Evaluation of Information

In this Chapter, you will:

1. learn about the extraction and summarization of information found;
2. learn to analyze information by means of content analysis and quantitative analysis using Personal Computer (PC) tools;
3. learn to evaluate information based on the context of information source;
4. learn about the multiple criteria considerations for decision making in information evaluation.
Objective 1: Extraction and Summarization of Information Found

In the previous section, we have studied the access of information. One can always make use of the resources in library, the Internet, search engines or study groups to access his/her information required. However, how can one tell if the information collected is of his/her need in tackling the problems which (s)he is facing? Before we reach the answer, the extraction and summarization of information found forms a vital stage before the information evaluation process, which is sometimes referred to as the preprocessing stage of information evaluation.

Information Preprocessing

In this stage, the primary goal is to locate the information that we are looking for according to the tasks which we are given, the assignments which we have to do or the problems which we are facing. Therefore, we first need to define what information we actually need. In the case of a student working on a project on “Ancient Egypt”, for instance, Egypt is the first term that (s)he should be kept in mind. After an intensive search from the Internet or the library (Refer to the previous section for details), a pile of information on Egypt might be gathered, among which may contain different aspects of Egypt such as its economy, geography, population, military development, political status, plantation, living conditions, history, etc. Since it is the ancient Egypt that is actually required, the pile of information must now be divided into two categories, namely, ancient Egypt information and non-ancient Egypt information. On our dividing the information on Egypt into the two groups as mentioned, we have imposed a scope on the information found. Therefore, scoping information is the first step in information extraction after the definition of our actual need.

After scoping, we have probably eliminated all information unrelated to our definition of actual need. The information left, however, might be too much to fulfill our need, or it might be too little for us. In case of the information remained being too little, we can always return to the resources which we have used before in order to gather more results, which are then followed by scoping. On the other hand, if we have too much information after scoping, we probably need to perform information sub-categorization and focus selection. Reasons for these two steps vary among users. Time is always the prime concern though. Another major reason could have been limited human and/or capital resources. Referring to the project on Ancient Egypt, the student has found that after scoping, there still much information left but (s)he needs to rush for the deadline of the project submission. So, (s)he now needs to sub-divide the information remained according
to its nature and then choose one or two focuses to be the topic of his/her homework. The following are what (s)he have found after sub-division of information,

- Ancient Egyptian History (some information)
- Ancient Egyptian Invention (some information)
- Ancient Egyptian Landscape (little information)
- Ancient Egyptian Living Style (plenty of information)
- Ancient Egyptian Food (plenty of information)
- Egyptian Plantation (plenty of information)

Among which, the student finally decided to put focus on Ancient Egyptian Food and Living Style because of his/her interest and, much more practically concerned, because of the information amount being manageable by him/her. If the focus which the student selected (e.g. Ancient Egyptian Landscape) does not contain enough information for his/her project but (s)he still wish to work on it due to personal preference, another information search becomes inevitable. However, the search criterion must be modified as “Ancient Egyptian Landscape”, instead of “Ancient Egypt”.

The processes of sub-categorization of information and category-selection form the second step of information extraction. After this step, our criteria of information will also be refined due to our focus on a particular aspect of information. The following diagram summarized the entire information extraction process. (Fig. 1)
Followed by the extraction of information, we should be much more certain of the scope and the direction of our needed information. However, among what we gathered so far, there might exist duplication of ideas, or there might be many similar ideas. Our need to eliminate all those similarities pushes us to set information summarization to be our next goal. In the case of the Ancient Egypt project, after the student established the focus on food and living style, (s)he will start reviewing all articles concerning these two aspects. The student can make use of a blank white sheet of paper to jot down the ideas, recording just once for all similar ideas. If two or more articles mentioning the same idea, (s)he needs to record the name of those articles as well. After the summarization process, the student will have a generalized piece of information on top of his/her pile of articles on food and living style. As the pile of articles after summarization might still be required before an assigned task ends, one should keep the information for possible future references. After all, in the end of most tasks or academic work, we will often be required to provide a list of references. If we dump our searched results once the summarization process is over, we will not be able to construct a list of references in the end. Details on how a list of references should be presented will be introduced in the next topic on “Communication and Presentation of Information”.

Department of Information Systems
Objective 2: Analysis of Information

Upon what has accomplished in objective 1, we will take a much more detailed look at our information, attempting to gain a much deeper understanding of what it is all about. Content analysis and quantitative analysis are two normal approaches we would adopt in face of this need. As the information we gathered is used to appear as either textual or numerical information, we will use content analysis for textual information and quantitative analysis for numerical information.

Content Analysis[3]

Content analysis is a research tool used to determine the presence of certain words or concepts within texts or sets of texts. Researchers quantify and analyze the presence, meanings and relationships of such words and concepts, and then make inferences about the messages within the texts, the writer(s), the audience, and even the culture and time of which these are a part. Texts can be defined broadly as books, book chapters, essays, interviews, discussions, newspaper headlines and articles, historical documents, speeches, conversations, advertising, theatre, informal conversation, or really any occurrence of communicative language. Texts in a single study may also represent a variety of different types of occurrences. To conduct a content analysis on a piece of text, the text is coded, or broken down, into manageable categories on a variety of levels – word, word sense, phrase, sentence, or theme – and then examined using one of content analysis' basic methods: conceptual analysis or relational analysis. Conceptual analysis can be thought of as establishing the existence and frequency of concepts – most often represented by words of phrases – in a text. For instance, you have a hunch that your favorite poet often writes about hunger. With conceptual analysis you can determine how many times words such as “hunger,” “hungry,” “famished,” or “starving” appear in a volume of poems. In contrast, relational analysis goes one step further by examining the relationships among concepts in a text. Returning to the “hunger” example, with relational analysis, you could identify what other words or phrases “hunger” or “famished” appear next to and then determine what different meanings emerge as a result of these groupings.

Perhaps due to the fact that content analysis can be applied to examine any piece of writing or occurrence of recorded communication, content analysis is currently used in a dizzying array of fields, ranging from marketing and media studies, to literature and rhetoric, ethnography and cultural studies, gender and age issues, sociology and political science,
psychology and cognitive science, and many other fields of inquiry. Additionally, content analysis reflects a close relationship with socio- and psycho-linguistics, playing an integral role in the development of artificial intelligence. The following list (adapted from Berelson, 1952) offers more possibilities for the uses of content analysis:

- Reveal international differences in communication content
- Detect the existence of propaganda
- Identify the intentions, focus or communication trends of an individual, group or institution
- Describe attitudinal and behavioral responses to communications
- Determine psychological or emotional state of persons or groups

*Quantitative Analysis*

Quantitative analysis is the scientific approach to decision making. Whims, emotions, and guesswork are not of the quantitative analysis approach. This approach starts with numerical information. Similar to raw material for a factory, the information are manipulated or processed into the results valuable for decision making. This processing and manipulating of raw data into meaningful information is the heart of quantitative analysis. Computers have been instrumental in the increasing use of quantitative analysis. [4]

After acquiring numerical information to solve a particular problem, one can always develop a model (usually mathematical model) in order to better understand the message which the numerical information is conveying. Within the range of the numerical information, prediction can also be made. If model is not to be constructed, one can still analyze the trend of the numerical information by grouping them into different categories.

Returning to the example of the Ancient Egypt project, the student may find some numerical information of life expectancy as follows (Table 1),

<table>
<thead>
<tr>
<th>age</th>
<th>female</th>
<th>male</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-4</td>
<td>193</td>
<td>193</td>
</tr>
<tr>
<td>5-9</td>
<td>45</td>
<td>53</td>
</tr>
<tr>
<td>10-14</td>
<td>22</td>
<td>29</td>
</tr>
<tr>
<td>15-19</td>
<td>28</td>
<td>29</td>
</tr>
<tr>
<td>20-29</td>
<td>69</td>
<td>49</td>
</tr>
<tr>
<td>30-39</td>
<td>73</td>
<td>56</td>
</tr>
</tbody>
</table>
By means of computer software such as SPSS or Excel, the following result will be obtained. (Fig. 2)

Although the above example is only a trivial example of the use of quantitative analysis, it demonstrated how quantitative analysis brought deeper understanding to our tasks if applied appropriately. Popular computer software application programs dedicated for quantitative analysis include SPSS, SAS, Microsoft Excel, Minitab, Maple, MathLab or Lisrel.
Objective 3: Information Evaluation Based on the Context of the Source of Information

Objective 4: Multiple Criteria Considerations for Decision Making in Information Evaluation

Learning how to evaluate what one see, read, and hear is the key to being information literate. Often student's evaluation skills are the “weakest link”. The first step in evaluation is deciding on the basis of evaluation. One’s evaluation criteria should change depending on one’s needs, but usually include reliability, validity, accuracy, authority, timeliness, and point of view or bias.

Evaluating something one find on the Internet is not much different than evaluating information in books, magazines and other sources. The main difference is that print material has usually been vetted by an editor, but internet information may not have been done so.

Another important concern is one’s own information need. For example, if I was using the internet to find a recipe for pesto sauce, I would evaluate the information differently than if I was wondering if I should go for laser eye surgery. One of the best articles about evaluating web resources is written by Robert Harris on “Evaluating Internet Research Sources”. He presents an evaluation model called CARS for Credibility, Accuracy, Reasonableness and Support.

A way in approaching any information (especially internet resources) is by asking the following questions.

Why is the information being provided?

There are many motivations for publishing on the Web or writing a book or an article: to attract users and therefore advertisers, to sell something, to enhance a service, to advocate a personal position or to disseminate research.

Specifically, web pages can be divided into the following groups (as with all classifications, sometimes sites will be a combination of more than one type):

- Advocacy Web Pages
- Business / Marketing Pages
- Information Web Pages
- News / Journals
- Personal Home Pages
- Entertainment
In fact, one can tell a little about the motivation and source of a document by the address. Common domain name suffixes are: .com (for a commercial enterprise); .gov (for government for U.S. only); .org (for organization). Hong Kong domain names often end in .hk, but do not necessarily have to end in .hk.

*Multiple Evaluation Criteria*

The following few points enlisted a certain evaluation criteria with suggested questions associated.

**Credibility**

- What is the author's training or experience with the topic?
- Is the author a recognized authority on the topic of the document?
- Is the author affiliated with an educational institution, research laboratory, governmental agency, or other reputable organization related to the topic of the document?

**Interactivity**

One thing that separates the Internet from traditional resources is the Internet's amazing ability to create communities.

- Does the site allow open comments and feedback?
- Can you view and comment on what other people are saying?

**Relevance and Scope of Content**

The quality of the information within a document is related to one’s needs.

- Is the information sufficiently current?
- Does the document provide any new information on the topic?
- Are there obvious gaps or omissions in the coverage of the topic?

**Content Validity**

- Is the methodology used to develop the resource described and appropriate to the content?
- Has the document been linked to or referenced by recognized authorities? (for example, does the document show up on more than one list of resources on the topic?)
One can check who has linked to a site by using search engine like Google or AltaVista.

- Can one find a review of the document in quality web directories such as the Scout Report Signpost, or Librarian's Index to the Internet?
- Is the document a primary (original, unfiltered material) or secondary (modified, selected, or rearranged information about primary materials) source?
- Does the information provided contradict or confirm information from other sources?
- Does the author provide references or links to confirm the accuracy of the information?
- Does the author provide verifiable statistics to support conclusions?

Accuracy and Balance of Content

- Are all sides of controversial issues presented, or is it necessary to seek alternative views?
- If the document deals with controversial issues, is the bias of the author clearly identified?
- Are there indications of careless or hasty preparation, such as spelling or grammatical errors?

Quality of the Links (specific to the Web resources)

- Are the links visible and understandable?
- Are links annotated?
- Are links provided primarily to resources?
- How reliable are the links?
References


[3] Colorado State University. Writing@CSU: Writing Guide. http://writing.colostate.edu/references/research/content/index.cfm


