes est d'intérêt général**
  (par conséquent pas liés à un cursus spécifique)
- **L'objectif est suffisamment petit pour être couvert par une (ou quelques) leçon(s)**
  (ou un chapitre, ou quelques sections d'un chapitre)

Composants cohérents :
= réseau de liens autour d'une idée centrale (voir le champs de connaissance, dia 3).

Composants clairement définis :
= recyclable.

Inclus des critères d'évaluation :
= voir note finale… (à faire).

Intérêt général :
= pas de contrainte induit par un enseignement donné.

Objectif petit :
= une granularité optimum.

Malheureusement, depuis la parution de l'article en 2006, B. Meyer n'a pas proposé de réalisation concrète de son « TRUC », à par certains cours offerts dans le cadre de l'ETH.

Source :
Testable Reusable Units of Cognition
Bertrand Meyer, ETH Zurich
Computer
Volume 39, Issue 4, April 2006 Page(s): 20–24
http://se.ethz.ch/~meyer/publications/computer/truc.pdf

Little BANG – la théorie

Les technologies convergentes :
ce qu'elles sont et leur addition qui fait BANG !

Le vrai pouvoir de la nanoscience réside dans la convergence de plusieurs technologies – **biologie, sciences cognitives, informatique**, robotique, etc. – dont la nanotechnologie est l'élément habilitant essentiel. La logique inhérente à la convergence technologique, c'est que **les constituants de toute matière, d'une importance fondamentale pour toutes les sciences, prennent naissance à l'échelle nanométrique.**

Les scientifiques et les gouvernements d'Europe et des États-Unis ont pour stratégie de fusionner les sciences fondées sur « l'unité du matériau à l'échelle nanométrique ». **Puisque tous les matériaux et processus opèrent du bas vers le haut (à partir des atomes qui se combinent pour former les molécules et toutes les structures supérieures), les tenants de la convergence pensent qu'ils peuvent contrôler les réactions à l'échelle macroscopique en manipulant les réactions à l'échelle nanométrique**. Selon cette vue réductionniste, toute substance – comme tout système biologique ou culturel – est le résultat de processus moléculaires opérant à divers niveaux.

Ces éléments de définition sont tirés de « Un infiniment petit guide d'introduction aux technologies à l'échelle nanométrique…et à la théorie du petit BANG », ouvrage stigmatisant l'usage inconsidéré des nanotechnologies et la posture des milieux transhumanistes (http://fr.wikipedia.org/wiki/Transhumanisme) qui en promeuvent la version anglaise sous l'acronyme NBIC (Nanotechnologie, Biologie, Informatique, sciences Cognitives).
Je n'en retiens ici que les éléments en rapport avec ma réflexion sur la granularité du savoir.

Biologie , science cognitive et informatique touchent également les sphères de ma question.
L'hypothèse est que tout peut se résoudre par la convergence au niveau nano.
Est-ce vrai au niveau de la gestion des concepts ?

Source :
Un infiniment petit guide d'introduction aux technologies à l'échelle nanométrique …et à la théorie du petit BANG
etcGroup, Juin 2005
http://www.etcgroup.org/en/materials/publications.html?language=French

Little BANG – le schéma

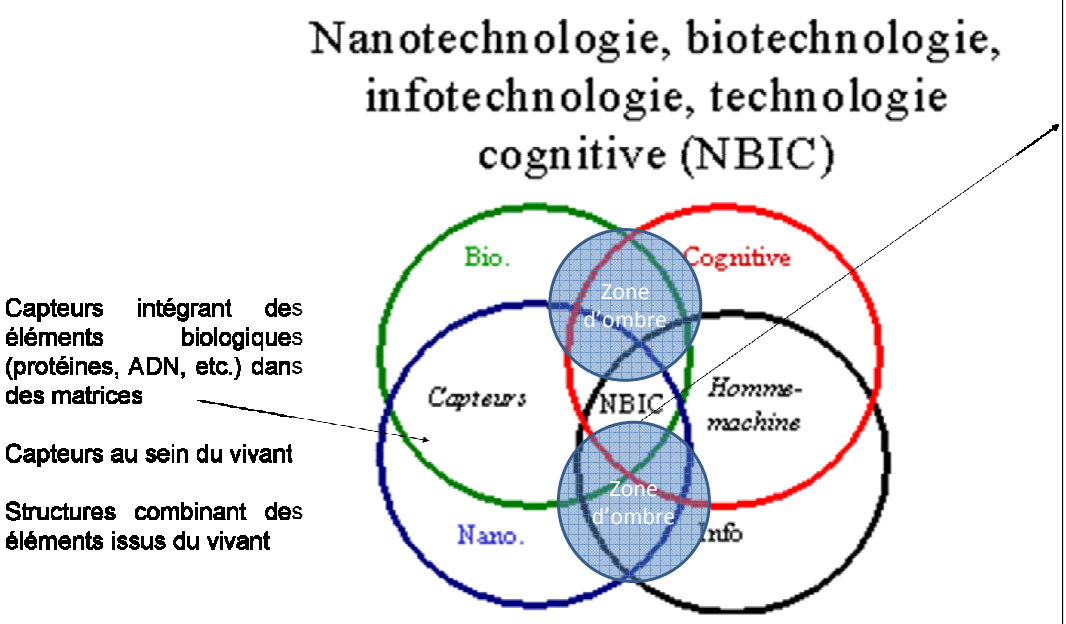
Un choc atomique qui fait BANG!
ETC Group qualifie la convergence de BANG.
Bits, Atomes, Neurones et Gènes s'additionnent pour former une théorie du petit BANG – la quête technologique du contrôle sur toute matière, toute vie, tout savoir.

La technologie	de l'information contrôle les	**Bits.**
La nanotechnologie	contrôle et manipule les	**Atomes.**
Les neurosciences cognitives	permettent de contrôler l'esprit en manipulant les	**Neurones.**
La biotechnologie	contrôle et manipule la vie en réorganisant les	**Gènes.**

On remarquera que la convergence porte sur les technologies et pas sur les connaissances qui les sous-tendent.

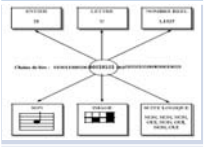
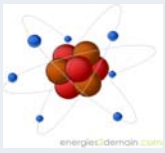
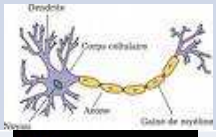
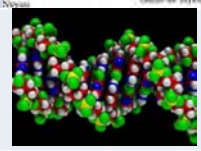
L'intrus sont les sciences cognitives et dont le niveau technologique (clic) ne permet pas le contrôle de l'esprit tel qu'affirmé. Nous allons voir aussi que le neurone ne se trouve pas au même niveau que les autres objets.

Little BANG – Interaction des niveaux



Ce schéma montre également des zones d'ombre qui montre que le concept NBIC/BANG peine à couvrir **réellement** tout le champ.

Little BANG – Mise à l'échelle

éléments	composés de	image	composent	
Bit	0 / 1 mot de 8 bits	0010010110101010 1001101100101001 010101100101010	Suite d'octets (alphabit)	
Atome	3 éléments (proton, neutron, électrons)		Molécules	
Neurone	Axone, dendrite, synapse **(complexe)**		**???**	
ADN (Gène)	4 bases (adénine, thymine , cytosine guanine)		Gènes	21'000 Chez l'homme

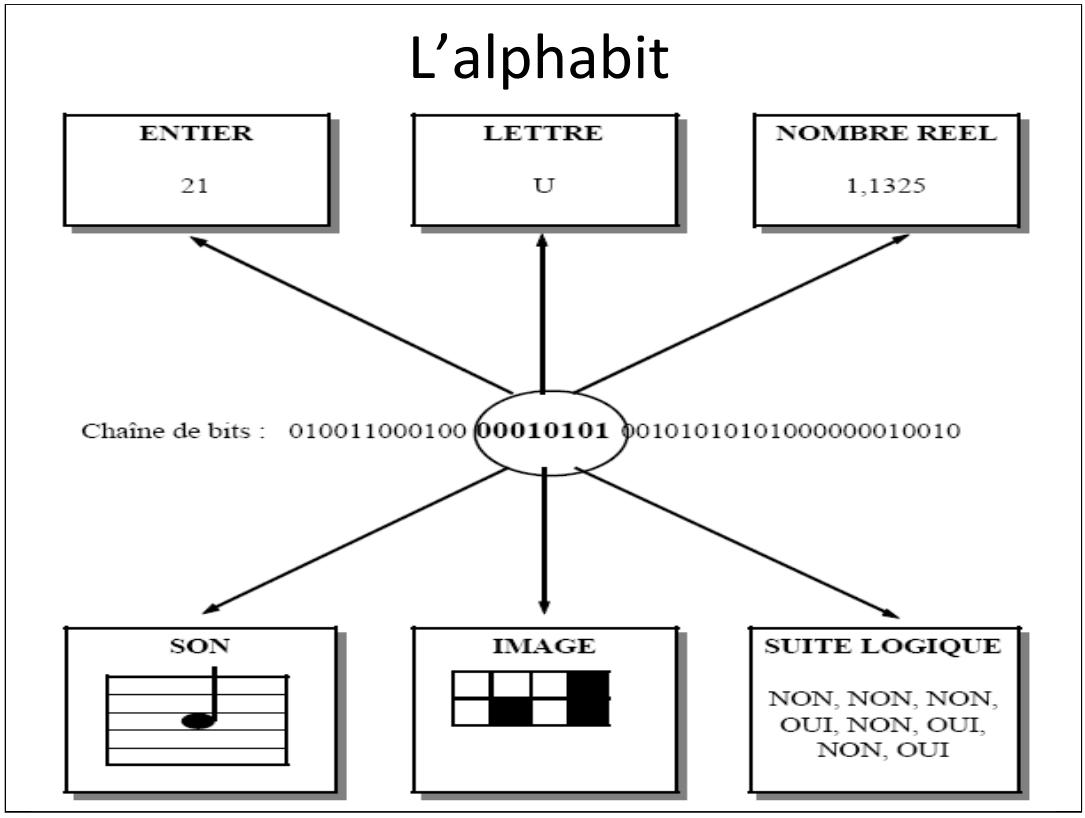
La mise en tableau de ces concepts montre :

Qu'à part le neurone les autres domaines reposent sur une combinatoire de composants peu nombreux.

Que la volonté de trouver un acronyme a fait intervertir les concepts d'un niveau : le gène se substituant à l'ADN en tant que composant de base du plan génétique.

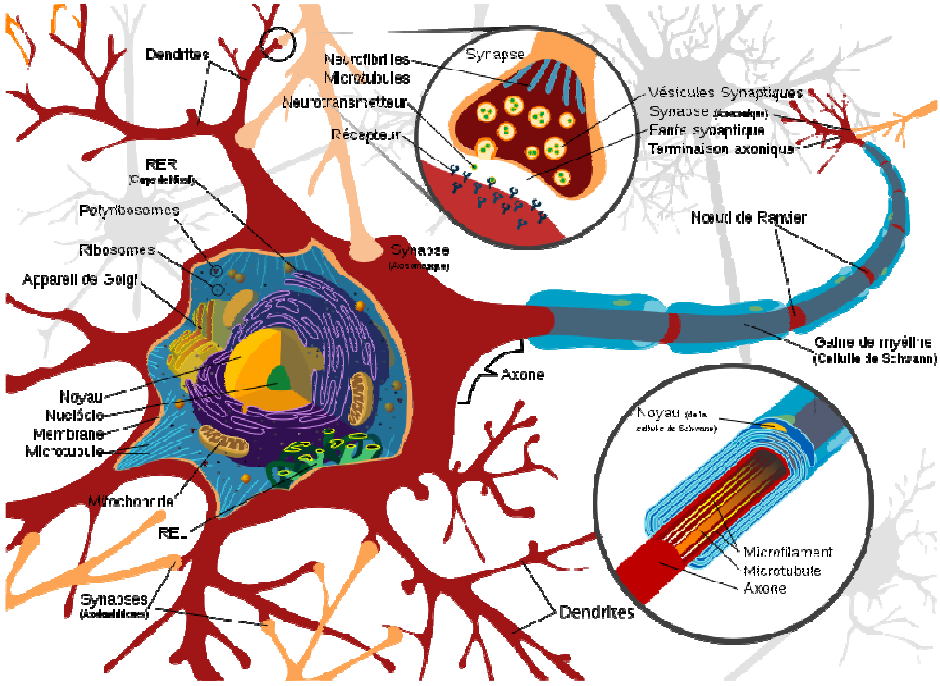
Que le neurone est l'intrus à cause de sa complexité par rapport aux éléments simples des autres niveaux et par le fait qu'il n'entre pas actuellement dans une logique combinatoire (???).

2 zooms :
- Sur « l'alphabit » : pour information (clic)
- Sur le neurone : qui n'est pas une suite d'éléments combinables mais un système complexe. (clic)

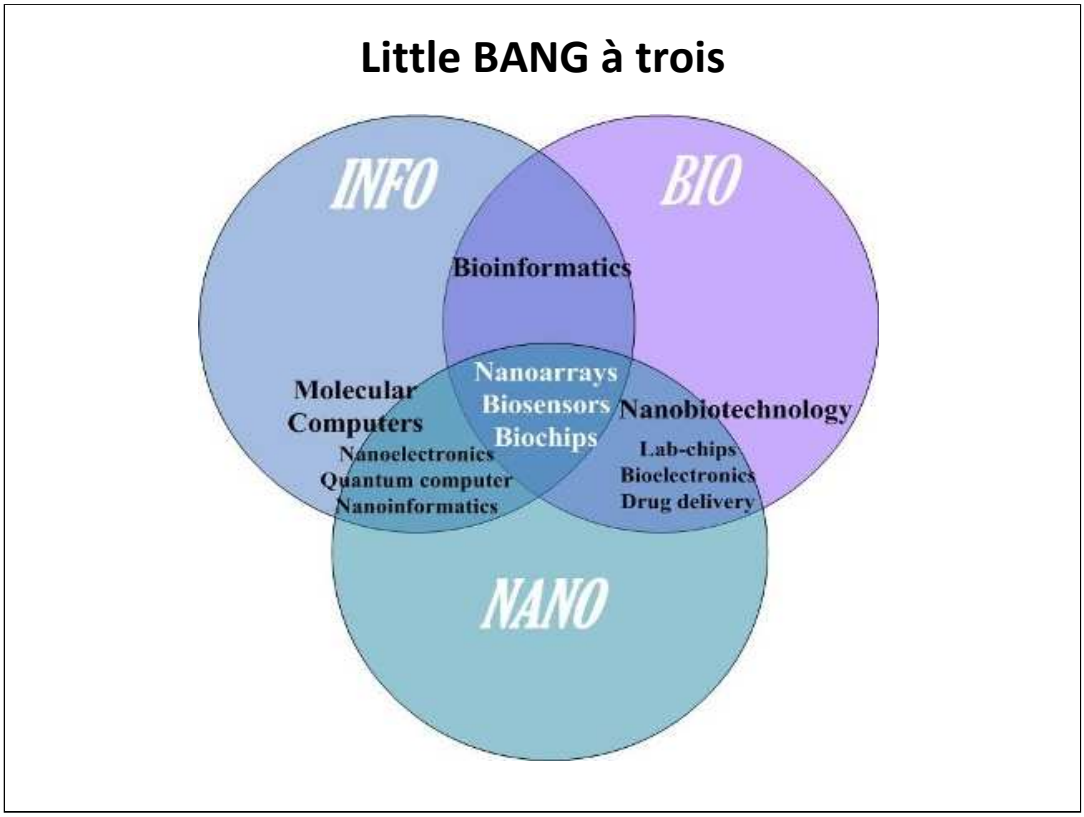


Une même chaîne de bits peut représenter des éléments différent selon l’environnement logiciel.

Le neurone –un système complexe



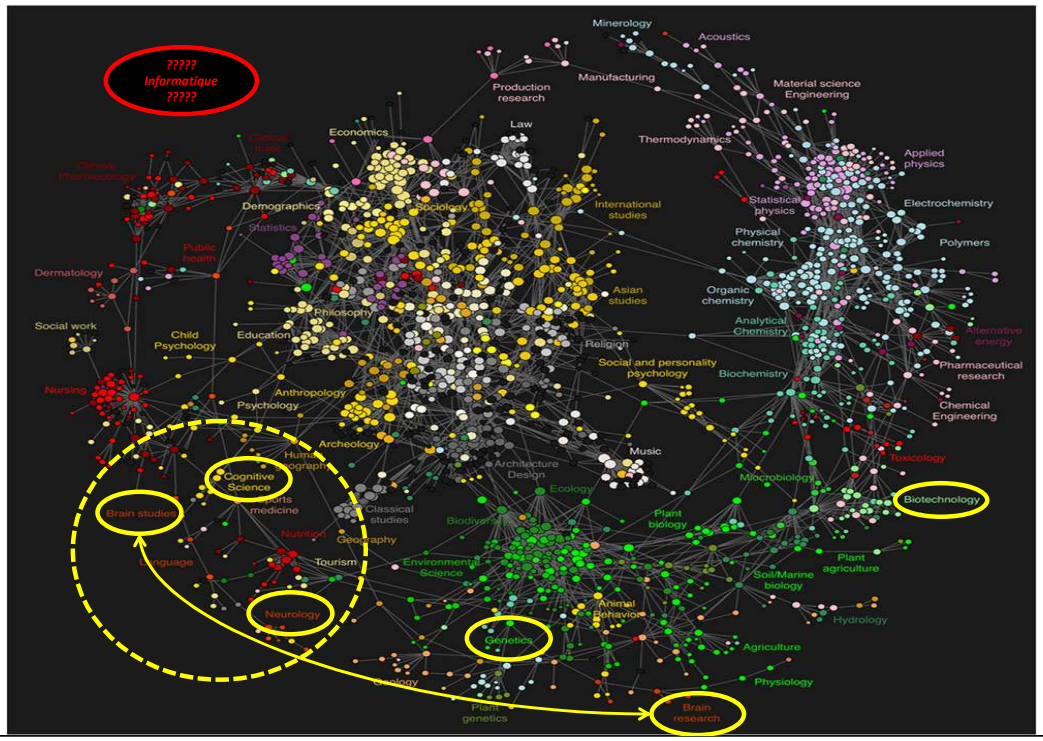
On peut difficilement assimiler ce système complexe à une combinatoire.



Voici un schéma issu de milieu technologique qui a exclu la dimension cognitive, ce qui semble confirmer l'inadéquation de cette liaison, qui n'est qu'un modèle projectif.

Cela semble confirmer l'artificialité du concept NBIC/BANG, ce qu'une étude transversale sur les publications scientifiques montre… (clic).

Map of Science



Cette carte illustre l'analyse d'un milliards de requêtes produites entre 2006 et 2008 par la communauté scientifique.

Les différents domaines du concept BANG apparaissent ici comme peu corrélés. (clic)

La seule proximité concerne des concepts intrinsèquement proches (brain studies, cognitive sciences, neurology). (clic)

Avec qui plus est l'apparition d'un doublon entre « brain research » et « brain studies ». (clic)

De plus, le domaine informatique semble noyé dans le reste des concepts abordés.

Il faut être prudent dans l'interprétation dans la mesure où cette carte ne représente pas un modèle des liens entre les concepts scientifiques dans l'absolu, mais l'image des liens que les chercheurs font actuellement entre les publications disponibles sur INTERNET.

Cependant, cet amas constellaire est ce qui se rapproche le plus de ce que devrait être le champ de connaissance tel que je le conçois…

Et dont voici l'état actuel …

Source :
Los Alamos Researchers Create 'Map of Science'
LOS ALAMOS, N.M., March 11, 2009 —
Data provides high-resolution picture of scientists' information retrieval habits
http://www.lanl.gov/news/index.php/fuseaction/home.story/story_id/15960

Une fiche descriptive d'objet de connaissance

Form Protégé

titre
résumé
référence (provisoire)
lieu
collation
collection
ISBN
Url standard
Url local
autres Url
consulté le
thème
sous-thème
type de ressource (publication)
type de contenu de la ressource
index
index ontologie
rédacteur notice
statut notice
publication /de la notice
notes

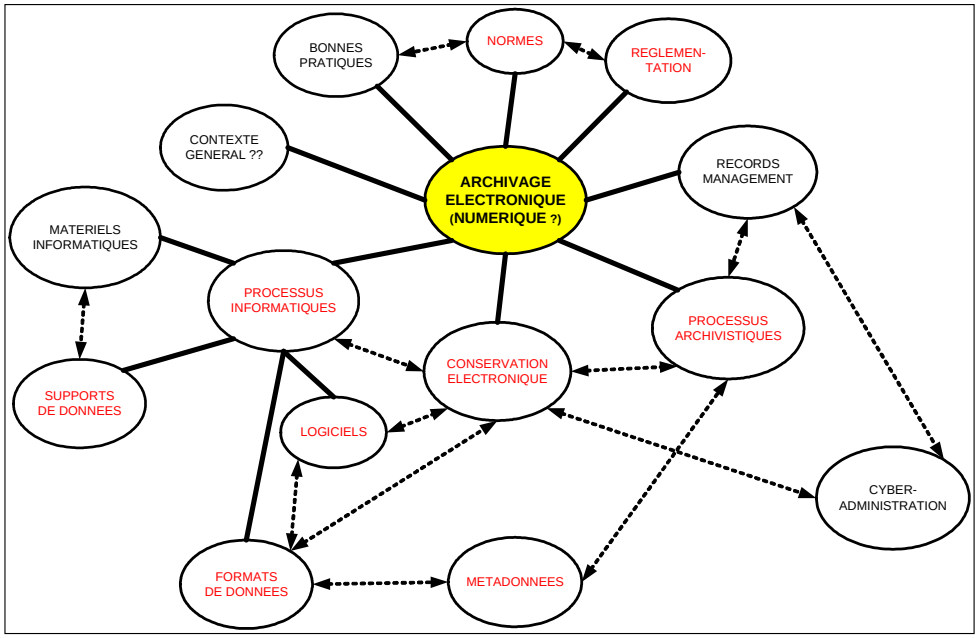
La première partie (jaune) ne diffère pas d'une notice catalographique standard.

La seconde partie permet de gérer des liens aux textes de référence stockés en dehors de la base.

La troisième partie (orange) est cruciale pour gérer des hiérarchies de liens conceptuels et qualifier la nature des documents/ressources. C'est le cœur de l'outils.

La drnière partie est un volet de gestion.

Un embryon d'ontologie



Dont il faut encore qualifier les types de liaisons.

Développé avec un programme de Stanford (Protege).

Voir démo dans le programme Protege.

Conclusion(s)…

Provocation à la discussion :

Le nœud du problème…

…c'est le(s) LIEN(S)