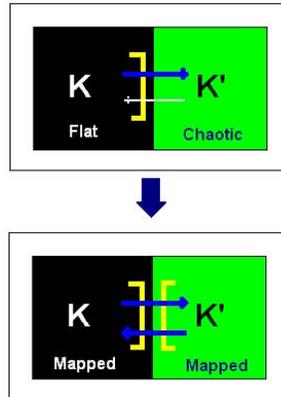


Towards a New Approach to Knowledge Management

A set of Conjectures

[Juan Chamero](#), CEO [Intag](#), December 3rd 2003

Note: updated as of March 2006



Actual versus new Knowledge Management paradigm is depicted in the figure above.

Established Human Knowledge K hosted in the Web space (black region) interacts with million of semantic pieces of **People's Human Knowledge K'** connected to the Cyber space (green region). Most actual interaction proceeds unidirectional broadcasting from K to K'.

K' appears as chaotic meanwhile K appears as structured even though "flat", in only one level. People are only enabled to query K. Data and Intelligence only flow unidirectional from K to K'.

Under a new paradigm enabled to use Internet at its full capacity win-win Human-Networking matchmaking scenarios could be attained optimizing the overall Human Knowledge cognitive offer and the Human learning process accordingly. Throughout "**e-membranes**" (yellow) Expert Systems aided by Agents may map K in as many levels as inherently exist. Once K is mapped its content is fully accessible, understood and retrievable from K' and in its turn K' becomes semantically coherent as seen from K side and Agents may proceed now to map it. Once both sides are structured and mutually understood and known, information, intelligence, and even wisdom may flow freely from K to K' and conversely from K' to K. The teaching-learning paradox alive in a sort of thermodynamic equilibrium is attained.

Abstract

Darwin, which stands for *Distributed Agents for Retrieval the Web Intelligence*, is a methodology to manage the HK, Human Knowledge. We distinguish two types of HK, the “**Established**” HK hosted in the **K Realm**, normally structured, and the “**People’s HK**”, somehow and somewhere hosted in the **K’ Realm** that looks like unstructured, chaotic, and dispersed. This methodology deals with these two realms trying to unveil, detect, and classify their main features. The core of the methodology is an Expert System aided by several families of Agents that work over these two realms throughout an e-membrane that resembling a bio membrane facilitates the open and free communication between them trying to create win-win scenarios without perturbing the actors. The Expert System works based on a set of common sense “Conjectures” that are now in process of being scientifically tested by the **AI-Lab of the CAECE University** from Argentina. We state that it is possible to automatically “retrieve” content and intelligence from data reservoirs. As an example, given the Web we may automatically generate its inherent “**Web Thesaurus**”. As long as people start to communicate openly and freely through these e-membranes their main behavior patterns could also be detected, unveiled, and classified. The semantic triad: Logical Tree, Thesaurus, and Cognitive Objects unequivocally identify any type of Knowledge.

Our finding could be a milestone that points to improve social communications: on one side we have juridical and physical persons selling, teaching, communicating, in some extent broadcasting their Truth, allegedly the “Established One”, and on the other side the People, buying, searching, attending, communicating, and continuously issuing opinions about everything conceivable, their believes and their personal truths.

We have created a first prototype, for a single discipline: Computing, that has an e-membrane where from agents have automatically retrieved 53,000 concepts and settled to learn as much as possible from users. This map, restricted to Computing, evolves by itself via agents. You may see it at www.intag.org.

Introduction

I will try to summarize the core concepts we were working on, related to IR, KM, Web Intelligence, Web Matchmaking, Expert Systems and Agents’ architecture and buildup. Everything started trying to optimize Web Matchmaking via Expert Systems aided by Agents. The more we advanced the more we realized that everything was closely related to the Human Knowledge understanding. The Web allows us the possibility to deal with a credible sample of it as a bulk and with its tiniest components, websites and pages at the same time. Our first work concerning the purpose of this document dealt with a huge American industrial and commercial database storing more than 10 million registers designed to be matched by millions of concurrent users. At that time and at that scale we realized of the existence of two realms, users and owners with an extremely poor matchmaking, more than expected based on their reasonable semantic differences. Our suspicion was stated in the following terms: people “talk” and even “think” different when acting either as owners or as users. If that’s true perhaps it was a trivial fact we didn’t aware of it yet. Not so trivial was to focus our attention to build a man-machine user interface between them capable to register any type of conceivable event at both sides of the interface. At that time we were also astonished by the pronounced apparent asymmetry of the matchmaking itself: the “owners” were by “de facto” not only enabled to broadcast anything they wanted but now with an Expert System empowered by a the “e-membrane” they could learn “as much as possible” from their users besides!.

We believed that in the long range this communication asymmetry will not be as good as an open and fair symmetry in the sake of the overall health of the man-machine coupled system. Asymmetries determine win-lose scenarios which at large become bad for both, owners and users. We intuited that we were in the right track to obtain a general solution to the matchmaking problem trying to envisage open, fair, balanced and symmetric systems instead. We suspected that a big part of the asymmetry and the poor matchmaking detected in many Internet services was due to the lack of “thematic” order in the Cognitive offer at the “machine” side, and because jargons and even ways of “thinking” at both sides of the man-machine interface were substantially different.

Conjectures Set

Conjecture 0: the semantic triad Logical Tree, Thesaurus, and Cognitive Objects unequivocally identifies any type of Knowledge

It is unchanged in its essence. However it implies a new Thesaurus definition encompassing the whole triad. In the beginning we accepted the notion of a “flat” Thesaurus as a single set of concepts for a given discipline. Then we upgrade it defining a structured Thesaurus with a set of keywords for each subject of a given discipline. Ultimately we may define a Knowledge Thesaurus tree like structure that in each node has: a) a rigorously defined subject; b) an associated well defined keywords profile set, semantic skeleton or “fingerprint”; and c) an associated BVL, Basic Virtual Library of meaningful documents (“Authorities”), each one with their specific associated “fingerprint” . This type of Thesaurus behaves like a “Knowledge Tree”. In figures let’s talk about Medicine. A typical curriculum would have about 1.500 subjects, 100,000 keywords, and from 30 to 70 specific (non-shared) keywords per subject in the average. By “specificity” we mean concepts that are closely related to a given subject and that probably were born within its realm. A typical Medical BVL would have from 10,000 to 20,000 Authorities. This triad identifies Medicine, behaving like the semantic skeleton of a reservoir of at least 30 million documents dealing with Medicine.

Note 1: in all these conjectures the crucial point to take care is the use of term “keyword”, a representation of a meaningful concept in a given discipline, by symbols, acronyms, words and precise strings of words. For instance NIC, Bell, IBM, parallel processing, gray codes, cardiovascular disease, zen fitness, e-business, n-tier, are keyword examples. Humans used them to document ideas combining and threading sequentially “common words” and keywords following very specific literary, writing styles, protocols, and syntax rules in documents dealing with specific “subjects”. Each subject is a keyword but the contrary is not always true, most keywords are not subjects or somehow are in process of becoming subjects. In our approach subject is a category of understanding and wisdom perhaps a semantic hierarchy, namely: data, information, knowledge, understanding, wisdom. As we will see in next Conjectures subjects could be considered crucial themes societies need to master to survive. The society established order must provide them, in fact knowledge structured via subjects. At their turn the subject semantic skeleton is its “fingerprint”, a weighted set of specific (non shared) keywords (see “wavelets”).

Note 2: Most conventional Search Engines index by words instead of keywords. Some of them only offer, from time to time, lists of most frequent users’ words/sequence of words, as highly probable keywords. They don’t differentiate them neither by subject nor by discipline and they don’t “learn of them” to improve the documents indexing and in some extent the documents interpretation. For this reason we say that their contents are “flat”, unstructured.

Conjecture 1: Website Owners “speak” and “think” rationale in terms of their objectives and in terms of their matchmaking policies.

It means that they send programmed “messages” (broadcasted or at demand) to their users. Its thinking is usually represented by “Menus” of features, attractions, associated with a sort of most times hidden “Logical Tree”. They think as hunters preparing their hunting strategies. Even in the case of educational sites it happens pretty much the same: teachers have to succeed “broadcasting” and at large inoculating pre established knowledge in their students. The same happens concerning governors governed matchmaking. From a simplified semantic point of view the owner’s behavior is structured and thought as something fixed: the “Established Truth”. This structure could be represented by a Logical Tree. These logical structures are well suited to describe the “offer”, at large the Cognitive Offer. If we were talking of Catalogs, associated Logical Trees describe them thoroughly, the main lines of products, their subdivision by type, material, functionality, from “root” to “leaves”. When talking of Knowledge hosted at the Established Side we deal with logical trees of “Major Subjects or Disciplines” as roots going down to sub subjects until we arrive to very specific leaves, namely specialties, niches.

Note 3: We conjecture that the human being as a user, in the role of asking for something, is not well suited to ask for subjects belonging to the K Realm. Our strong conjecture concerning ways of speaking and thinking is based on the intuition that People’s Subjects are substantially different from Established ones. In fact this Conjecture proposes the detection and unveiling of the People’s Thesaurus instead.

Note 4: We say rationale because we are convinced that “persons” behind the “Established Side”, think along a tree and speak like automata answering users’ queries like “gaming” governors versus governed. As a matter of fact things in K Realm are somehow “fossils”: laws, regulations, codes, prescriptions, are the consensual truth at a given moment inherited from the past. Their destiny should be “the change” to adapt better to changes in K’. In K things are frozen at a given time, for instance “as of ...”. That’s not bad!. It’s the eternal game of evolution. In K’ Realm, on the contrary life is continuously brewing: new ideas, suggestions, opinions, all came from K’.

Conjecture 2: People as users “speak” and “think” rather chaotically in terms of their passions, desires, their necessities at large.

As a corollary, owners and users speak and even think different. A smart observer, looking from a virtual “e-membrane” at the owners’ side, with analytical spirit, may arrive to the conclusion that any offer has a well defined purpose and that its “hunting” strategy could be precisely inferred. On the contrary, it would be practically impossible for any observer located in the same place to infer any type of order at the users’ side without accumulating extensive man-machine interaction experience and being smart enough to use advanced computing techniques and bizarre Human Behavior theories. Users try to satisfy their curiosity, and ultimately their needs, gaming against the “other side” at their own risk with their talent, their knowledge, and with strategies changing from time to time and from situation to situation. Users speak with personal “keywords”, a word or string of words –interpreted by K- that point to satisfy their needs. These user keywords (or their components) must exist in the owners’ side in order to get meaningful answers. For that reason the whole collection of possible user keywords should be thought as created at the owners’ side and organized in a People’s Thesaurus, associated throughout a sort of “non inheritance” rule to a users’ Logical Tree. By non-inheritance rule we mean “keywords specificity”, namely the right use of them at the right level of comprehension within human expressions of ideas. Let’s imagine these expressions as documents either too general or too focused. If too general authors would use general level keywords avoiding, as much as possible, the use of too focused keywords. Similar but reflexive could be the rule concerning too focused documents: authors would use focused keywords instead avoiding the use of too general keywords.

Note 5: Users need “solutions”, prescriptions that supposedly are stored and publicly offered in the K Realm, in most cases openly and free. To get them the only tool they have at hand are symbols, buttons, links, and at large string of meaningful words of K Realm side. So they are somehow obliged to know precisely the K jargon to succeed.

Conjecture 3: Users’ interactions along sessions are strings of semantic molecules of two types, users’ keywords, and navigation instances. These sessions’ strings are the representation of the users’ strategies to satisfy their needs.

With these two conjectures in mind it is possible to fully satisfy basic users’ needs as “predicted” at the owners K side. Well written but not existent keywords will cause a mismatch condition, and it’s a warning for a potential evolutionary change to be taken at K side.

This split suggests two separate basic users’ patterns: keywords’ patterns, and navigation’s patterns. Instances before and after a given user query, slightly modify both the meaning and the result obtained via a given keyword. Concerning searching strategies the experience tells us that in long sessions keywords may or may not belonging to the same discipline- doesn’t matter too much. The same applies to strings like [k1, k2, k3]: all its permutations could be considered similar strategies. A strong supposition is that the interleaving of instances within a session doesn’t affect the strategies outcomes significantly.

For all these reasons users sessions’ tracks of the form [iikikkikiikiii....], where k stands for keyword and i for instance, even ignoring their related outcomes, provide us a primal source to infer users’ behaviors without interfering with them!. Large numbers criteria come in our help: frequent sequence patterns could be associated to succeeding querying strategies that at their turn could be checked by blind type tests. The sequences of k n-ads: monads, dyads, triads, and so forth, detected in users’ sessions, could be considered like Tarzan’s expressions of the users’ Jargon.

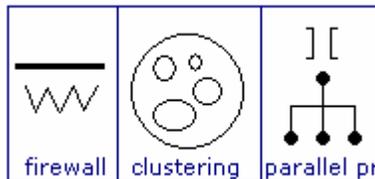
Note 6: All these considerations will strongly depend of what side you are in. A given word or string of words could be considered as a guess, the right “key” to open the Pandora box of K. If one user succeeds he/she may infer then that the key used is a keyword. From K side, you as an observer, may infer that a given word or sequence of words is a keyword if and only if it has been issued a significant number of trials by a significant number of users. Here we may envisage then two types of keywords, k and k’ which stand for keywords of K Realm and keywords of K’ Realm respectively. As we will see we may also define two Thesauruses, the K Thesaurus (Web Thesaurus) and the “people” K’ Thesaurus.

Conjecture 4: Cognitive Objects, documents, are expressed as strings of two semantic molecules, Common Words, belonging to a given Jargon, and keywords.

These cognitive objects are closely related to subjects of the Logical Tree, and this relation could be either explicitly or implicitly stated. We are especially interested in documents that behave as “Authorities” for each subject of the Logical Tree. An ideal authority should deal with only one subject of a given discipline and its keywords should be as specific as possible, it should use keywords that probabilistically tend to appear at a given level of the tree and strongly concentrated over only one node of this level. It means from minimal to zero presence of upper and lower level keywords. From this specificity emerges the concept of node/subject “neighborhood”. This neighborhood is defined in the three-dimensional semantic space Subjects – Keywords – Documents, generally as “related” documents existent around a given node with strong fingerprint similarity to a set of selected documents.

Note 7: Concepts resemble Chinese and Japanese ideograms that generally represent processes. Semantically they resemble and even in some extent behave like “wavelets”: sometimes like well defined semantic particles, and sometimes with diffuse wave like semantic contours. They are “approximately represented” by chains of common words/and or neologisms/acronyms. Perhaps in a near future it will be better represented graphically by symbols like the ones depicted below.

| | | |
|------------|-----------|----------------|
| 人 Human | 男人 Man | 女人 Woman |
| 家 God | 神 Home | 活化 Activate |



Note 8: These Conjectures could apply to the newspaper industry. Newspaper marketing people used to talk of the very “cognitive” content as “blank”, only attractors viewed as a physical media to carry ads and businesses. From this analogy we may then imagine two visions: for regular users newspapers is a media that reduces our ignorance, our uncertainty, and ads, all kind of propaganda are frequently considered blank, and only eventually cognitive objects, just the contrary that for owners!.

Conjecture 5: It is possible to enable a Full Duplex Type communication between Websites and their users throughout an e-membrane, enabling the free flow of content and its associated intelligence between them.

In Darwin architecture all the interactions are performed through an e-membrane. This membrane resembles a living semi permeable bio membrane where all the messages going forth and back through it are processed trying to extract as much information as possible from them and at the same time providing the best service to both sides. Let's see what it means.

When the message is a potential keyword coming from users' side, the membrane performs all possible "static" statistics at both sides, namely: it accounts for Web Thesaurus keyword use at owner side, makes a request to offer all documents available, and accounts for users' history. Besides these statistics keywords are analyzed and matched versus a set of suspected Users' Thesauri. Each keyword within a session is potentially considered as part of a "speech" or part of a sort of interception game (a "Get the Truth" game), and accordingly it could be accounted in more than one Users' Thesauruses.

This e-membrane was thought as a smart, subtle, and polite mechanism. Its aim is to obtain as much information as possible from the users' side without interfering with them, or at least minimizing interferences. The core of its "smart" built in strategy is to keep it from spying users. No cookies, no brute "bail and catch" hunting devices. Total absence of owners' side messages such as: Come here again!, You are Welcome, and even the mild May I help you?. To offer an open and free interface the e-membrane is built to stop and/or discourage any type of intrusion such as: How was it?, Was it helpful to you?, and never issue of "offenses" like "Please tell us a little about you". The dominant policy is "let it be", "let it flow freely", concerning communication between owners and users.

Note 9: We may envisage hitherto what symmetry of K and K' realms imply. In K Realm meaning begins with the knowledge of the Web Thesaurus and the comprehension of pairs [k, s]. In K' Realm meaning will begin when we get the building of People's Thesaurus and the comprehension of pairs [k', s'] instead.

Conjecture 6: Intrusions in communications cause serious troubles that go deeper and far than a local perturbation. The slightest intrusion may make invalid not only sessions but prevent users from communicating freely. They distort the static statistics and users' strategies as well.

Nota 10: This is trivial. More than a conjecture it could be considered something universally desirable, something that has gained global consensus. However these expected effects should be investigated and measured because there is a wrong belief nurtured by surveys and polls. Conventional surveys and polls are based on intrusions, sometimes generating visible harassment. These intrusions condition the answers. How much?. It has to be measured in order to eradicate it as much as possible as a credible methodology to know the K' Realm.

Conjecture 7 Human Knowledge, at a given time, is bounded.

We were talking about the practical use of the Digitalized HK in physical terms. One approach to have a metric of it could be the volume of documents necessary to describe it. If we estimate the total amount of written documents in 12,000 millions –each one contributing with a piece of "valuable" information- we may ask ourselves then: How many of all those documents could be considered "Authorities"? We are also talking about the "Established" Human Knowledge; let's say the sum of all the structured knowledge as "Disciplines", the Major Subjects of the Established Human Knowledge. In this sense a Major Subject could be Medicine, Engineering, Philosophy, and the basic Curricula we learn in the universities (we have to add more subjects to the classic list, like for instance all derived from entertainment issues, the "homo ludens" activities).

Along the last decade there were some approaches to measure this Human Knowledge variety. The Britannica for example tells us about 8 super disciplines in the upper level: history, philosophy, art, society, technology, religion, science, and mathematics. Some of them open too much like technology, and science meanwhile some as religion, and mathematics open in a few branches.

We have found serious essays that talk of HK spectrum ranging from 150 to 200 Major Subjects. In its turn Major Subjects could be opened in Logical Trees from 600 to 1,000 subjects each. Well let's talk now a little about Authorities.

We mean by authority a document dealing with approved/certified authoritativeness about a given subject. If the subject is up in the hierarchy the document could be physically organized as a Manual, or as a Treaty. Going downwards in the tree we may talk of Essays, and Papers.

As you know this is not strict but a trend. Anyway we are talking of whole documents not single pages. This comment is addressed to the Web, where the unit of retrieval via search engines is the page. Authorities are generally hosted in Web sites not in single pages.

We estimate 3 authorities in the average per subject, in order to have a BVL Basic Virtual Library for a given discipline. Talking in Web terms we mean 3 Websites per subject. You may imagine this sizing exercise how a university professor average library would look like for comparison. For a discipline of 1,000 subjects that would mean 3,000 non repeated "e-books" to cover the whole spectrum. Not bad!. However this is not enough to be fully satisfied. We need more than that; we need redundancy, complementary and similar documents. Let's see how we could cope with this necessity.

The next step is to have one idea of the size of the Human Knowledge Thesaurus, the whole collection of the whole HK (from the "Established" side point of view). Let's perform some calculations to estimate it:

Major Subjects: 200

Subjects per Major Subject: 1,000

Authorities per Subject: 3

BVL of the HK: 600,000 Authorities

Extended HK: 200 million documents (similar, found by Darwin HMM, Hidden Markovian Model agent)

Note 11: We estimate that the average single discipline Thesaurus will have about 50,000 keywords, so we are talking of a Web Thesaurus: of nearly 10,000,000 keywords.

Note 12: Now going back to our need of more information to enrich our BVL. We have two options. First one is to retrieve similar authorities to the one we have at hand. Second strategy is to look for the best documents that satisfy our initial curiosity. In both cases our starting point is a basic document "a priori" considered good. How many documents could we retrieve with almost the same quality (authoritativeness)?. The answer is: as much as we decide depending of the existence of them and the quality resolution of the retrieval procedure. Our optimistic estimation is 20 per query, so extending the limit of our Extended HK to 200 millions of documents. Not too bad again!. This assertion is especially worthwhile if we are talking of a current **HKM, Human Knowledge Map** continuously maintained, updated with the news of it, and corrected. That means an evolutionary HKM.

Note 13: In Darwin methodology Thesauruses evolve along the time, new concepts arise, many concepts age too much, many subjects obsolesce and even die, their interrelation changes continuously, weak branches become new roots, important branches disappear, many merge among them defining new subjects. In the K' Realm things happen similarly: activities, professions, artisans mastering, evolves along the time. However at a given time activities and know how are bound even though growing in diversity and sophistication but not so fast as we may believe. The antique Alexandria library kept more than 300,000 Authorities, nothing negligible!. The AI Lab of the CAECE University have accounted for nearly 6,000 different human economical, social and political activities, a big amount but bounden and easily manageable.

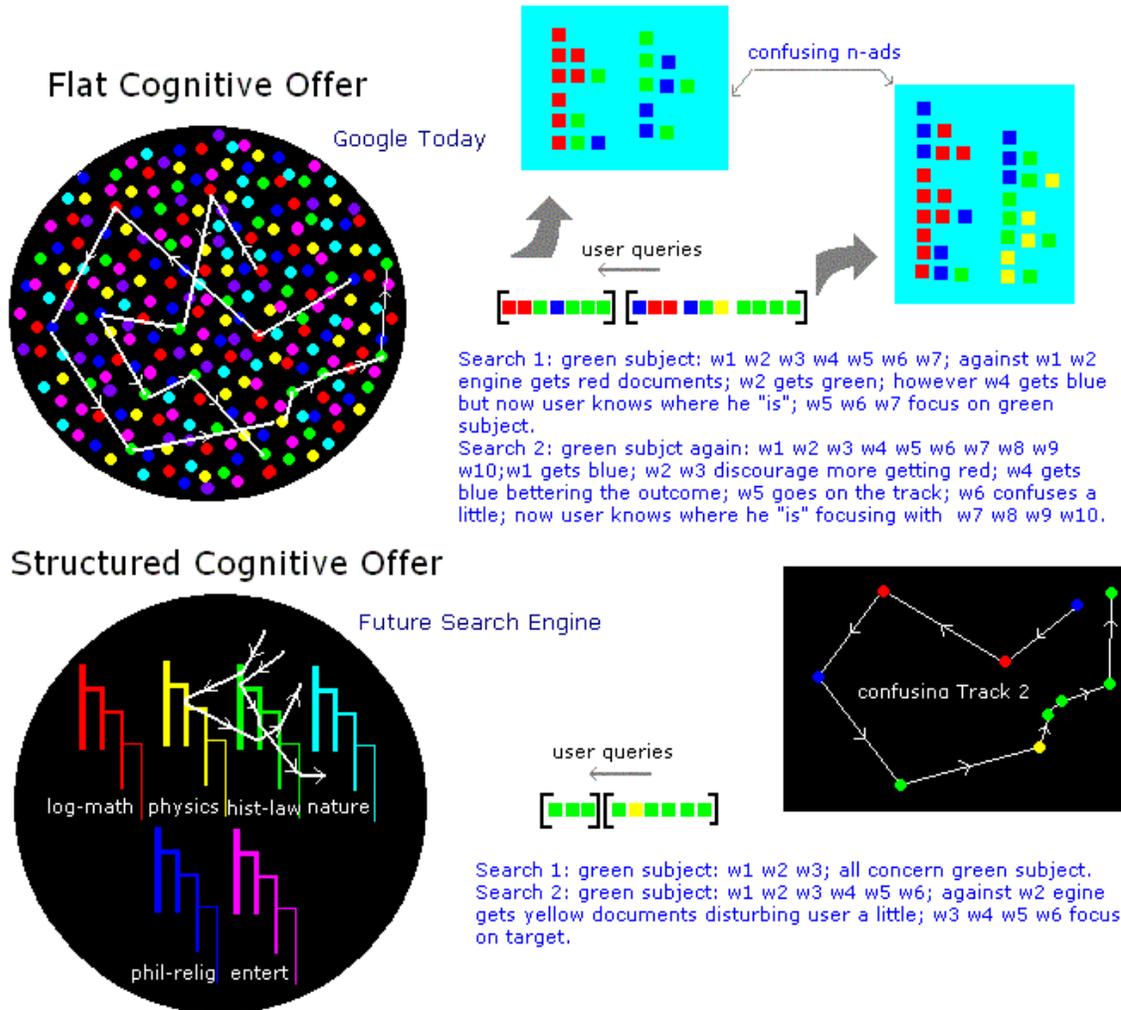
Conjecture 8: Given a LT, Logical Tree of a given discipline we may automatically generate its related Thesaurus

This conjecture was confirmed. Our first Thesaurus of 53,000 keywords for the ICT, Information, Computing, and Telecommunications discipline was initially determined partly by humans and partly by agents. With this experience we have designed and implemented a procedure to perform the automatic generation of Thesaurus given a LT provided we have access to a reservoir (the Web in this case) with "sufficient" data where from to retrieve the semantic components for a given discipline.

Be that reservoir the Web and the LT of Medicine. We need a huge amount of suspected "Authorities", for instance more than 100,000, to apply Darwin. Once the Thesaurus is generated we may proceed then to determine the BVL and the EL, Extended Library of Medicine.

Be the reservoir a proprietary database and the task the engineering/reengineering of a Catalog: The same considerations apply!.

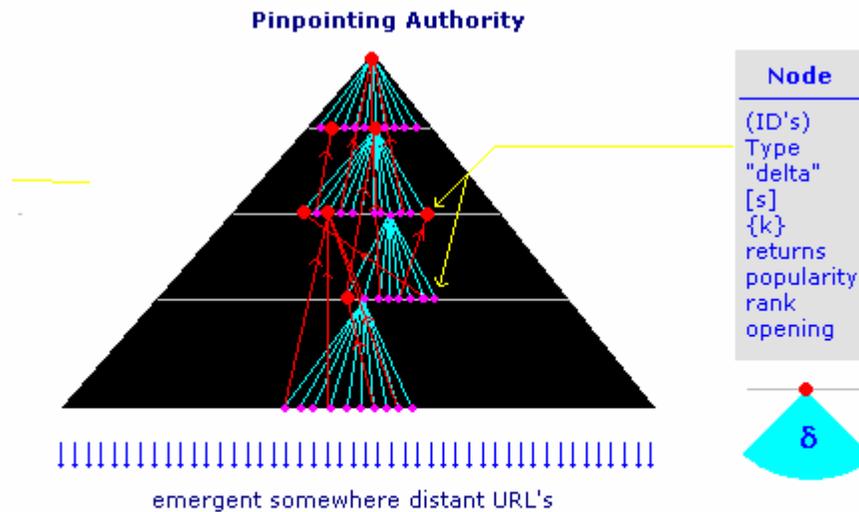
Note 14: This Conjecture presupposes the existence of a huge reservoir of documents dealing with subjects belonging to the knowledge we want to map. However this conjecture is hard to generalize. In fact Darwin is one procedure to perform that in a series of steps as follows: From LT agents extract a flat and almost "naked" Thesaurus seed; with this seed new agents extract a first KB, Knowledge Base; from this knowledge base the first agents extract now a whole flat Thesaurus (not a seed but still unstructured); with this flat Thesaurus we extract a better KB Knowledge Base; with this knowledge base guided by LT a complex clustering algorithm working statistically with Sparse High Order Binary Matrices a new type of agent structure the whole thesaurus, let's say a feasible set of semantic pairs [k, s].



Note 15: The figure above depicts the difference when searching using a conventional Search Engine respect to one with its content mapped in as many levels as its inherent intelligence permits. Users Instead of issuing fuzzy and anodyne queries will issue the best path guided by a smart Wizard, making possible the GWYN, Get What You Need in only one click!.

Conjecture 9: Given a Historical Reservoir we may generate its related Thesaurus.

We are now pursuing this conjecture to become a reality. The crucial problem to solve is the automatic LT generation. We may enhance the cycle by human intervention as a first approach and agents offering a set of weighted LT alternatives. We are now trying to go farther: Recovering the "Hidden Intelligence" of a collection of documents, "a priori" considered "sufficient" to represent a given discipline. No curricula at hand, only documents, only a set of suspected "Authorities". If the collection is sufficient we may argue that from the "Union of Authorities" we may retrieve a first approach of the hidden LT but flat, unstructured, all possible subjects in the same ground level. We intuit that this way being feasible and precise is better than beginning with a pre-established Curriculum. Knowledge pieces interrelate among them like graphs not as ideal trees. Because this enforced beginning our mapping based on curricula has by "de facto" an original vice. It would be better trying to unveil the best graph that fits better the existing KB.



- * PU-Pyramidal Unit: from 10,000 to 200,000 node-members (N)
- * N suspected {k} sets and [s]'s
- * PUT-Pyramidal Unit Thesaurus: U/{k}
- * PUB-PU Boolean Vector size (PUT), pointing to existence and abundance of k's in PU {k}'s

Note 16: The picture above depicts how an agent analyzes a "Pinpointing Authority", namely a Web Portal dealing with the discipline to unveil. These analyses are multiple correlated among them in order to detect potential monads [s], [k] to become feasible pairs [s k].

Bibliography

1. [ICONS](#) , [Intelligent Content Management System](#) , Feb 2003, Rodan Systems (PL), The Polish Academy of Sciences (PL), Centro di Ingegneria Economica e Sociale (IT), InfoVide (PL), SchlumbergerSema (BE), University Paris 9 Dauphine (FR), University of Ulster (UK); It deals with Access Algorithms and Data Structures Underlying a Distributed Knowledge Base.
2. Web Epistemology, <http://webepistemology.org/main> , is a platform for studying the impact of the web on the mechanisms of collective knowledge production, organization and distribution. A major part of this project is devoted to analysing the stakes and goals of scientific research management and policies with regard to the web.
3. [Distributed Knowledge Modeling through the World Wide Web](#) , by Mildred L. G. Shaw and Brian R. Gaines, from Knowledge Science Institute, University of Calgary, Alberta, Canada T2N 1N4, {mildred, [gaines](mailto:gaines}@cpsc.ucalgary.ca)}@cpsc.ucalgary.ca ,
4. Castells, M. (2001), "The Internet Galaxy – Reflections on the Internet, Business, and Society", London; Oxford University Press.
5. [Web Hunting: Design of a Simple Intelligent Web Search Agent](#) , by [G. Michael Youngblood](#) .
6. [KWIC](#) , Keyword in Context, an old but always present Concordance concept and tool, as a parsing general utility.
7. Knowledge Management – [Digital Cloning \(KES\)](#) , Digital Informational Experts Created from Prior Personal Knowledge, where KES stands for Knowledge Expert Systems. It's a reflection about knowledge, personal and collective, and what information and knowledge management are.
8. [Actualizing Context for Personal Knowledge Management](#) , by Kenneth A. Berman and Fred S. Annexstein, Department of ECECS, University of Cincinnati, Cincinnati, OH 45221. ken.berman@uc.edu , fred.annexstein@uc.edu .
9. [KAON](#) is an open-source ontology management infrastructure targeted for business applications. It includes a comprehensive tool suite allowing easy ontology creation and management and provides a framework for building ontology-based applications. An important focus of KAON is scalable and efficient reasoning with ontology.
10. [Classification and Clustering](#) , extracted from Kiduk Yang's Doctoral Student Thesis, [North Carolina University](#).
11. "[El Espacio Web y la Noosfera](#)", (The Web Space and the Noosphere), by [Dr. Juan Chamero](#) and published by SCIPPOOL–Red Científica, a bilingual Spanish-English scientific community, 2001.
12. "Towards a New Digitalized Human Knowledge Paradigm", published by SCIPPOOL –Red Científica, a bilingual Spanish-English scientific community. You may see it in [English](#) at and in [Spanish](#)
13. "Towards a New Knowledge Management Paradigm", ISSN 109-2750, presented at WSEAS, the World Scientific Engineering Academy and Society Multi Conference held in Miami, USA, April 2004, WSEAS Transactions On Computers, Issue 5, Volume 3, Page 1488. It's available for members at www.wseas.org and as an abstract in www.wseas.com .
14. "[How Case Studies Methodology embeds with continuity within the millennial Teaching Learning Paradigm](#)" , published by SCIPPOOL –Red Científica, August 2004, a bilingual Spanish-English scientific community.
15. Presented as a New Generation Search Engines Tutorial in the Argentine Expocomm 2004 by the IEEE Buenos Aires Chapter, held in Buenos Aires, Argentina, September 2004, You may obtain a copy of it by asking for <http://expocomm.com/argentina/> or to the lecturer juan.chamero@intag.org .