**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.

**Table 1. Excerpt of Hurricane Ideas Vocabulary.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Hurricane Ideas** | |  | **1990-2013** |  |
| **ALPHABETICAL ORDER** | | | **FREQUENCY ORDER** |  |
| **Term** | **Freq** |  | **Term** | **Freq** |
| academic | 1 |  | Hurricane Disaster Ideas | 2784 |
| access | 4 |  | disaster | 108 |
| accident | 1 |  | health | 61 |
| accurate | 1 |  | flood | 47 |
| action | 4 |  | experience | 46 |
| activation | 1 |  | stress | 41 |
| active | 4 |  | provide | 39 |
| activities | 2 |  | mental | 37 |

**Example from Hurricane Research:** The data considered in this book are taken from scholar publications dealing with disaster-related events entered into PubMed (1990-2013). The analysis involved over 29,000 documents (1990-2013). There were over 2700 ideas involving the term – hurricane. Disaster was linked with hurricane 108 times while health and hurricane occurred 61 times. Disaster science was selected because of the interest in developing more effective ways of dealing with natural and manmade crises. As an example, Table 1 shows an excerpt of the vocabulary used by authors in forming ideas describing characteristics of hurricanes. These data are part of an ongoing construction of a database of disaster-related ideas for the National Center for Emergency Preparedness, Vanderbilt University.

The large number of ideas that must be processed to find the ones that are relevant is a condition of text processing. Of the various ways used to describe documents, accuracy and completeness are accomplished best by identifying the authors’ ideas. To do that means that each document must be read, the ideas extracted and organized for future use. The use of software enabled this effective triaging in a rapid and efficient fashion.

Table 2 shows an excerpt of the ideas involving hurricane. The table shows ideas involving the central term – hurricane. Each record contains that term plus a related one and the locational data. The first sentence shown deals with document identified as PMID 1485290 in the PubMed repository. That sentence contained six informative terms. Hurricane is the primary term of interest and is linked with the five related terms. Those are highlighted in the identified sentence. The sentence was -- ***In this article we describe the efforts of local authorities to detect and treat casualties caused by a hurricane that struck the west coast of Norway January 1st, 1992 and prevent further injuries. [*[Ranhoff AH](http://www.ncbi.nlm.nih.gov/pubmed?term=Ranhoff%20AH%5BAuthor%5D&cauthor=true&cauthor_uid=1485290),** [**Naustdal H**](http://www.ncbi.nlm.nih.gov/pubmed?term=Naustdal%20H%5BAuthor%5D&cauthor=true&cauthor_uid=1485290)**,** [**Skomsvoll JF**](http://www.ncbi.nlm.nih.gov/pubmed?term=Skomsvoll%20JF%5BAuthor%5D&cauthor=true&cauthor_uid=1485290)**. [Injuries following a hurricane in Nordmøre].** [**Tidsskr Nor Laegeforen.**](http://www.ncbi.nlm.nih.gov/pubmed/?term=1485290) **1992 Dec 10;112(30):3777-80. [Article in Norwegian] PMID 1485290]** These six terms involving hurricane are shown in Table 2 and identified as being in sentence 1. The ideas from sentence 4 for PMID 2372228 also are shown in the Table. These ideas illustrate organization by columns labeled -- Hurricane, Related, Ident and Sentence.

**Table 2. Excerpt of Ideas Dealing with Hurricane.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Hurricane** | **Related** | **Year** | **Ident** | **Sentence** |
| hurricane | authorities | 1992 | 1485290 | 1 |
| hurricane | casualties | 1992 | 1485290 | 1 |
| hurricane | cause | 1992 | 1485290 | 1 |
| hurricane | injuries | 1992 | 1485290 | 1 |
| Hurricane | prevent | 1992 | 1485290 | 1 |
| hurricane | concept | 1990 | 2372228 | 4 |
| hurricane | coping | 1990 | 2372228 | 4 |
| hurricane | disaster | 1990 | 2372228 | 4 |
| hurricane | mechanism | 1990 | 2372228 | 4 |
| hurricane | nurse | 1990 | 2372228 | 4 |
| hurricane | population | 1990 | 2372228 | 4 |
| hurricane | psychosocial | 1990 | 2372228 | 4 |

**Selection of Vocabulary:** Table 1 showed the selection of vocabulary based on the frequency of linkage with the primary term (hurricane). Table 2 showed an excerpt of the ideas extracted from the text based on the vocabulary. The process involved in selecting the vocabulary is the first of the algorithmic approaches. Two measures are considered. The first is the ***frequency of use of the idea***. The ideas used most often tend to represent a consensus acceptance of relevance given a selected primary term. This relationship is considered to be ***direct***, that is X is used frequently by authors with Y. The second is the ***consistency of use*** across different situations. That depicts another form of consensus where the term is used with the different terms linked to the selected primary. This relationship depicts an indirect link with the primary term. That is, Z is related to the set {X} that are, in turn, directly related to Y.

Table 3 shows a summary of the expansion and use of the term-idea identification process. Each column label occurred with high frequency with the primary (hurricane). The ideas associated with each column heading were retrieved and the higher frequency terms-ideas are merged into the table. Table 1 showed the terms linked frequently with hurricane. Table 3 shows the terms (row variables) linked with each respective column variable. Ideas below the level defined as ***higher frequency*** are assigned the value of zero in this analytic table.

**Table 3. Analysis of Terms-Ideas to Identify Vocabulary.**

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Term** | **Hurricane** | **Disaster** | **Health** | **Flood** | **Experience** | **Stress** | **Provide** | **Mental** | **Response** | **Consistency** | **Sum** |
| health | 61 | 1100 | 0 | 156 | 295 | 164 | 863 | 1707 | 519 | 8 | 4865 |
| disaster | 108 | 0 | 1100 | 138 | 374 | 250 | 608 | 622 | 849 | 8 | 4049 |
| mental | 37 | 583 | 1707 | 91 | 161 | 214 | 269 | 0 | 204 | 8 | 3266 |
| public | 33 | 306 | 2024 | 77 | 0 | 0 | 288 | 224 | 358 | 7 | 3310 |
| emergency | 25 | 477 | 794 | 0 | 328 | 0 | 711 | 209 | 728 | 7 | 3272 |
| provide | 39 | 498 | 863 | 78 | 210 | 0 | 0 | 269 | 339 | 7 | 2296 |
| response | 36 | 679 | 519 | 70 | 0 | 121 | 339 | 204 | 0 | 7 | 1968 |
| experience | 46 | 374 | 295 | 71 | 0 | 166 | 210 | 161 | 0 | 7 | 1323 |
| medical | 31 | 578 | 544 | 0 | 195 | 0 | 579 | 0 | 411 | 6 | 2338 |
| clinic | 20 | 0 | 318 | 0 | 210 | 0 | 303 | 150 | 108 | 6 | 1588 |
| management | 0 | 491 | 293 | 75 | 139 | 0 | 242 | 0 | 144 | 6 | 1384 |
| nation | 0 | 318 | 526 | 0 | 97 | 0 | 242 | 171 | 234 | 6 | 1109 |
| cause | 25 | 253 | 276 | 199 | 0 | 114 | 0 | 190 | 0 | 6 | 1057 |
| problem | 0 | 255 | 575 | 65 | 119 | 0 | 0 | 303 | 0 | 5 | 1317 |
| stress | 41 | 260 | 0 | 0 | 169 | 0 | 0 | 235 | 137 | 5 | 842 |
| person | 0 | 321 | 415 | 0 | 235 | 0 | 234 | 0 | 135 | 5 | 838 |
| model | 20 | 0 | 0 | 142 | 0 | 0 | 188 | 161 | 138 | 5 | 649 |

The two summary measures – ***consistency*** and ***frequency*** -- are shown. Consistency is determined as the number of column variables having ideas with each row label. This measure seeks to determine the number of different situations (i.e., column headings) coupled with each row variable. The maximum value would be the number of column labels. The frequency of ideas (SUM column) is determined by adding the row frequencies across the columns. This measure assesses the global interest in a row variable by subject specialists across the terms linked with hurricane. A relevant subset of terms directly or indirectly linked with the primary (hurricane) would include the higher frequency of use terms (direct links with hurricane) plus the higher consistency of use terms (indirect links with hurricane).

**Exhibit 1. Document 10895916 Published in 2000.**

**Source:** [**Wylie T**](http://www.ncbi.nlm.nih.gov/pubmed?term=Wylie%20T%5BAuthor%5D&cauthor=true&cauthor_uid=10895916)**,** [**Cheanvechai D**](http://www.ncbi.nlm.nih.gov/pubmed?term=Cheanvechai%20D%5BAuthor%5D&cauthor=true&cauthor_uid=10895916)**,** [**Seaberg D**](http://www.ncbi.nlm.nih.gov/pubmed?term=Seaberg%20D%5BAuthor%5D&cauthor=true&cauthor_uid=10895916)**. Emergency response team: Hurricane Georges in Key West.** [**Prehosp Emerg Care.**](http://www.ncbi.nlm.nih.gov/pubmed/?term=10895916) **2000 Jul-Sep;4(3):222-6. PMID 10895916**

**Sentence 2: Lower Florida Keys Hospital, which serves Key West and the Lower Keys, had previously been evacuated of inpatients and staff.** (2/2)

**Sentence 3: An emergency response team composed of three emergency medicine (EM) physicians and four EM nurses was sent at the request of the state to maintain emergency department (ED) operations at the hospital.** (5 / 6)

**Sentence 6: Initially, patients requiring hospitalization were evacuated, but as the storm neared, this was stopped**. (2/2)

**Ideas in Document 10895916 – Sentence 2**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Primary** | **Related** | **Year** | **Ident** | **Sentence** |
| patient | hospital | 2000 | 10895916 | 2 |
| hospital | patient | 2000 | 10895916 | 2 |

**Ideas in Document 10895916 – Sentence 3**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Primary** | **Related** | **Year** | **Ident** | **Sentence** |
| nurse | Emergency | 2000 | 10895916 | 3 |
| operation | Emergency | 2000 | 10895916 | 3 |
| response | Emergency | 2000 | 10895916 | 3 |
| emergency | Nurse | 2000 | 10895916 | 3 |
| operation | Nurse | 2000 | 10895916 | 3 |
| response | Nurse | 2000 | 10895916 | 3 |
| emergency | Operation | 2000 | 10895916 | 3 |
| nurse | Operation | 2000 | 10895916 | 3 |
| response | Operation | 2000 | 10895916 | 3 |
| emergency | Response | 2000 | 10895916 | 3 |
| nurse | Response | 2000 | 10895916 | 3 |
| operation | Response | 2000 | 10895916 | 3 |

**Ideas in Document 10895916 – Sentence 6**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Primary** | **Related** | **Year** | **Ident** | **Sentence** |
| patient | hospitalization | 2000 | 10895916 | 6 |
| hospitalization | Patient | 2000 | 10895916 | 6 |

**Verification of Informative Terms and Ideas:** Exhibit 1 shows an example of the verification process using a 2000 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 identified sentence(s) were shown with the informative terms highlighted. Each term is identified once within a sentence.
3. The idea records dealing with that sentence were shown.

Behind each sentence is a ratio giving the number of terms identified by the software relative to the number of terms available for identification. That ratio determines the precision associated with correctly identifying informative terms. The selection of informative terms can be evaluated by comparing the terms selected with those used by the subject specialist-authors. In the sentences in Exhibit 1, terms are color coded to represent different situations. The terms highlighted in red indicate that the software properly selected them. Those in green denote that the software missed them. Terms highlighted in blue were used by the authors, but, were not included in the vocabulary. The blue colored terms may be less frequently used in ideas, be inconsistent, and/or be general in meaning.

While pattern recognition should be 100%, text is prepared using a number of different programs. Some introduce hidden symbols that interfere with recognition and authors use punctuation in various ways so that the pattern is different from the one being sought. As such, correct recognition can vary from 66% to 100% with a median of 85% across different texts. The term, hospital, was missed by the software in sentence 3. The accuracy score was reduced to 83%. The 6th sentence contained a word – storm – that was not included in the vocabulary. It satisfied the criterion of specificity relative to the topic. It did not satisfy the criterion of frequency. Hence, the software performed correctly.

**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.

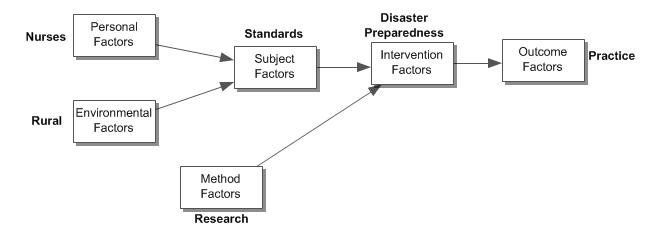
**Exhibit 2. Example of Terms in Sentence Indicating Research Strategy.**

**Source:** [**Anderson DA**](http://www.ncbi.nlm.nih.gov/pubmed?term=Anderson%20DA%5BAuthor%5D&cauthor=true&cauthor_uid=22320158)**. Using disaster exercises to determine staff educational needs and improve disaster outcomes in rural hospitals: the role of the nursing professional development educator.** [**J Contin Educ Nurs.**](http://www.ncbi.nlm.nih.gov/pubmed/?term=22320158) **2012 Jun;43(6):284-8. doi: 10.3928/00220124-20120201-24. Epub 2012 Feb 8. PMID 22320158**

**Sentence 3: This article explains the role of the rural nursing professional development educator as a disaster preparedness educator, facilitator, collaborator, researcher, and leader, using the American Nurses Association's Nursing Professional Development: Scope and Standards of Practice.**

Since ideas are independent building blocks they can be combined in various ways depending on the intent of the analyst. Exhibit 2 shows a sentence containing terms and ideas suggesting the initiation of a research strategy. The terms are highlighted in red and additional possibly informative terms not satisfying the selection criteria are shown in blue.

Figure 1 shows the terms and ideas, from the example sentence, arranged as a research strategy describing rural nurses as the study subjects, American Nursing Standards as the subject material, disaster preparedness education as the intervention, and nursing practice as the affected outcome. The method identified was research. The research strategy tends to combine high frequency and low frequency terms. In this example, rural, disaster, preparedness, research, American, nurse, and practice are the high frequency terms. The low frequency ones are nursing, professional, development, scope, and standards. When viewed this way, the research question would be – ***Are the nursing professional development scope and standards effective in enhancing nursing practices in the event of a disaster?***

**Figure 1. Research Strategy Based on Terms and Ideas.**

The dimensions shown in Figure 1 represent specific components in describing a desired action. The study subjects (***personal and environmental factors***) are presented with specific information or actions (***intervention factors***) based on existing knowledge (***subject factors***). The result of this is seen as a change in status (***outcome factors***). The procedures involved in accomplishing this action are ***method factors***.

The template approach can be employed to develop and evaluate various research strategies. As such, it is useful in the cognitive process leading to effective generation of new knowledge. It is a form of simulation enabling the analyst to visualize the system that will be transformed into specific procedures by the construction of the research protocol.

**Summary**

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. An initial reaction would be to challenge the veracity of the formal process. That challenge is precisely the reason for the establishment of the process. 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 at all levels 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.