

**Linda F. Burghardt**  
**Long Island University, Brookville, N.Y.**

## **An Epistemological Analysis of Holocaust Survivor Transcripts**

**Abstract:** The power of depth indexing for access to complex resources is apparent, yet tools for depth indexing are not well-developed. This study uses term generation, concept extraction, and facet identification derived through latent coding, and with each facet epistemologically defined a knowledge organization system is created that enables depth indexing of the archives.

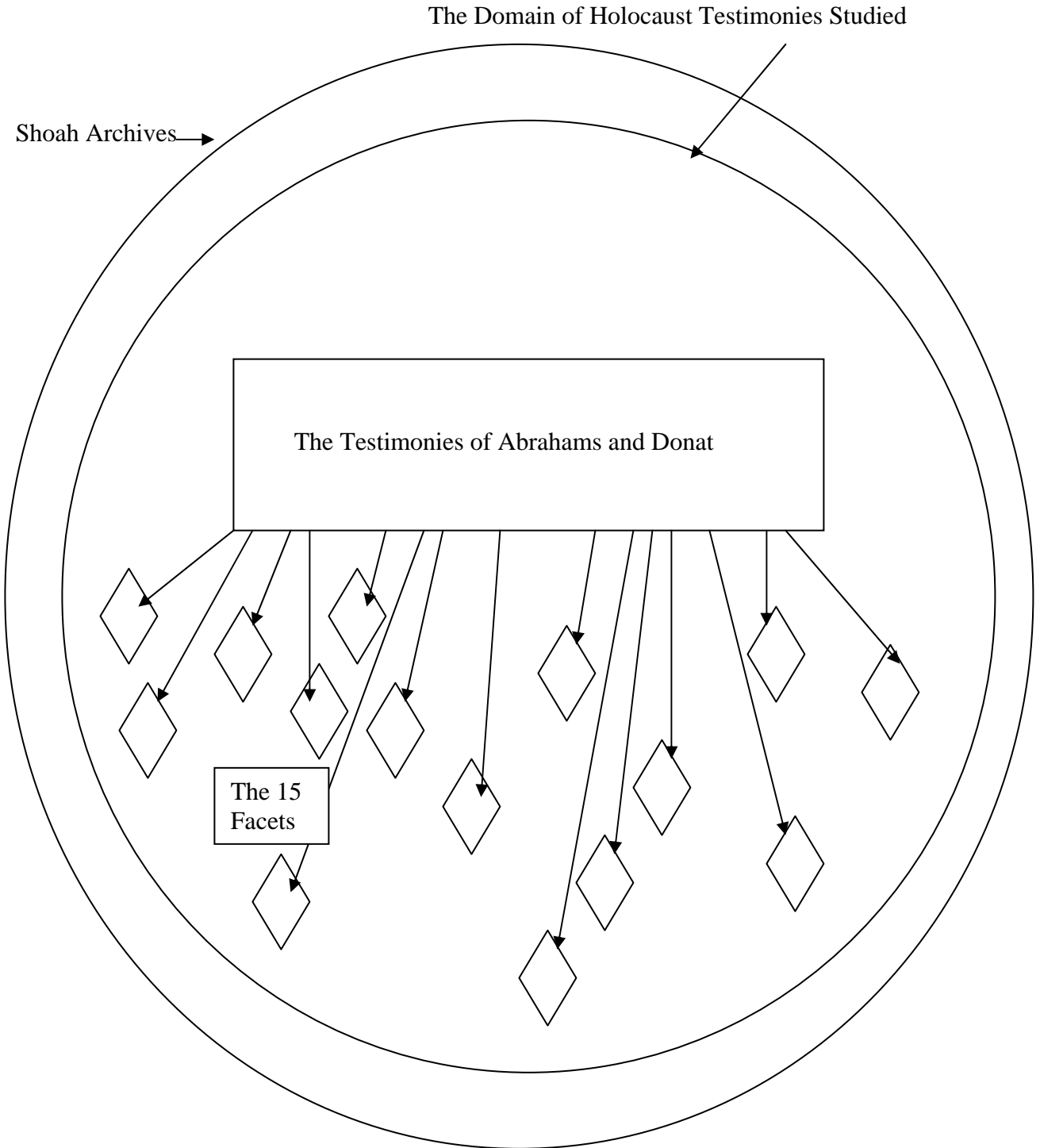
**Résumé :** La puissance de l'indexation exhaustive aux fins d'accès à des ressources complexes est évidente, mais les outils d'indexation exhaustive ne sont pas suffisamment développés. Cette étude fait appel à la génération de termes, à l'extraction de concepts et à l'identification de facettes issues d'un codage d'arrière-plan. Pour chaque facette définie de façon épistémologique, un système d'organisation de la connaissance est créé permettant l'indexation exhaustive de l'archive.

The genesis of this research study flows from the simple observation that while the power of depth indexing for access to complex resources is apparent, tools to perform depth indexing are not well-developed. Indexing for description of archival collections demands an appropriate knowledge organization system with an epistemological basis. One approach often suggested (LaBarre 2007a, 2007b, Broughton 2006) is the development of facets based on content of the archival resources.

The CAIS 2010 call for papers highlights the society's interest in "Synergy Through Diversity," with a "focus on innovative research and on information science as an evolving field." In addition, the CFP names Information Organization as a conference theme, with Cataloguing and Classification as one component. The content of this research paper focuses on these conference goals in three significant ways. First, it directly addresses the Cataloguing and Classification component by developing a method to create a faceted classification based on information gained in the process of performing an epistemological analysis, one that produces an appropriate knowledge organization system with a sound epistemological basis. Second, the paper supports and reinforces the conference's main focus, Information Science as an Evolving Field, by showing how research that builds upon previous work extends and expands the limits of the discipline by teasing emergent concepts out of old methods and gradually developing them into new, evolving applications. Finally, this project builds on the values inherent in the conference title, "Information Science: Synergy Through Diversity," by creating a system that harnesses the strengths of synergy through the joint action of the individual facets uncovered in the study, facets which when taken together increase each other's effectiveness, producing a total effect that is greater than the sum of the individual elements, in true synergistic fashion.

This study uses term generation and facet identification derived through latent coding from two different survivor testimony transcripts from the Shoah Archives, a large, elaborate Holocaust collection. Concept-extraction leads heuristically to a faceted indexing array, with each distinct facet epistemologically defined (Hjørland 1997), as demonstrated in Diagram 1, below.

**Diagram 1: A Faceted KOS for the Shoah Archives**



The results show that an epistemological analysis of latent coding yields facets that enable effective depth indexing of the studied section of the archival resources, making possible an increase in the sophisticated analysis of this digital archives.

As technology has developed to the point where large digital multimedia archives can be offered as on-line video streams, the issues of creation, storage, management and distribution of these collections has become a monumental logistical challenge (Spark 2003). Multimedia archives, like other collections of materials, rely on indexing to serve as the prime navigational device. Epistemological analysis is more complicated, of course, than flat term indexing, which results only in keywords, but it yields far better results. Because facets are sets of terms that are united by sharing similar characteristics, they are named for the totality of the contents, which adds an important and useful additional layer to the description of the domain. The utility of this additional layer is seen in the improved indexing that is created for the specific narratives within the domain. Analyzing them epistemologically enables the relationship between these facets to be shown, which provides even more information for the searcher. This occurs because through applying Hjørland’s (1997) epistemological divisions to the facets, their social and cultural activity becomes clear.

In this study, latent coding techniques teased out the fifteen facets that together defined the domain of the first-person Holocaust narratives that were analyzed, narratives that told the survival stories. The facets – the epistemologically-generated perceptual component parts of the domain – were generated by thinking of them functionally: for example, as actors, their activities, the locations of these activities, the dates on which they took place, the emotions that were experienced. The latent coding resulted in terms extracted from the narratives that organized themselves around axes exhibiting strong emotional components. The facet identification, then, named the particular groups around which terms clustered in a specific text. How the first rough set of axes were organized is demonstrated in Table 1, below.

**Table 1: The First Set of Axes**

|                 |             |                    |
|-----------------|-------------|--------------------|
| Danger          | Work        | Places             |
| Emotion         | Money       | Jewish identity    |
| Power           | Escape      | Sustenance         |
| Actors          | Papers      | The Germans        |
| Dates and times | Information | Sports and culture |

For the narrators of the testimonies, escape was the primary means of survival, and each of the narrations revolved around the series of events that comprised this action, enabled it, or acted as obstacles to it. Thus the simple terms that acted as the gathering points around which the content elements clustered became the fifteen final axes as their names were fleshed out to better represent the contents, the nuances contained within and the emotional intensity that gave them both shadow and light, as represented in Table 2, below.

**Table 2: The Fifteen Final, Named Axes**

|   |                                      |
|---|--------------------------------------|
| Intimations of Danger                     | The Need for Papers                  |
| Signs of Emotion                          | The Search for Information           |
| The Pursuit of Power                      | Locations: Geopolitical Designations |
| Actors: Friends, Family, Helpers, Enemies | Jewish Identity                      |
| Dates and Times                           | Food and Drink                       |
| Occupations and the Opportunity to Work   | Actions of the Germans               |
| The Role of Money                         | Recreation                           |
| Actions: Attempts to Hide or Escape       |                                      |

Epistemology was critical in organizing this information because it explains the framework Hjørland (1997) created to understand how knowledge is acquired and the foundation that supplies a theoretical base for action. His socio-cognitive theory of knowledge and the epistemological framework for information science that grew out of this theory seeks knowledge as a product of history, culture, and social interaction. This activity theory studies collective knowledge structures such as the Holocaust narratives, and the framework is based on these knowledge domains and the role they play in society. Because knowledge is seen as a social or cultural process, the uses of knowledge within the context of human activity are critical. This is why performing an epistemological analysis – sorting through materials and placing them into categories, with terms being grouped according to their content (Hjørland 1997) – is so useful for increasing the depth and sophistication of the Holocaust narrative indexing. The epistemological analysis created a way to sort through the facets and place them into categories based on how the domain knows what they are and how they came to be, whether through empiricism, historicism, rationalism, or pragmatism (Hjørland 1997). These approaches, or paradigms, refer to the ways in which the knowledge represented by the terms becomes known, as in Table 3, below.

**Table 3: Epistemological Analysis Results: Hjørland’s four traditional epistemological positions and the facets that belong in each**

|   |   |
|---|---|
| <b>Empiricism</b>                         | <b>Rationalism</b>                      |
| Actors: Friends, Family, Helpers, Enemies | Signs of Emotions                       |
| Actions: Attempts to Hide or Escape       | Intimations of Danger                   |
| Actions of the Germans                    | Jewish Identity                         |
|   |   |
| <b>Historicism</b>                        | <b>Pragmatism</b>                       |
| Locations: Geopolitical Designations      | The Role of Money                       |
| Dates and Times                           | Occupations and the Opportunity to Work |
|   | Food and Drink                          |
|   | Recreation                              |
|   | The Need for Papers                     |
|   | The Search for Information              |
|   | The Pursuit of Power                    |

For example, those grouped under empiricism were derived through observation and are purely objective; those under rationalism were found through human-centric thought; those in the historicism category are tied to past events; the pragmatic approach looks at what works. Individual facets literally represent a different point of view; together they create a knowledge base of a domain. In this way an appropriate knowledge organization system for depth indexing of the archives was created.

This paper adds to the value of the visual history domain by proposing a way to make the testimony of witnesses to history more accessible and thus significantly increasing their worth. It would enable researchers to employ an enhanced indexing system that lets them locate, by forwarding to a specific place on the tape, the particular information they are seeking. Often the next step after creating facets and performing an epistemological analysis is the creation of a faceted classification based on information gained in the process, as the facets together form a class or a particular subject, in the meaning created by Ranganathan (1965), who developed the fully faceted approach of classification, called the Colon Classification. The successful classification scheme based on this conceptual framework would be constructed from a combination of the fifteen individual facets, but would differ from a traditional classification in that it would not necessarily follow a set pattern of structured hierarchy. In this additional way this research contributes to the conference theme of “Information Science as an Evolving Field” by providing the basis of what could be developed into a fully faceted classification that would greatly enhance the value of the historically significant Shoah Archives, the world’s largest collection of video oral histories, and also form a workable model for other large, complex resources to fulfill their purpose by making their holdings more accessible.

## References

- Broughton, V. 2006. The need for a faceted classification as the basis of all methods of information retrieval. *Aslib Proceedings* 58 (1/2): 49-72.
- Hjørland, Birger. 1997. *Information Seeking and Subject Representation: An Activity-Theoretical Approach to Information Science*. Westport, CT: Greenwood Press.
- La Barre, Kathryn. 2007(a). Faceted navigation and browsing features in new OPACs: Robust support for scholarly information seeking? *Knowledge Organization* 34 (2): 78-90.
- La Barre, Kathryn. 2007(b). The heritage of early facet classification in document reference retrieval systems, 1920-1969. *Library History* 23 (2): 129-49.
- Ranganathan, S.R. 1965. *The Colon Classification*. New Brunswick, N.J.: Rutgers University Press.
- Spark, David. 2003. Digital asset management: off the shelf or on your own. *EContent* 26(8/9): 38-42.