**Chapter 1 -- Idea Analysis – A Research Tool**

**Introduction**

Given the exponentially increasing amount of information available, the individual-- student or professional -- is faced with the challenge of identifying, extracting, organizing, and utilizing facts and opinions to form new descriptions of a topic. Research plays an important role in dealing with the deluge of information. In a real sense, research is a formalized learning process, pushing back the unknown and discovering aspects of new topics.

Many books have been written describing research methods. Few however have focused on the full range of tasks and how they can be performed using the most effective capabilities. Those tasks are:

1. ***Developing a learning resource***. This repository of essential data provides the ingredients needed to construct descriptions of the topics and to develop new strategies leading to advancing knowledge. This construction process can best be accomplished by software designed to identify, extract, and organize specific data from the scientific publications describing the topic. By using software, there is enhanced transparency of procedure. The measures, criteria, and decision-rules must be specified in advance and use of those performed in a consistent fashion.
2. ***Using the learning repository.*** The data in the learning resource represent the totality of information regarding the topic. The student can rapidly learn the basic description of the topic by focusing on the information considered most relevant by the subject specialists who developed it. This specialist-guided mentoring and instruction expands on the work of an author of a textbook by providing a worldwide view of the topic rather than an individual’s interpretation.

The professional can rapidly access the subsets of data considered relevant in addressing a specific topic. In addition to saving time and effort, the subsets selected will be more complete than an individual’s personal perspective of the knowledge, thus providing more opportunities to expand the body of relevant information.

1. ***Formalized learning.*** With the learning repository available, the individual’s time and energy can be shifted from ***acquiring*** data to ***using*** the higher cognitive functions. The first of these is synthesis, the construction of new descriptions of the topic or issues in the topic based on arrangements of the essential data from the learning repository. In contrast to previous methods, this version of synthesis would lead to many different possible arrangements. This array of new descriptions offers the opportunity to compare, evaluate, and judge each leading to a ranking denoting most applicable to least.

The selected application could be called a testable hypothesis or the question to be studied in a research program. It is the ***best*** arrangement of the existing facts and represents the most plausible portal to new understandings. Those results will be determined by the new study. In this sense, the process of developing a study hypothesis and the process involved in populating that hypothesis with new information are transparent, quality-controlled events. Learning accomplished using this approach can be described as a system of formalized performance of specific tasks, operationalized critical thinking, or simply as research.

**Separating the Processing Tasks:**  The approach described in this discussion, i.e., idea analysis, using computer support, allows separation of the individual tasks and assigns each to the most capable and effective methodology. The implication of this separation is the ability to consider the details and issues associated with each cognitive function. This emphasis leads to a more open process and could lead to a formalized version of critical thinking and its companion, creative thinking.

Critical and creative thinking are recognized desirable skills applied in various ways in higher education. However, a significant challenge is the translation of this behavior from an ideal to an operational entity.

**Example from Dog Cancer Research:** Table 1 shows the different cancers studied and the frequency of ideas involving each term for the period 1980-2013. These data are part of an ongoing construction of a database of dog-diseases. The complete set of ideas totaled 2,857,264 from 27,926 documents entered into PubMed during 1980-2013.

**Table 1. Excerpt of Cancer Ideas From Dog Disease Vocabulary.**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Term** | **1980-4** | **1985-9** | **1990-4** | **1995-9** | **2000-4** | **2005-9** | **2010-13** | **Sum** |
| bladder | 4 | 4 | 1 | 8 | 1 | 6 | 5 | 29 |
| breast | 8 | 6 | 18 | 7 | 13 | 51 | 33 | 136 |
| cervical | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 3 |
| colorectal | 0 | 0 | 0 | 3 | 1 | 2 | 2 | 8 |
| hodgkins | 0 | 1 | 2 | 0 | 0 | 1 | 0 | 4 |
| liver | 1 | 1 | 0 | 1 | 2 | 4 | 4 | 13 |
| lymphoma | 0 | 2 | 8 | 2 | 2 | 13 | 0 | 27 |
| mammary | 6 | 6 | 10 | 9 | 5 | 49 | 0 | 85 |
| melanoma | 0 | 1 | 3 | 3 | 1 | 3 | 0 | 11 |
| ovarian | 0 | 0 | 0 | 3 | 3 | 1 | 2 | 9 |
| pancreatic | 2 | 0 | 2 | 0 | 0 | 3 | 2 | 9 |
| prostate | 0 | 0 | 1 | 20 | 19 | 34 | 19 | 93 |
| prostatic | 0 | 0 | 1 | 3 | 3 | 6 | 0 | 13 |
| sarcoma | 2 | 0 | 7 | 7 | 4 | 14 | 2 | 36 |
| thyroid | 1 | 0 | 8 | 0 | 0 | 3 | 8 | 20 |

Each of the terms shown in column 1 was linked with the central term – cancer – to form an idea and label for the involved cancer. For example, there were 136 ideas involving the pair – cancer & breast – in the dog disease literature during the period 1980-2013.

The large number of ideas that must be processed to find the ones that are relevant is a condition of text processing. The only way to accurately describe each document is by identifying the authors’ ideas. To do that means that each document must be read, the ideas extracted and organized for future use. As a result, the dog cancer literature was found to contain 4782 ideas that described attributes of cancer. The use of software enabled this effective triaging in a rapid and efficient fashion.

Table 2 shows an excerpt of the ideas involving cancer. This document was entered into PubMed during 2013. The associated sentence is shown in Exhibit 1. The terms identified by the software are highlighted in red. The terms that could have been identified but were missed are identified in blue. The capture accuracy for this sentence was 71% (5/7). Overall, capture accuracy had a median of 85% (66% to 99%).

**Table 2. Excerpt of Ideas Dealing with Dog Cancer.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Cancer** | **Related** | **Year** | **Ident** | **Sentence** |
| cancer | angiogenesis | 2013 | 23561040 | 1 |
| cancer | cell | 2013 | 23561040 | 1 |
| cancer | invasion | 2013 | 23561040 | 1 |
| cancer | tumor | 2013 | 23561040 | 1 |
|  |  |  |  |  |

**Exhibit 1. Ideas and Sentences for Document 23561040 (2013).**

***Source:*** [***Król M***](http://www.ncbi.nlm.nih.gov/pubmed/?term=Kr%C3%B3l%20M%5BAuthor%5D&cauthor=true&cauthor_uid=23561040)***1,*** [***Majchrzak K***](http://www.ncbi.nlm.nih.gov/pubmed/?term=Majchrzak%20K%5BAuthor%5D&cauthor=true&cauthor_uid=23561040)***,*** [***Mucha J***](http://www.ncbi.nlm.nih.gov/pubmed/?term=Mucha%20J%5BAuthor%5D&cauthor=true&cauthor_uid=23561040)***,*** [***Homa A***](http://www.ncbi.nlm.nih.gov/pubmed/?term=Homa%20A%5BAuthor%5D&cauthor=true&cauthor_uid=23561040)***,*** [***Bulkowska M***](http://www.ncbi.nlm.nih.gov/pubmed/?term=Bulkowska%20M%5BAuthor%5D&cauthor=true&cauthor_uid=23561040)***,*** [***Jakubowska A***](http://www.ncbi.nlm.nih.gov/pubmed/?term=Jakubowska%20A%5BAuthor%5D&cauthor=true&cauthor_uid=23561040)***,*** [***Karwicka M***](http://www.ncbi.nlm.nih.gov/pubmed/?term=Karwicka%20M%5BAuthor%5D&cauthor=true&cauthor_uid=23561040)***,*** [***Pawłowski KM***](http://www.ncbi.nlm.nih.gov/pubmed/?term=Paw%C5%82owski%20KM%5BAuthor%5D&cauthor=true&cauthor_uid=23561040)***,*** [***Motyl T***](http://www.ncbi.nlm.nih.gov/pubmed/?term=Motyl%20T%5BAuthor%5D&cauthor=true&cauthor_uid=23561040)***. CSF-1R as an inhibitor of apoptosis and promoter of proliferation, migration and invasion of canine mammary cancer cells.*** [***BMC Vet Res.***](http://www.ncbi.nlm.nih.gov/pubmed/?term=23561040) ***2013 Apr 5;9:65. doi: 10.1186/1746-6148-9-65. PMID: 23561040***

**Sentence 1: *Tumor-associated macrophages (TAMs) have high impact on the cancer development because they can facilitate matrix invasion, angiogenesis, and tumor cell motility.***

**Selection of Vocabulary:** The selection of informative terms is based on the authors’ use of these. Four measures are considered. The first is the ***grammatical class*** of the term. Informative terms are members of the nouns, adjectives, or gerunds classes. The second is the ***frequency of use of the idea***. The ideas used most often tend to represent a consensus acceptance of relevance. The third is the ***consistency of use*** across different situations. That depicts another form of consensus wherein the term is combined with other relevant terms to describe a variety of situations. The fourth is the ***specificity of the terms*** in the ideas. The terms that are more specific to the topic tend to be involved in a larger number of ideas within the involved sentences. In the sentence shown in Exhibit 1, the term – ***motility*** – lacked specificity and was used infrequently in the document set. Similarly, the term – ***macrophage*** – while specific was used infrequently.

**Verification of Informative Terms and Ideas:** The selection of informative terms can be evaluated by comparing the terms selected by the software with those used by the subject specialist-authors. This is an important part of the quality-control process. By comparing terms captured by the two methods, the performance of the software can be determined. In the sentences that follow, terms are color coded to represent different situations. The terms highlighted in red indicate that the software properly selected them. Those in blue denote that the software missed them. Terms highlighted in blue were used by the authors, but, were not included in the vocabulary.

Exhibit 2 shows examples of the verification process using a 1984, 2004, and 2009 publication. The format is as follows:

1. The source data were given including the identification number assigned by PubMed (PMID). That number makes retrieval rapid and accurate.
2. The involved sentence was shown with the informative terms highlighted. Each term is identified once within a sentence.

Behind each sentence is a ratio giving the number of terms identified by the software relative to the number of terms available for identification.

# Exhibit 2. Sentences and Verification of Informative Term Capture by Software.

***Source:*** [***Couto CG***](http://www.ncbi.nlm.nih.gov/pubmed/?term=Couto%20CG%5BAuthor%5D&cauthor=true&cauthor_uid=6547126)***,*** [***Cullen J***](http://www.ncbi.nlm.nih.gov/pubmed/?term=Cullen%20J%5BAuthor%5D&cauthor=true&cauthor_uid=6547126)***,*** [***Pedroia V***](http://www.ncbi.nlm.nih.gov/pubmed/?term=Pedroia%20V%5BAuthor%5D&cauthor=true&cauthor_uid=6547126)***,*** [***Turrel JM***](http://www.ncbi.nlm.nih.gov/pubmed/?term=Turrel%20JM%5BAuthor%5D&cauthor=true&cauthor_uid=6547126)***. Central nervous system lymphosarcoma in the dog.*** [***J Am Vet Med Assoc.***](http://www.ncbi.nlm.nih.gov/pubmed/?term=6547126) ***1984 Apr 1;184(7):809-13. PMID 6547126***

***Sentence 3: Two dogs were given systemic anticancer chemotherapy, and 4 dogs were given a combination of systemic chemotherapy, intrathecal cytosine arabinoside, and craniospinal irradiation, resulting in marked improvement of the clinical signs. (8/8)***

**Source:** [**Gagnon J**](http://www.ncbi.nlm.nih.gov/pubmed/?term=Gagnon%20J%5BAuthor%5D&cauthor=true&cauthor_uid=15635895)**1,** [**Bouchard F**](http://www.ncbi.nlm.nih.gov/pubmed/?term=Bouchard%20F%5BAuthor%5D&cauthor=true&cauthor_uid=15635895)**,** [**Landry M**](http://www.ncbi.nlm.nih.gov/pubmed/?term=Landry%20M%5BAuthor%5D&cauthor=true&cauthor_uid=15635895)**,** [**Belles-Isles M**](http://www.ncbi.nlm.nih.gov/pubmed/?term=Belles-Isles%20M%5BAuthor%5D&cauthor=true&cauthor_uid=15635895)**,** [**Fortier M**](http://www.ncbi.nlm.nih.gov/pubmed/?term=Fortier%20M%5BAuthor%5D&cauthor=true&cauthor_uid=15635895)**,** [**Fillion L**](http://www.ncbi.nlm.nih.gov/pubmed/?term=Fillion%20L%5BAuthor%5D&cauthor=true&cauthor_uid=15635895)**. Implementing a hospital-based animal therapy program for children with cancer: a descriptive study.** [**Can Oncol Nurs J.**](http://www.ncbi.nlm.nih.gov/pubmed/?term=15635895) **2004 Fall;14(4):217-22. PMID 15635895**

***Sentence 11: Stemming from a nursing initiative started in 1999, this project aims to promote the well-being of children living with cancer during their hospitalization, reduce their emotional distress and facilitate their adaptation to the therapeutic (psychological, physical and social) process by promoting the emergence of special bonds between children and animals. (10/12)***

**Source:** [**Kim JH**](http://www.ncbi.nlm.nih.gov/pubmed/?term=Kim%20JH%5BAuthor%5D&cauthor=true&cauthor_uid=19959182)**1,** [**Yu CH**](http://www.ncbi.nlm.nih.gov/pubmed/?term=Yu%20CH%5BAuthor%5D&cauthor=true&cauthor_uid=19959182)**,** [**Yhee JY**](http://www.ncbi.nlm.nih.gov/pubmed/?term=Yhee%20JY%5BAuthor%5D&cauthor=true&cauthor_uid=19959182)**,** [**Im KS**](http://www.ncbi.nlm.nih.gov/pubmed/?term=Im%20KS%5BAuthor%5D&cauthor=true&cauthor_uid=19959182)**,** [**Sur JH**](http://www.ncbi.nlm.nih.gov/pubmed/?term=Sur%20JH%5BAuthor%5D&cauthor=true&cauthor_uid=19959182)**. Lymphocyte infiltration, expression of interleukin (IL) -1, IL-6 and expression of mutated breast cancer susceptibility gene-1 correlate with malignancy of canine mammary tumours.** [**J Comp Pathol.**](http://www.ncbi.nlm.nih.gov/pubmed/?term=19959182) **2010 Feb-Apr;142(2-3):177-86. doi: 10.1016/j.jcpa.2009.10.023. Epub 2009 Dec 2. PMID 19959182.**

***Sentence 3: The present study investigates TILs and the presence of related cytokines, as well as the expression of breast cancer susceptibility gene-1 (BRCA1), in canine MGTs. (4/6)***

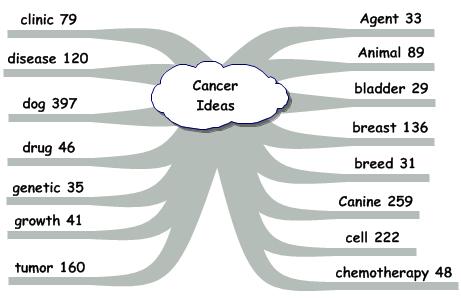
The sentence from the 2004 document showed three possibilities. The term – stem – is highlighted in green. That indicates that the software correctly identified the term but, it was used to denote another situation and not the one intended by the authors. Stem was included in the vocabulary to describe the combination – stem & cell. The authors used the verb – ***stemming*** – instead of the more commonly used verb – ***resulting***. This showed a problem with pattern recognition. The terms in red indicate correct identification by the software. The two terms in blue show the omission of terms with general rather than specific meaning in dog cancer.

The document from the 2009 document illustrates a problem associated with the use of a variety of text editor programs in preparing the abstracts. Each program introduces hidden characters that hinder recognition of terms. In this case, the term – ***breast*** – was missed. The term – gene – was not included in the vocabulary due to infrequency of use.

**Idea Maps:** Idea maps are graphic displays of relationships between a central term or combination of terms (i.e., central idea) and related ones. They are formed by considering the ideas presented by subject specialists. The displays can take various formats. These include:

1. Alphabetic array of the terms linked with a central term.
2. Frequency array of terms linked with a central term.
3. Dimensional array of terms linked with a central term.

**Figure 1. Alphabetic Array Using Informative Terms From Cancer Related Ideas.**

****

The ideas shown in Figure 1 depict relationships between the central term – cancer – and the informative terms occurring with higher frequency. The terms are arrayed beginning on the right and continuing to the left. Each is shown with the number of times that term was linked with cancer. The alphabetic array is useful as a type of subject index providing an assessment of the importance assigned to each idea by the subject specialists, assuming that importance is determined by frequency of use. The alphabetic array is helpful to the subject specialist interested in a specific idea by highlighting its existence and frequency.

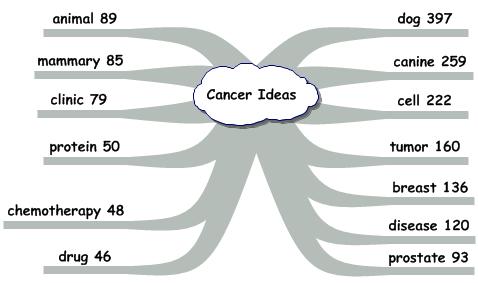
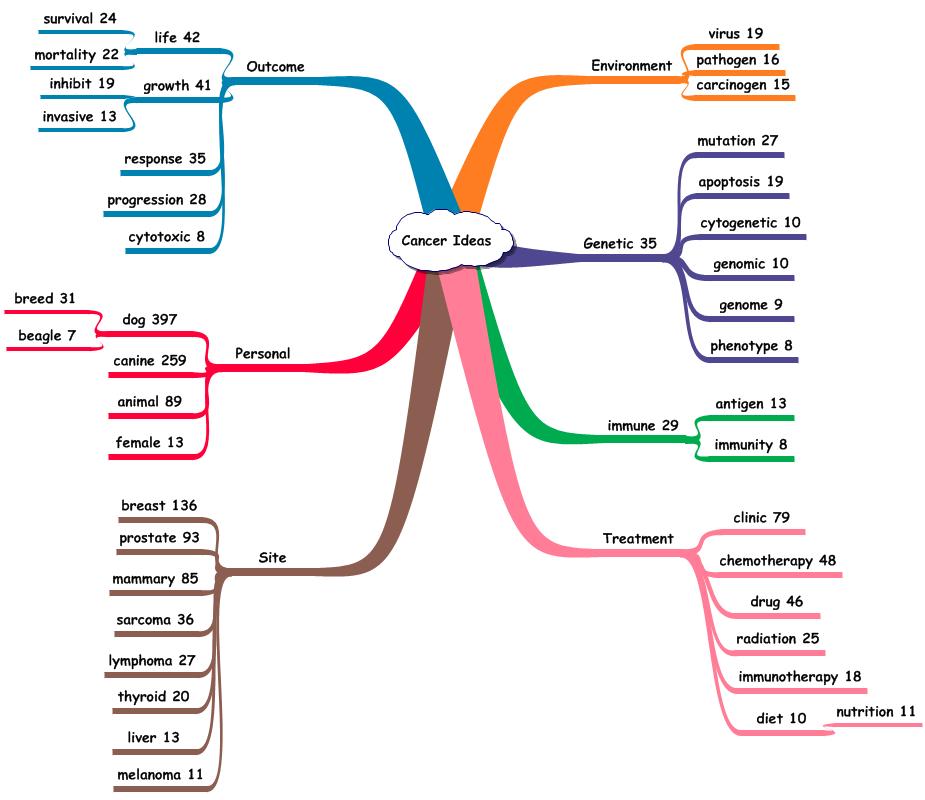
**Figure 2. Frequency Array Using Informative Terms From Cancer Related Ideas.**

Figure 2 shows the higher frequency terms related to the central term – cancer. The terms are shown with the most frequently occurring ones on the right and proceeding down and around on the left. This display serves as an index showing the ideas considered in terms of relevance (i.e., frequency) by the subject specialists. This display is most useful to the student interested in acquiring an understanding of a new topic. By focusing on the highest frequency ideas, the student can rapidly prepare a description of the topic.

Figure 3 shows the higher frequency ideas involving the central term – cancer – arranged in dimensions representing major components of the topic. Those dimensions are shown as arms in the idea map. The terms were classified using the meaning and/or function. This classification scheme is useful in formulating research designs. The dimensions represent components that must be addressed. The personal factors describe the characteristics and attributes of the individuals or materials chosen to study. The environmental factors describe the physical, social, and psychological characteristics of the environment in which the individuals live. The subject factors represent the physical, chemical, social and psychological attributes that make up the body of knowledge regarding the topic. The intervention factors are physical, chemical, biological, psychological, or educational acts used to influence change in specific subject factors. The outcome factors represent the characteristics or attributes being affected by the interventions. The methods describe the procedures used in the study.

**Figure 3. Dimensional Array Using Informative Terms From Cancer Related Ideas.**

****

**Idea Maps and Research:** By identifying the ideas within each dimension, different constructs can be considered. For each, the higher frequency ideas form a basic structure of acceptable and relevant knowledge. Often, lower frequency ideas will be introduced as testable elements. The composite will determine the impact of the low frequency ideas in expanding or enhancing the knowledge structure.

Idea maps serve to provide a gestalt view of the knowledge structure based on the ideas. Figure 4 shows an expansion of the dimensional map. Each of the terms was linked with cancer. An initial question is the interrelationships among the terms. In addition to being linked with cancer, have they been linked by the subject specialists to each other? To illustrate that possibility, the terms shown were linked with cancer and with tumor. Those triadic ideas, i.e., ***cancer & tumor & related term***, provide more specific information about the knowledge structure.

These triadic ideas can be used to construct sentences describing the topic. Example sentences might be:

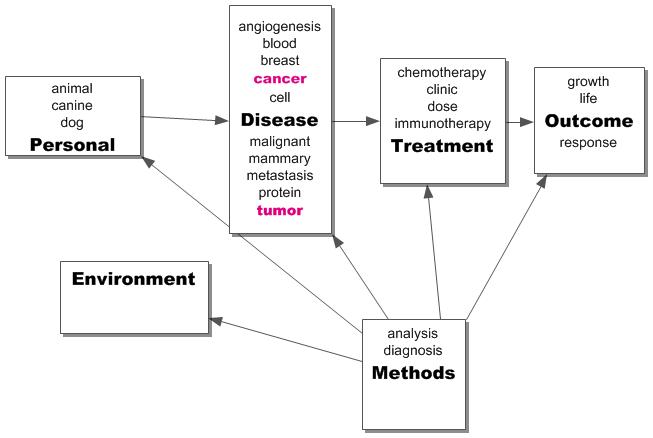
**Informative terms**: Animal 🡪 ***cancer 🡪 tumor*** 🡪 malignant 🡪 chemotherapy 🡪 response.

**Possible Sentence: *Animals with malignant cancers may benefit from chemotherapy in achieving a response shown by a decrease in the size of the tumor?***

**Informative terms:**Canine 🡪 ***cancer 🡪 tumor*** 🡪 malignant 🡪 immunotherapy 🡪 growth

**Possible Sentence:** ***Canines with malignant growths of tumors may benefit from cancer immunotherapy?***

**Figure 4. Dimensional Idea Map Including Terms Linked with Cancer and with Tumor.**

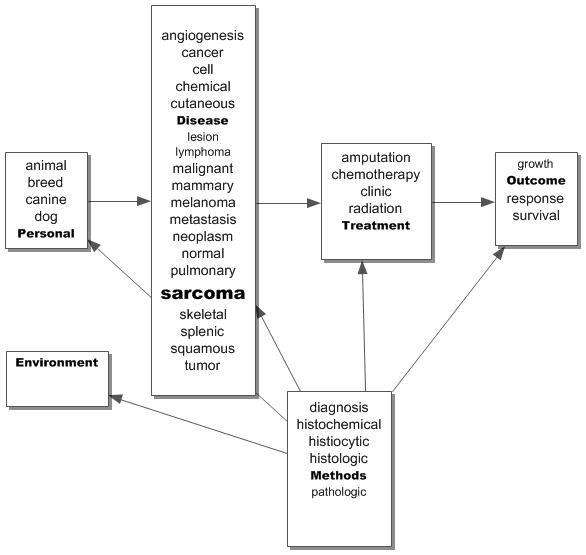
****

**An Example Using the Term -- Sarcoma:** The ideas involving the central term – ***sarcoma*** – were retrieved and organized. Figure 5 shows the higher frequency terms, each linked with the term – sarcoma -- and organized into dimensions representing different aspects of the sarcoma relationship. The ideas were classified using the meaning and/or function of the related term.

Figure 5 versus Figure 4 shows the expansion of ideas when the role of sarcoma is considered. The study groups added breeds. The interventions included amputation and radiation. The subject factors included sites, tissues, and types of cancer. The methods were expanded. The idea map in Figure 5 could serve to stimulate development of a number of interventional studies. The interventions could be explored to determine the ones that would be most effective.

**Uniformity of Starting Knowledge:** Figures 1-5 showed the value of the idea map in identifying and describing existing knowledge. With those, the individual could develop new research strategies. The savings in time and energy and more importantly, the changes in focus to essentially fulltime attention to the higher cognitive functions are significant.

**Figure 5. Idea Map Showing Terms Linked with the Central Term – Sarcoma.**

****

Given the existence of the idea database, uniformity of knowledge can be realized at least at the outset. With a comprehensive and accurate representation of the subject specialists’ ideas in a usable form, each investigator starts with an accepted and comprehensive set of specific ideas, concepts, and issues. What is done subsequently is a matter of individual intellectual prowess. Working with ideas offers the opportunity to make those subsequent cognitive functions more transparent and subject to quality-control. In that way, uniformity of approach to critical and creative thinking is a possibility. The findings can be diverse depending on the measures, criteria, and decision-rules employed. The uniformity stems from the application of the specific procedures in carrying out each cognitive function.

**Uniformity of Study Design:** Without the centralized idea database, coordination of research required at least two levels of oversight. One associated with funding and the second with the directions of research. With the centralized idea database, the changing idea map serves to coordinate the research by displaying the work accomplished and enabling investigators to identify the elements needed. Communication among the investigators via email and at conferences could fulfill the need for interaction and assignment of future work. In this milieu, funding of the preliminary studies is not only acceptable but important in rapidly describing and developing definitive studies.

**Communication Among Investigators:** The ability to work in the cooperative group setting while at the same time performing vital preliminary descriptive-intervention studies is important in expanding the knowledge regarding a particular topic and in benefiting the individuals effected by the problem. The central database of ideas would assist in bringing a consensus among the experts with respect to the starting level of knowledge and the need for coordinated investigations. The intellectual benefit would be seen in a transparent, quality-controlled approach to critical and creative thinking.

The idea database is a feasible approach to the important first steps in the organization and utilization of information:

1. The process of building the idea database is accurate and comprehensive. Of importance, the quality control process enables determination of the degree of accuracy and the extent of comprehensiveness.
2. The resulting ideas are relevant to the topic. The authors demonstrate their intentions relative to the topic by linking the topic descriptor to related terms.
3. The time and effort involved is shorter than other available approaches. The analysis of the text and subsequent organization takes about 0.3 minutes.
4. The associated costs are considerably cheaper as the software is performing the bulk of the effort. The cost of the individual’s time and energy is transferred from the clerical tasks involved in retrieval to the intellectual ones.

**Research Agenda:** Once the idea database is operational, the development of a research agenda is efficient because of the ease in arranging the ideas to form different syntheses. The higher frequency ideas provide a consensus description of what’s known at a given time. The lower frequency ideas represent two situations – one depicting the disappearance of ideas with time. The second represents the emergence of new ideas. Temporal analysis can readily identify the two situations. The impact of the new ideas on the established structure can be assessed and those likely to represent important change can be considered in new research strategies.

Since ideas are independent building blocks they can be combined in various ways depending on the intent of the analyst either to duplicate existing idea structures or to build new ones. The central database and the higher cognitive function algorithms create a different type of expert. Traditionally, a specialist would be chosen for his/her wisdom and opinions. With the tools described above, the expert is needed to clarify and expand the higher cognitive functions while the more clerical/mechanical ones are converted to a transparent, evidence-based system. By making the analysis transparent and quality-controlled, the needed documentation is readily available. This facilitates the shift to true intellectual prowess by the expert.

In a similar fashion, the student, new to the subject, can begin learning by solving problems and by building new idea structures. The ability to acquire, organize, and utilize the ideas enhances the transformation from novice to professional. The need to spend long hours in the library stacks is replaced by a need to spend time thinking and researching. Those actions could yield an operational description of critical and creative thinking.

With the ideas organized and usable, students can build numerous idea structures representing a given topic. This array shifts the focus from a single answer to a spectrum of possibilities each with desirable and undesirable characteristics. Weighing those attributes and developing rationales is an example of the critical thinking process. Translating those functions to transparent, quality-control procedures is an example of the Idea Analysis approach.

**Transforming Ideas to a Research Design:** Ideas can be explored by considering how the authors used them in sentences. Authors select and arrange the informative terms to express the ideas of interest in a coherent way. The arrangements are not the only ones possible. Since ideas are independent building blocks they can be combined in various ways depending on the intent of the analyst.

**Multi-Dimensional Research:** An important feature of research involving multi-dimensional knowledge is the realization that numerous subject factors and interventions may be required before significant change is observed in an outcome factor. This form of research differs from the traditional laboratory experiment where a single factor is influenced to accomplish a specific change. The complexity of multi-dimensional research requires that learning progresses in phases.

1. The first phase deals with identification of possibly ***relevant attributes*** of characteristics.
2. The second phase deals with determination of ***possible changes*** in those attributes.
3. The third phase deals with comparison of ***degree of change***.
4. The fourth phase deals with ***short term and long term effects*** of the change.

**Combinational Designs:** Frequently, there is an attempt to combine those phases to save time. The accepted combination is an interventional study requiring large samples of study subjects, comparable interventions designed to show the influence of the new one, and longer periods of observation. Those studies could be expensive involving numbers of personnel, more measurements, and questionable results. Without the proper problem descriptors, the success of the project is unclear.

**The Descriptive-Interventional Design:** The distinction between descriptive and interventional studies is in the acceptance of current knowledge from the literature. Descriptive studies are based on the assumption that the evidence regarding a phenomenon is weak and more is needed. Intervention studies are based on the assumption that the current description is adequate to develop a test of effectiveness. As a result of the investigation, the description is enhanced and the evidence dealing with the relationships is increased. Intervention studies are of value in building an evidence-based, data-driven description of the phenomenon.

Researchers often adopt descriptive studies instead of interventional ones. The descriptive ones were intended to develop the knowledge structure representing the topic. Often, these studies took on a life of their own, and became the end rather than the starting point originally intended.

An alternative to this emphasis on description is the preliminary interventional study. This design combines the descriptive elements with an intervention. The sample size is relatively small and the measurement set restricted. The objective of the study is to develop a ***more specific description together with an initial assessment of the degree of change*** that might be observed. The results of this type of study could lead to other small descriptive-interventional studies. The findings would eventually be sufficient to formulate a definitive study with the essential subject factors and interventions in place. That study would conform to the requirements of sample size, personnel, conduct, and duration. The chances of developing interpretable results would be enhanced.

This descriptive-interventional study design often is described as preliminary or pilot. The study is effective in developing more specific questions for study but rarely produces answers. While these preliminary studies provide insights with respect to the characteristics needed for accurate description plus estimates of the degree of change offered by the interventions, funding agencies tend not to fund them. In terms of fundable studies, there are two primary types – the large survey (descriptive) study and the large intervention study with randomized assignment to the comparable interventions.

**Advantages of the Descriptive-Interventional Design:** The advantages of this study process are numerous. The descriptive-interventional study process can be carried out in many environments providing insights across different study groups and physical settings. A measure of ‘truth’ in science is the ability to replicate the findings in different situations. This consistency may be of greater interpretability than a single large study.

A second advantage may be the identification of different relevant attributes and characteristics together with the measurements/observations depicting them. Having different investigators pick, independently, the measures, criteria, and decision-rules in each study offers a more complete picture of the knowledge structure than if everyone employs the same study method, particularly early on. Uniformity is appropriate when the definitive study is launched. The preliminary studies should employ different methods in order to determine the information provided by each.

A third advantage is the ability to employ a variety of interventions. These findings would be relevant in determining the better ones to employ in the definitive study. The combination of subject factor to be influenced, the intervention chosen to be the influencer, and the outcome factor representing the degree of change provide an important component in the determination of the ultimate study. Having different investigators determine these interventions again offers a form of consensus when the same intervention is studied in different places and a diversity of results when different ones are studied.

**Essential Difference – Now vs. Proposed:**  Small studies have been performed through time and the results published. The uncertainty in this process lies with the question of use of the information. Since these studies are small in size and may be unable to declare definitive findings, the attention they receive by other investigators could be variable. In contrast, if these small descriptive-interventional studies are part of a coordinated approach, the findings are added to the knowledge base and used to develop new studies. Progress is coordinated and knowledge acquisition accelerated.

**Idea Maps, Research Studies and Uniformity:** The advantages of uniformity are important in science and learning. However, the emphasis on uniformity has been misplaced. The importance of a uniform knowledge base of data describing the topics cannot be dismissed. Previously, the uniformity was seen in the development of bibliographic repositories, containing documents and records describing the presumed content of those documents. This warehousing approach was a valuable first step. In numerical problems, the concept was changed to a combination of warehousing and knowledge base offering credible data available via a number of portals. This combination has not yet been accomplished with text and is the objective of the idea analysis approach.