**Chapter 2. Power of Ideas**

**Abstract**

**This report illustrates Contextual Analysis, a form of artificial intelligence methodology designed to facilitate transparent and quality controlled learning. The data employed are the subject specialists’ ideas as presented by combinations of informative terms (nouns, adjectives, and gerunds) within the authors’ sentences. This approach assumes that the author-specialist and the editorial and peer review approve of the terms and their combinations as presented. As such, these combinations are considered contextually relevant. A distinct feature of software is the ability to recognize patterns, thus, enhancing recognition of ideas in the sentences. Capture of these ideas and subsequent organization as data records means that text analysis can essentially begin with the higher cognitive functions. This eliminates the traditional repetitive search, retrieval, and analysis associated with manual text analysis. Ideas serve as building blocks leading to new descriptions of topics or new knowledge generating strategies and new research. This form of learning (called gestalt) is superior to linear. By providing the student with a more complete understanding of the subject at the outset, the ideas involved can be used to demonstrate competence in the subject.**

**Introduction**

**With each increase in the use of text, identification of essential ideas required more effort and took more time. Programs of education were extended to deal with the increase in information while at the same time procedures designed to deal with information did not keep up. Effective processing of information was assumed to be an integrative component of learning while each student was left to independently work out the methods to be employed. As a result, the transformation of information to knowledge required longer periods of time. There was increased emphasis on clerical activities (retrieval and analysis) rather than a focus on the higher cognitive functions (synthesis, comparison, evaluation, judgment, and application).**

**Formalized methods of information processing focused on terms and combinations of terms in identifying blocks of text. This approach was incorporated into software (i.e., search engines) and became the standard in identification of information. Two major problems were identified. *The first was misdirection*. The search engine did not identify the terms actually used by the authors so a form of “blind-man’s bluff” was created. This problem has since been corrected by expanding on the term provided by the user and displaying terms and combinations involving the supplied term.**

***The second was misinformation*. Terms can have different meanings depending on the context adopted by the author. Identification of terms didn’t take those different meanings into account. Thus text was identified that had little or no relevance relative to the user’s intent. As a result, as little as 10% of the retrieved information, might actually be informative assuming a specific context.**

**These problems can be eliminated by focusing on the ideas actually provided by an author. To do this, the text must be read, and informative vocabulary (nouns, adjectives, and gerunds), and ideas provided by authors in sentences must be identified and organized. Software can be built to accomplish these clerical tasks. Informative terms can be identified by the patterns of letters. Ideas can be identified by pairing the informative terms provided within each sentence. Not all of these pairs (i.e., ideas) will be considered important by the originating authors. However, time has an influence on the interpretation of importance of an idea. By capturing all of the ideas – seemingly important and unimportant – the scene is set for tracking change in the relevance of an idea. Importance of an idea can be estimated by the number of times author-specialists cite the idea in sentences. The high frequency ideas represent a consensus of knowledge while low frequency ideas may indicate growing or fading interest. Frequency through time is an effective measure of relevance.**

**The large volume of identified ideas is a new challenge identified by this approach. Fortunately, the ideas provide a solution. The authors combined terms to provide specific meanings. These relationships extend beyond the specific use provided by an author. As a result, ideas can be considered independent of a particular author’s use. Ideas can be organized in various ways, each offering a particular interpretation or description. This capability supports a transformation from the identification and analysis functions to the higher cognitive functions – synthesis, comparison, evaluation, judgment, and application. The shift not only involves change in functions but also a differential in effort. Idea processing is associated with a 90% effort in the higher cognitive functions instead of the typical 90% effort associated with identification and analysis.**

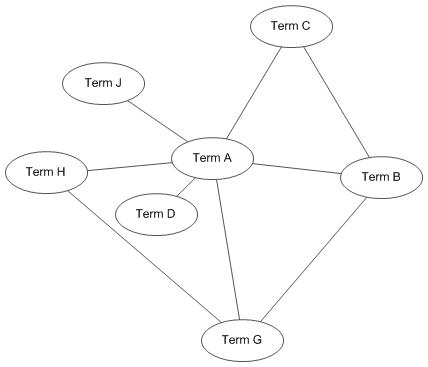
**Working with Ideas**

**Ideas represent a succinct way of communicating. The two terms linked together to form an idea can be expanded to a more conventional message. For example, the pair – A & B – can be expanded to – *Author X expressed the relationship, A linked with B, in a sentence from text 123.* This translation states a fact that can be supplied by novice or specialist. This feature enhances learning by a student as well as evaluation and interpretation by experts.**

**While a single idea might depict a situation, more frequently, a swarm of ideas form a concept that more completely describes the topic. That swarm is made up of related ideas.**

**For example, consider a term A linked with the other terms and is called the *primary* term. The authors who worked with the topic have identified the links. The network of links is determined by combining ideas containing a common primary term. At a particular time, not all of the terms will have been linked suggesting that the concept map is not yet complete. The missing links might serve as the basis for new research.**

**A complete map acts like a neuro net with the exception that each link has been identified and ‘verified’ by investigators so that the relationships exist in a variety of situations. That phenomenon has been used for years in establishing truths.**

**In contrast to statistical linking, the links in a concept map undergo scrutiny to determine the reality of the relationship and the relevance of the link in describing the concept. The concept map is a description of the details representing a topic. A hierarchy of descriptions thus exists involving the term or phrase representing the idea:**

1. **The ideas involved.**
2. **The network of links providing the details.**
3. **The text expansion of the ideas.**

**By completing this hierarchy for a topic, a more complete and rapid transformation of information to knowledge is possible.**

**In learning a new topic, the most effective way is to begin with the ideas, form a concept map of the ideas, expand the concept map by adding text, and finally summarizing the process by a single term or phrase. This contrasts with the established process which begins with a full text description. This is broken down into smaller text which represents one or more ideas. When the set of ideas is identified, they might then be organized as a concept map. While the same outcome is possible from both approaches, the established method is considerably longer and involves more clerical activity. This process is fatiguing, leaving little time or desire for engaging in higher cognitive functions.**

**An Example using Documents Dealing With Artificial Intelligence.**

**Illustration of the newer technical approach uses scientific reports from Artificial Intelligence entered into PubMed during 2013 - 2017. This bibliographic resource contains over 26 million documents covering a vast number of health related subjects. A particular advantage is that the new analytic approach to scientific literature provides the wherewithal to process large volumes of information. The results of an analysis are expected to organize the information contained in the text by differentiation, classification, translation, and interpretation. These acts are behaviors performed by subject specialists. If the methods are accurate and effective, software could accomplish these functions, hence the label – artificial intelligence.**

**In order for software to identify and deal with contextual relationships, there must be a structure enabling recognition. That structure involves the ideas presented within the domain of a sentence. Informative terms (nouns, adjectives, and gerunds) placed in association with each other in the sentence can be operationally defined as ideas. The assumption is that author-specialists place specific terms within the sentence because they are contextually related. A recognized feature of software is the ability to determine patterns and thus recognition of informative terms and their combinations.**

**These ideas can be identified, extracted, and organized to form data records, containing a pair of these informative terms (the idea), the year published, the document containing the data and the sentence involved. These records can be stored in data repositories for further use. The software is able to recognize and extract over 95% of the ideas presented in the document, thus eliminating the traditional repeated search, retrieval, analysis, and extraction associated with manual text analysis. As a result, the analyst can shift effort from these clerical functions to the higher cognitive tasks – synthesis, comparison, evaluation, judgment, and application. In addition, the procedures involved can be more transparent enabling quality control and monitoring by third parties.**

**The resulting creative paths can be managed in a fashion comparable to the proof of theorems. The inventor develops the results and provides the thought process enabling validation of the proclamation. Once established, countless numbers of students and professionals can perform the same acts yielding the same results. In addition, by using the same thought procedures, these third parties can expand the results by contributing new knowledge.**

**An advantage unique to the contextual analysis approach is a significantly shorter time directed to the cognitive functions. The idea structure is compatible with established data management methods – copying, sorting, combining, and classifying. These add accuracy and transparency as well. The savings in time and effort associated with easier retrieval allows the analyst to essentially begin with synthesis. The shift in analytic procedures also provides time and energy to tackle the true intellectual aspects of each higher function. These include – development of measures to describe the attributes of interest in each cognitive function, the development of criteria in dealing with the measures, and the development of decision-rules to interpret the measures and criteria.**

**Table 1. High Frequency Terms in Artificial Intelligence Documents 2013-2017.**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Term** | **2013** | **2014** | **2015** | **2016** | **2017** | **Sum** | **Dimension** |
| **Total** | **527368** | **421016** | **459451** | **133032** | **51020** | **1591887** |  |
| **method** | **17944** | **15210** | **16084** | **4315** | **1699** | **55252** | **method** |
| **model** | **15835** | **12740** | **14835** | **4336** | **1715** | **49461** | **method** |
| **network** | **10335** | **8843** | **9753** | **2943** | **1222** | **33096** | **method** |
| **learning** | **7197** | **8487** | **9631** | **2647** | **803** | **28765** | **method** |
| **machine** | **7488** | **8023** | **9275** | **2390** | **775** | **27951** | **method** |
| **analysis** | **8490** | **6132** | **7014** | **1981** | **953** | **24570** | **method** |
| **algorithm** | **7492** | **6588** | **7187** | **1823** | **720** | **23810** | **method** |
| **process** | **6499** | **5566** | **6324** | **1979** | **675** | **21043** | **method** |
| **support** | **6572** | **5809** | **6107** | **1636** | **598** | **20722** | **method** |
| **information** | **5810** | **5776** | **6264** | **1906** | **462** | **20218** | **method** |
| **classification** | **6075** | **5603** | **6404** | **1297** | **583** | **19962** | **method** |
| **neural** | **6557** | **6784** | **7923** | **2298** | **1004** | **24566** | **subject** |
| **performance** | **6310** | **5613** | **5855** | **1382** | **566** | **19726** | **outcome** |

**A Unique Feature of Ideas**

**Table 1 shows the higher frequency informative terms provided by author-specialists in their writings of artificial intelligence. The total number of involved ideas was 1,591,887. In addition to the frequency of citing each term, the terms are classified into dimensions comprising the information provided. Three dimensions are shown – methods, subject, and outcome. The methods dimension contains terms describing the procedures employed in performing the cognitive functions. The subject dimension contains terms that describe characteristics of the subject being studied. The outcome dimension describes the results of interactions amongst the dimensions. In this case, the interactions are between the methods and subject dimension. Other examples might show interactions involving the other remaining dimensions – personal factors, environmental factors, and interventional/treatment factors.**

**The choice of elements in a study would depend on the analyst. Selection might be based on one of three conditions:**

1. **Missing findings involving a term/dimension.**
2. **Inconsistent previous findings.**
3. **Previous findings in new environment.**

**Table 2. Potential Study Structures Based on Dimensions and Frequency.**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Methods** | **Freq** |  | **Subject** | **Freq** |  | **Outcome** | **Freq** |
| **method** | **55252** |  | **neural** | **-24566** |  | **performance** | **19726** |
| **model** | **49461** |  |  |  |  |  |  |
| **network** | **33096** |  |  |  |  |  |  |
| **learning** | **28765** |  |  |  |  |  |  |
| **machine** | **27951** |  |  |  |  |  |  |
| **analysis** | **24570** |  |  |  |  |  |  |
| **algorithm** | **23810** |  |  |  |  |  |  |
| **process** | **21043** |  |  |  |  |  |  |
| **support** | **20722** |  |  |  |  |  |  |
| **information** | **20218** |  |  |  |  |  |  |
| **classification** | **19962** |  |  |  |  |  |  |

**With high frequency terms as in Table 2, the third rationale might be considered. High frequency terms are typically used to develop new descriptions of concepts, rather than new study strategies. If a new research study is contemplated, the corresponding study hypothesis can be formed by reversing the study structure. That is, outcome 🡪 subject 🡪 method. The resulting sentence might be – *artificial intelligence (method) is used to determine performance (outcome) as a result of forming a neural structure (subject) determined by a specific method (method).***

**Table 3 shows an expanded set of higher frequency informative terms. In addition to the added terms, the set includes additions to the outcome dimension. The study structure would be – *Artificial Intelligence 🡪 Methods 🡪 Subject 🡪 Outcome.* The hypothesis skeleton would be – *Outcome 🡪 Subject 🡪 Methods*. A corresponding hypothesis sentence might be – *Performance (accuracy and ability) is the result of interaction between artificial intelligence methods and the neural net established.***

**Table 3. Relationships Among an Expanded Set of Informative Terms.**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Term** | **Analysis** | **Algorithm** | **Process** | **Support** | **Information** | **Classification** | **Dimension** |
| **Total** | **953** | **720** | **675** | **598** | **462** | **583** |  |
| **method** | **26** | **29** | **15** | **15** | **11** | **32** | **Method** |
| **model** | **21** | **19** | **22** | **26** | **8** | **18** | **Method** |
| **machine** | **11** | **26** | **4** | **55** | **2** | **19** | **Method** |
| **learning** | **9** | **27** | **18** | **8** | **7** | **18** | **Method** |
| **vector** | **7** | **8** | **1** | **55** | **2** | **11** | **Method** |
| **network** | **17** | **16** | **9** | **12** | **9** | **10** | **Method** |
| **image** | **14** | **7** | **8** | **3** | **6** | **20** | **Method** |
| **algorithm** | **6** | **0** | **9** | **9** | **8** | **12** | **Method** |
| **classification** | **11** | **12** | **4** | **14** | **4** | **0** | **Method** |
| **classifier** | **5** | **5** | **3** | **16** | **2** | **16** | **Method** |
| **analysis** | **0** | **6** | **17** | **9** | **10** | **11** | **Method** |
| **process** | **17** | **9** | **0** | **1** | **16** | **4** | **Method** |
| **provide** | **10** | **6** | **9** | **8** | **11** | **3** | **Method** |
| **artificial** | **13** | **9** | **6** | **8** | **6** | **2** | **Method** |
| **search** | **13** | **6** | **7** | **3** | **9** | **3** | **Method** |
| **signal** | **5** | **2** | **22** | **0** | **4** | **5** | **Method** |
| **support** | **9** | **9** | **1** | **0** | **2** | **14** | **Method** |
| **information** | **10** | **8** | **16** | **2** | **0** | **4** | **Method** |
| **detection** | **4** | **4** | **5** | **6** | **3** | **7** | **Method** |
| **experiment** | **7** | **5** | **4** | **5** | **1** | **6** | **Method** |
| **selection** | **6** | **5** | **4** | **5** | **5** | **9** | **Method** |
| **neural** | **14** | **15** | **10** | **10** | **9** | **9** | **Subject** |
| **ability** | **8** | **9** | **9** | **6** | **5** | **6** | **Outcome** |
| **accuracy** | **7** | **14** | **4** | **6** | **4** | **27** | **Outcome** |
| **performance** | **5** | **18** | **9** | **8** | **1** | **15** | **Outcome** |

**The neural net would involve the informative terms from Table 3, arranged in a simple structure as shown in Table 4.**

**As seen in Table 3, each of the terms is related to all of the other terms forming a complete network.**

**Table 4. Informative Terms Forming Nodes in the Neural Network.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Methods** |  | **Subject** |  | **Outcome** |
| **algorithm** |  | **neural** |  | **ability** |
| **analysis** |  |  |  | **accuracy** |
| **artificial** |  |  |  | **performance** |
| **classification** | |  |  |  |
| **classifier** |  |  |  |  |
| **detection** |  |  |  |  |
| **experiment** | |  |  |  |
| **image** |  |  |  |  |
| **information** | |  |  |  |
| **learning** |  |  |  |  |
| **machine** |  |  |  |  |
| **method** |  |  |  |  |
| **model** |  |  |  |  |
| **network** |  |  |  |  |
| **process** |  |  |  |  |
| **provide** |  |  |  |  |
| **search** |  |  |  |  |
| **selection** |  |  |  |  |
| **signal** |  |  |  |  |
| **support** |  |  |  |  |
| **vector** |  |  |  |  |

**Contextual Analysis vs. Statistical Analysis**

**Is Contextual Analysis more accurate in forming networks of nodes and relationships? Statistical analysis involves identification of high frequency terms and the correlations with the nodes in forming a network. The problem is simply identification of informative terms. High frequency occurrence in text is not sufficient in declaring a term to be informative. Some of those identified will be contextually relevant and others will be superfluous. Differentiation of the two types will require measures, criteria, and decision-rules apart from those inherent in the statistical analysis.**

**In contrast, Contextual Analysis identifies informative terms and relationships (i.e., ideas) provided by the subject specialist. That expert is supported by the publishing process including editorial and peer review. As such, the combinations of terms are more likely to be contextually relevant. An accepted measure throughout text analysis is the frequency of citation across authors. A primary criterion is the extensiveness of the relationships involving a particular term. The primary decision-rule is the ease in establishing a meaningful network. These conditions satisfy the requirement to form a new study strategy (including a hypothesis) or a new description of a topic. The process eliminates the need for months of literature review by relying on the consensus of author-experts, as determined by software designed to identify, extract and organize the ideas. In addition to being comprehensive, the process is instructive.**

**Ideas into Concepts**

**Table 5. Ideas Involving – Algorithm – as the Central Terms.**

|  |  |  |
| --- | --- | --- |
| **Term** | **Freq** | **Dimension** |
| **method** | **29** | **method** |
| **learning** | **27** | **method** |
| **machine** | **26** | **method** |
| **model** | **19** | **method** |
| **network** | **16** | **method** |
| **classification** | **12** | **method** |
| **computation** | **11** | **method** |
| **performance** | **18** | **outcome** |
| **accuracy** | **14** | **outcome** |
| **automatic** | **13** | **outcome** |
| **optimization** | **12** | **outcome** |
| **genetic** | **17** | **subject** |
| **neural** | **15** | **subject** |
| **Total** | **720** |  |

**An interesting aspect of working with ideas is their ability to form concepts simply by sorting on the related terms associated with a central informative term. Table 5 illustrates this using the term – algorithm – as the central one and by sorting the related terms from higher frequency to lower. This array serves to expand the meaning of the central term. In contrast to dictionary descriptions, the idea configuration depicts a more comprehensive set of attributes. By sorting on frequency of author citations, the more frequent terms provide the following description. Firstly, *algorithm* is a *method*. It is used to determine *machine learning*. It’s a *model* for *network classification* and *computation*. It offers *performance* and *accuracy* in what appears to be an *automatic* manner. It has been applied to *genetic* analyses and *neural* networks. While the most frequent terms are most descriptive, even the lower frequency terms provide insights. For example, one of the methods involving algorithms is *simulation*. That related term was offered seven times in the period 2013 – 2017.**

**Summary**

**Artificial Intelligence is composed of statistical and contextually relevant methods designed to reveal and expand knowledge. The Contextual Analysis approach, illustrated in this report, shows the relevance and ease of working with ideas. A primary advantage is the ability to recognize, extract, and organize the ideas presented by author-specialists. Another advantage is the accuracy and comprehensiveness of the extraction resulting in over 95% of the ideas being stripped from the original text and stored in databases, thus facilitating ready use by all concerned. That organizational process eliminates months of redundant clerical work and facilitates early institution of intellectual function.**

**More efficient retrieval and analysis is not the essential feature of Contextual Analysis. Eliminating most of the clerical effort in text analysis is desirable but incidental to the real benefit of the focus on ideas. These intellectual building blocks can be managed using accepted and effective data management tasks. They can be sorted to develop concepts. They can be classified to demonstrate the different sub-topics making up a topic. They can be counted to determine a form of consensus across author-specialists. These arrays can be accomplished in seconds enabling formation of insights into the meaning of terms and concepts within subjects. These intellectual benefits can be demonstrated using procedures enabling transparency and quality control. The mystery associated with private and personal mentation is eliminated. As such, the months or years required to transform the novice-student into the professional are reduced to days.**

**A different form of instruction can be contemplated. Consider an introductory instruction program that consists of the following:**

1. **On day 1, the instructor introduces the idea map describing the key informative terms and the relationships among these nodes. This gestalt view is associated with the requirement to describe and discuss the methods used to identify and organize these nodes.**
2. **On day 2, the instructor asks the students to consider ways in which the higher frequency nodes and ideas can be organized as new descriptions of the subject or as new knowledge generating strategies.**
3. **On day 3, the instructor asks the students to consider ways in which the lower frequency nodes and ideas can be organized as new descriptions of the subject or as new knowledge generating strategies.**
4. **On day 4, the students are asked to develop research strategies using the higher frequency terms and ideas.**
5. **On day 5, the students are asked to develop research strategies using the lower frequency terms and ideas.**
6. **On day 6, the students are asked to develop research strategies using combinations of higher and lower frequency ideas.**
7. **On day 7, they rested.**

**While some students may require more than a week and others less, the fact is that linear learning is considerably slower and more tedious than gestalt learning. Further, working with ideas and computer algorithms is faster and more transparent. The result is that in a significantly shorten time, understanding and importantly, the ability to develop new research strategies in a subject will be demonstrated. This demonstration is a hallmark of competence in dealing with the subject.**

**The results of such instruction would be a new cadre of prepared investigators introducing new knowledge in a condensed time. Problems requiring years for solution now, may be solved in weeks. The benefits to the global wellbeing may be significant and may eliminate the longstanding conflicts common today. These benefits associated with efficient and effective management of information and knowledge are available if attitudes about linear and gestalt learning are changed.**