RELATIONSHIP ANALYSIS OF IMAGE DESCRIPTIONS: AN ONTOLOGICAL, CONTENT ANALYTIC APPROACH

by

Allen C. Benson

BFA, University of Minnesota, Minneapolis, MN, 1974

MLS, University of Alabama, Tuscaloosa, AL, 1993

Submitted to the Graduate Faculty of the

Department of Library and Information Science

School of Information Sciences

University of Pittsburgh

in partial fulfillment of the requirements

for the degree of Doctor of Philosophy

University of Pittsburgh

2011

UNIVERSITY OF PITTSBURGH

SCHOOL OF INFORMATION SCEINCES

This dissertation was presented

by

Allen C. Benson

It was defended on

June 23, 2011

and approved by

Geoffrey C. Bowker, PhD, Professor and Senior Scholar in Cyberscholarship, Program in Library and Information Science

Daqing He, PhD, Professor, Program in Library and Information Science

Scott E. Fahlman, PhD, Research Professor, Language Technologies Institute and Computer Science Department, Carnegie Mellon University

Dissertation Advisor: Christinger Tomer, PhD, Professor, Program in Library and Information Science

Copyright © by Allen C. Benson

2011

ABSTRACT

RELATIONSHIP ANALYSIS OF IMAGE DESCRIPTIONS: AN ONTOLOGICAL, CONTENT ANALYTIC APPROACH

Allen C. Benson

University of Pittsburgh, 2011

The relationships humans express when describing images have powerful, but poorly understood, effects on how visual information is represented, structured, and processed in information systems. This study evaluates the benefits and difficulties of using content analysis and ontological analysis as predictors of relationship instances and types occurring in image descriptions. A random sample of 36 documented reference transactions from the administrative files of the Pittsburgh Photographic Library is analyzed in light of three describing contexts: image searcher, curator, and cataloger.

Through the qualitative and quantitative assessment of image descriptions, the research leads to several key findings and contributions. The most important findings vindicate the claim that recognition, capture, and classification of relationship instances can be empirically grounded utilizing content analysis and ontological tools and methods. Evidence comes in successfully ascertaining and capturing in a Corpus the existence of 1,655 relationship instances. Further, the analysis finds evidence of relationship types and subtypes of relationships whose members share certain recognizable properties in common. The study also brings useful, new insights to the capture of background information surrounding events using *situation-templates*, introduces methods for formulating case relations and image attributes as binary predicates, and it offers a new, finer-grained definition of relationship. Contributions of this study include a corpus of

relationship instances, an ontology of relationship types, and a methodological framework that provides significantly better results than earlier studies in the prediction of relationships (the architecture of which is depicted in Figure 22 on page 101).

There are a number of ways this research could be extended and corroborated. First, event analysis ought to be tied to a system of semantic frame analysis. Second, test the content analysis form against other texts, which will result in elaboration of the core ontology of relationship types. Third, expand image description analysis beyond the current domain to include image description in visual ethnography, art history and criticism, and photography practices. Fourth, test how inference engines reason over relationships in knowledge-based environments. Finally, to aid in the analysis of the meanings of relationships, more work is needed in formalizing the ontological concepts used in image descriptions.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I express sincere appreciation to my advisor, Dr. Chris Tomer, for his solid guidance, his unflinching support throughout this entire process, and for being my mentor. His high expectations and continuous questioning of my assumptions helped me reach my full potential. Dr. Tomer was incredibly generous with his time and provided me with a wealth of ideas that assisted me in narrowing my focus and refining my approach to research. Most importantly, he gave me considerable freedom to pursue the topic that was intellectually challenging for me.

I would also like to acknowledge and thank my committee members Dr. Geof Bowker, Dr. Daqing He, and Dr. Scott Fahlman. Each time I left Dr. Bowker's office, I came away inspired and with a fistful of notes and citations for readings in philosophy, sociology, history, and ethnography. Dr. He shared with me his valuable insights into information retrieval and the research process and led me to pursue a preliminary investigation, which helped me sort out countless problems in the content analytic approach before beginning the main body of this study. I express my thanks to Dr. Fahlman for inviting me to join his Scone knowledge base research group at Carnegie Mellon University, for answering countless questions in email, for meeting with me on almost a weekly basis throughout my entire doctoral program, and for being my mentor. His enthusiasm and willingness to share a variety of ideas regarding problems in natural language processing and knowledge representation made much of this thesis possible.

I would like to thank my friends for their companionship and unwavering support, especially Marty Tiani, Alexis, Jon, and Isolde Rittenberger, and David Kraeuter. Thanks to Dr. Kenrick Thompson, a valuable and trusted friend, who warned me about the pitfalls, rewards, and hard times that would surely accompany my pursuit of a PhD. I am indebted to him for his constant encouragement during my doctoral studies, for countless conversations in person and in letters, advising and directing me along the way.

From the Carnegie Library of Pittsburgh, I would like extend my gratitude to Gilbert Pietrzak, Curator, Pittsburgh Photographic Library, for providing me with access to numerous photographic resources throughout my doctoral studies. I further want to thank Dr. Wilfried Sieg, mathematical logician, Department of Philosophy, Carnegie Mellon University, for providing invaluable insight into how photographs can be viewed in relationship to subject. Similarly I would like to thank Dr. Ray Jackendoff, Philosophy Department, Tufts University, for sharing his insight into the layers of meaning in language describing pictures. I also want to thank Dr. Ernest Abadal, Universitat de Barcelona, Facultat de Biblioteconomia i Documentació, for his support and invitation to introduce my research to the Library and Information Science students and faculty at the University of Barcelona.

Above all, I am very grateful to the people to whom I owe the most – my family. They have supported me throughout my PhD program. My parents, Carl and Kathryn, who have both passed away, were with me every day nonetheless. They taught me to enjoy long days of hard work as I grew up on a farm in the Midwest. It seems to me they both viewed work as a celebration of doing what one likes best in life. Most importantly, they inspired me long ago to follow my passions, which I did in this endeavor. I miss them both very much and wish they could be here now to celebrate in my accomplishment.

Thanks to my nieces Amy, Lilly, Lori, and Nelly, for their encouragement to take on this challenge, and to my nephew Eric, who is now deceased, but forever on my mind. I am also grateful to my brother Gallen, for his insight into the research process, encouraging me to pursue it as a creative endeavor much like his own creative work as an artist and jewelry designer. Thanks to my sister Susie for long phone conversations, help with language translations, always asking me for updates on my progress, and congratulating me for my successes along the way. Without my brother and sister's unconditional support, I would not have been prepared to take on this challenge.

I cannot end without thanking my friend Cri. We met at the beginning of this journey, both in transition as students in Prague, trying to decide what adventures in life to pursue next. I am greatly indebted to Cri for her engaging insights into life and language—sometimes quirky, always brilliant.

1.0 INTRODUCTION

This dissertation examines instances of relationships that occur naturally in the discourse surrounding photograph archives. More specifically, the investigation focuses on what humans communicate about images during reference transactions in the Pittsburgh Photographic Library during the years 1963-2009. What is of primary interest here are the types of relationships that emerge when searchers, curators, and catalogers describe the image content of photographs and the extent to which these expressions convey visual information through linguistic means.

What is novel in this approach is the application of content analysis and ontological tools and methods as the framework for analyzing and structuring the data collected in this study. The result is a semantically rich set of relations and concepts recorded in a Corpus of Relationship instances and an ontology of relationship types capable of supporting hierarchically structured knowledge organization systems more specific than traditional controlled vocabularies. Engaging ontology's role in the structure and representation of relationships and their instances leads to information systems capable of cross-related and intersecting data searching, greater knowledge sharing, reuse, and interoperability among disparate systems. Through multiple, hierarchical inheritance and default reasoning with exceptions—processes made possible in computational systems by these more formal structures—the universe of knowledge can expand beyond what is stated explicitly in image descriptions.

The more traditional approach of only using image attribute structures and categories of image attributes for description and classification in Library and Information Science (LIS) is an old one, but capturing and attempting to classify the *types of relationships* that humans express when describing images and doing so in a semi-formal, systematic manner using ontological tools and methods has only recently been considered. Research in this area is therefore needed and is concerned with three problems, namely

(1) What kinds of relationships do humans express when they describe image content in photographs?

(2) What forms do these relationships take?

(3) Can the relationship types and their instantiations be captured and classified?

Examining the whole question of relationships as they are naturally expressed during reference transactions, however, requires that the boundaries between image content and language be viewed with a fresh eye. The goal is to take an investigative path that extends the body of literature concerned with experience and expertise in image searching and view image description as a social process. First, there is the searcher who presents a query describing known, unknown, and possibly non-existent images. Second, there is the curator who plays the role of mediator, describing and interpreting image content during the process of mediation. Finally, there is the cataloger whose role is to describe photographs in catalog records, which are consulted during search and retrieval by the curator and image searcher. It is with regard to this constellation of searcher, curator, and cataloger that the current study is carried out—a study that attempts to determine to what extent semantic relationships exist when humans describe image content.

The dissertation is structured around six chapters. The remainder of this introductory chapter provides an overview of the topic and explains why relationships in image descriptions are important. Next, there is a description of the nature and scope of the problem being investigated followed by assumptions, limitations, and significance of the study. Chapter two presents a review of the literature attempting to determine to what extent researchers' understanding of image behavior in fact informs relationship theory. Chapter three presents research questions and introduces the methodology and design of the study. Chapter four is devoted to an in-depth analysis of the data sample and brings to light the interesting instances and classes of relationships that emerge from the Corpus. It is interesting for another reason, however, because it shows what the problems in content analysis of image research discourse in general are and how they might be approached. Chapter five is concerned with modeling an ontology of relationships that emerges from the data analysis. Chapter six presents a summary and conclusion of the dissertation, discusses the study's limitations, and makes recommendations for future research.

1.1 OVERVIEW

It is important to begin by clarifying the terms *visual information*, *image description*, and *relationships. Visual information* is taken to mean the bits of data that make up the content of images. This dissertation is ontologically committed to occurrences of photographs, so the visual information of concern is the image content of photographs. *Image description* is intended to mean linguistic expressions that describe the meaning of images. The expressions may entail words describing image objects (picture elements) visually represented in the picture as well as mental constructions, such as angle of view, aspects of light and shadow, references to pre-photographic referents that exist outside the frame of the photograph, and relationships among these entities. The researcher takes as a reasonable working assumption that meaning is not embedded in the image content of photographs. Rather, viewers who interpret visual information bring meaning to the image. Furthermore, the meanings humans perceive appear to utilize Gestalt processes. That is, they are the result of an interaction between environmental input and activities of the brain that influence that input (Jackendoff, 1983).

This study adopts Rebecca Green's definition of *relationship*, which she states is an association between two or more entities or classes of entities (Green 2001, p. 3). There are some difficulties with this definition, however, that need refinement. When the word "relationship" is used in the domain of images it is not always clear what is meant. For one thing, it is not clear what entities are being associated with one another and what set of conditions is being placed on a relationship's meaning. The categories of relationships in current library indexing systems are few in number and overloaded with meaning, or, rather, mean too little.

Wittgenstein states there are some words that have a precise, strict meaning, but for the most part "the meaning of a word is its use in the language" (Wittgenstein, 2001, §43). Like tools, words have many possible uses. The word "relationship," it can be argued, is a word with more than one sense. Wittgenstein further suggests that words derive their strength from a series of crisscrossing, overlapping relationships and similarities. He calls this relationship a family resemblance relationship. Just as a family consists of several different members with crisscrossing and overlapping similarities, so it is with words. Saussure (1966) states similar views, that the manner in which humans understand and represent the world is formed by the language they use and that the meaning of words is established by how words relate to each other. During the course of this research one goal is to refine Green's definition of *relationship* by examining the everyday use of relationships in the language describing photographic images.

The relationship problem is not new to Library and Information Science and expressing relationships when describing photographic images has been a central aspect of the discourse surrounding photography since its early beginnings, yet relationships are rarely examined as phenomena in need of explanation. To begin answering what a relationship is in textual descriptions of photographs, consider *The Pencil of Nature*, originally published in London as six separate fascicles between 1844-1846. This is the first book illustrated with photographs, or *photogenic drawings* as their creator William Henry Fox Talbot called them (Talbot, 1969). Talbot accompanies the first photograph, shown in Figure 1, with the following textual description:

> This building presents on its surface the most evident marks of the injuries of time and weather, in the abraded state of the stone, which probably was of a bad quality originally. The view is taken from the other side of the High Street—looking North. The time is morning. In the distance is seen at the end of a narrow street the Church of St. Peter's in the East, said to be the most ancient church in Oxford, and built during the Saxon era. This street, shortly after passing the church, turns to the left and leads to New College.
>
> --William Henry Fox Talbot

The first relationship Talbot describes is spatial, making reference to the location of marks on the surface of the stone. Talbot then describes a *point of view* or location of a hypothetical observer of the scene when he states, "taken from the other side of the High Street—looking North." He expresses temporal relations between the photograph and the time it was taken—morning—and a time interval spanning the Saxon era. There is a set of spatial expressions that place the Church of St. Peter “at the end of a narrow street,” “in the distance,” and “in Oxford.” Finally, Talbot describes motion as a spatial path, “passing the church” and then turning to the left.

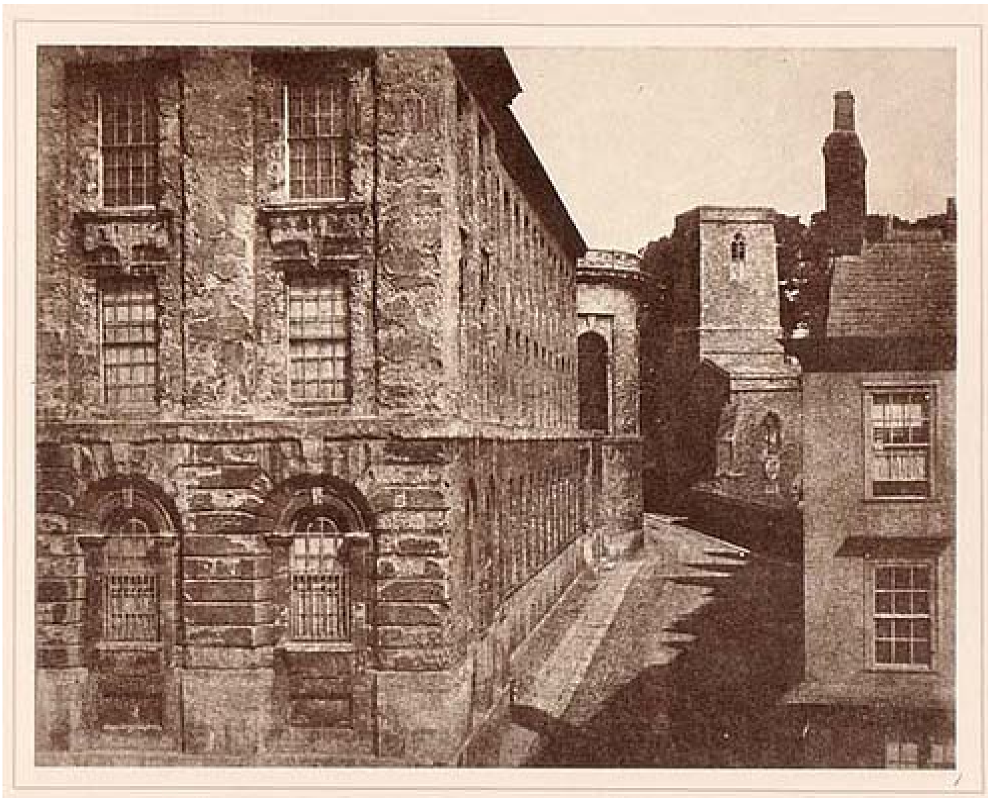


Figure 1: Henry Fox Talbot. Part of Queen's College, Oxford. [The Pencil of Nature, Part 1, pl. 1] n.d. Taken from the reproductions in Larry J. Schaaf, H. Fox Talbot's The Pencil of Nature; Anniversary Facsimile (New York: Hans P. Kraus, Jr. Inc., 1989). Used with permission.

Some relations are explicitly expressed in language, while others are implicit, existing in the mind of the viewer describing the image. Both contribute to the overall understanding and interpretation a viewer brings to the image. For instance, it is implicitly understood that when Talbot writes about the surface of the building and the marks left on stones he is talking about the outside of the building. Even though Talbot does not specifically mention the outside surface, viewers of this image understand his description to mean the outside surface of the stones. Other relations may go unnoticed unless a viewer makes them explicit. For example, an architectural historian might draw attention to the six-over-six double-hung sash windows in the building positioned on the right side of the print. Little is known about these possible alternative descriptions derived from background knowledge and inference. The aim of the current study is to isolate semantic relationships as they are expressed in image descriptions, both explicitly and implicitly, and to examine the contributions they make in conveying information about image content.

Thus, in order to describe image content in photographs, there must be an understanding of how entities and relationships between them are encoded in images and how people decode these and re-present and manipulate relationships in language. This dissertation is concerned with the latter. To be successful in this investigation, the representation of relationships is based on image descriptions—that is, the linguistic expressions of image searchers, catalogers describing image content, and curators mediating between these two groups. With at least minimal correspondence between the photographs that are saved in archives and the language of image descriptions as they occur naturally in the search and description processes, the expectations are that the results of this research will in some small way help the efforts to improve knowledge organization systems that assist researchers in locating and retrieving, with some degree of accuracy, the images they seek within large collections of images.

1.2 THE PROBLEM

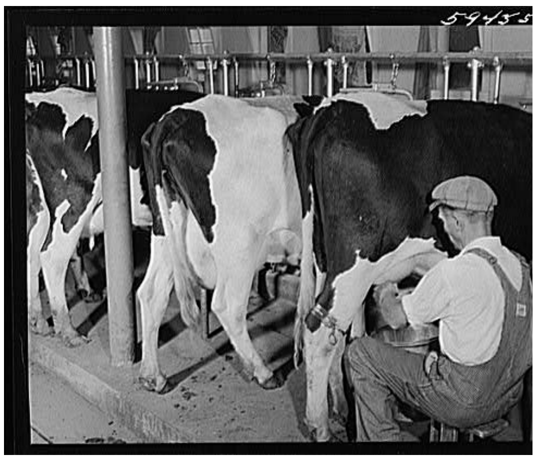
From the perspective of Library and Information Science (LIS), a common-sense knowledge about relationships is required for many practical tasks including viewing and interpreting photographs, describing photographs, modeling controlled vocabularies, responding to searchers' queries, and assigning subjects headings to photographs. Given the importance of relationships in image search and retrieval systems and a system's ability to locate relevant images, the problem of semantic relationships and how they are represented in controlled vocabularies and understood in search processes remains a difficult and time-consuming problem.

Existing approaches to image behavior research are primarily concerned with how image attributes are categorized and described. Collins (1998) expresses it succinctly when she concludes her study of image user queries saying, "Any amount of subject indexing, even of only the main subjects of photographs, could only improve access" (p. 53). This often-held position is based on the following assumption:

> *Image searchers describe images in terms of attributes and categories of attributes, so the more attributes that are identified and indexed, the better chance an image searcher has of finding and retrieving relevant images.*

The following example helps illustrate the problem that results when relationships that exist among image attributes—critical entities that integrate concepts into coherent wholes—are not also taken into account when designing information systems. The images presented in Figure 2 represent three photographs retrieved from a search conducted using the Library of Congress (LOC) picture catalog on the phrase "cows being milked in a barn."

Wolcott, Marion Post. (1941). *Cows by the windmill and water hole being driven to the barn to be milked. Two Rivers Non-Stock Cooperative, a FSA (Farm Security Administration) co-op at Waterloo, Nebraska.* (Library of Congress, Prints & Photographs Division, FSA/OWI Collection, [LC-USF34-059616-D (b&w film neg.)]



Wolcott, Marion Post. (1941). *Cows being milked in dairy barn at Two River Non-Stock, Cooperative, a FSA (Farm Security Administration) co-op at Waterloo, Nebraska.* (Library of Congress, Prints & Photographs Division, FSA/OWI Collection, [LC-USF34-059455-D (b&w film neg.)]

Wolcott, Marion Post. (1941). *Cows entering the barn to be milked. Two River Non-Stock Cooperative, FSA (Farm Security Administration) co-op, Waterloo, Nebraska.* (Library of Congress, Prints & Photographs Division, FSA/OWI Collection, [LC-USF34-059587-D (b&w film neg.)]

Figure 2: Sample of photographs and their captions retrieved from a search of the Library of Congress picture collection on the phrase, “cows being milked in a barn”

Two kinds of relationships are expressed in this query: (1) a case relationship that involves an agent—a farmer perhaps—an action "milking," and recipients of that action—cows, and (2) a spatial relationship expressed in the prepositional phrase "in a barn," which places one locative entity *cow* inside another spatial reference *barn*.[1]

The relevance of these photographs is determined by matching query terms with indexed terms describing picture elements identified in the photographs. It is important to note that locative entities (cows, windmill, water hole, and so on) are pictured in the photographs, but it is within the caption where the mental constructs linking these entities together are expressed as spatial relationships as shown in Table 1.

Table 1: Spatial relationships expressed in the Library of Congress picture captions

Spatial expression	Relationship
cows *by* the windmill	by
[cows by] the water hole	By
cows … *in* dairy barn	In
dairy barn *at* Two River Non-Stock, Cooperative	At
co-op at Waterloo	At

Jörgensen (2001) discovered in her studies that image users in fact describe content using expressions in the form of short narratives. Alicia Tribble's grammatical analysis on 80,000 phrases extracted from the Phetch data set substantiates this (Tribble, 2010, p. 45). Tribble

[1] The locative relationship in this example is based on definitions provided by Herskovits (1986) and analysis of the case relationship is based on Chaffin & Herrmann (1984).

discovered that contributors to the Phetch data set prefer describing images using short phrases consisting of two or more words and include dependency patterns such as "forest in background," "facing towards photographer," "he has short hair," and "tub has white rim" (p. 48).

Despite the prevalence of grammatical relationships and the relationships expressed in catalogers' descriptions and visual researchers' queries (Green, Bean & Myaeng, 2002; Bean & Green, 2001); despite the different disciplines that have made a study of relationships and their representations in language (Casagrande & Hale, 1967; Fahlman, 1979; Jackendoff, 1985; Herskovits, 1986); and despite the frequent acknowledgements that there is a need to closely examine relationships in the field of LIS in general and image behavior research specifically (Jörgensen, 2003, p. 261; Svenonius, 2000, p. 161); there remains a persistent lack of coherence and understanding about what relationships express in image content.

Care has been taken to frame the current study as an investigation of the semantic relationship problem, not a search and retrieval problem. While a searcher may find images s/he wants searching the Library of Congress catalog on the statement "cows being milked in a barn," the fact remains that there are many other images included in the results set that are not pictures of cows being milked in a barn; more importantly, the search engine is not concerned with interoperability and reasoning over semantic relations in unstructured data on large scales. Rather, the engine is engaging with a relatively small universe of database records. Thus, while controlled vocabularies and search and indexing systems are useful in locating and retrieving images containing named attributes in bibliographic records, they have several drawbacks in unstructured, natural language contexts. Relationships, therefore, constitute a problematic field for intellectual study in LIS.

The problem of categorizing and describing image descriptions goes much deeper than this. As a practical matter, there is general agreement that there are objects in the world and that objects exist in various relationships with one another. A photograph, the physical object, is a three-dimensional thing with a front, back, and edge. Its physical properties are definable in terms of length and width. Other properties that are easily categorized include surface type (e.g., glossy or matt), material base (e.g., fiber paper or Resin Coated), and light-sensitive chemistry (e.g., cyanotype, salt print, gelatin silver, and so on). These are recognizable properties of the physical manifestation of photographs and an important part of the discourse surrounding photography.

What is contained in the picture itself, however, presents a number of problems. The surface on which the image content lies is a geometric plane, yet humans use words and phrases to communicate information about picture elements lying on this flat surface as projections into the four dimensions of space and time. Viewers describe three-dimensional objects having relationships with other three-dimensional objects both within and outside of the boundaries of the physical photograph. The longstanding view in LIS believes that the structure of information describing this spatial environment can be organized into objectively defined categories, much like the physical characteristics mentioned earlier such as length, width, and material base. The argument made here, however, is that this is a highly problematic way of viewing image content.

Herskovits (1986) touched on this point when she was investigating linguistic expressions referring to objects in the spatial domain. She argued that maintaining an objectivist viewpoint of the spatial domain is in fact incomprehensible because locative expressions often refer to entities that are mental constructions, not concrete objects. Her argument can be extended to the spatial environment of photographic images. The language used for describing relationships in this space does not always reflect the physical nature of the image. Herskovits maintains that in general, words mediate between picture elements and referents. This study aims to further refine how one may usefully think of image description as a medium through which visual information might be adequately conceived.

1.3 SIGNIFICANCE OF THE STUDY

The establishment of image attribute structures and categories of image attributes are incomplete if the types of relationships that exist among them are not taken into consideration. The reality of how visual searchers think about and rely on relationship types and the degree to which these are put to work and recognized by curators and automated search and retrieval systems is far from clear. Linguists have shown, for example, that a particular spatial relationship has no significance unless it can express distinctions between right-of, left-of, in, at, under, and so on. These variations in spatial relationships allow visual data to be informative. If library information systems are not capable of representing different relationship types and their important differences, then the semantic units cannot inform anything.

This dissertation is only a modest beginning to the complex problem of relationships in image descriptions. The contributions it promises do not constitute a comprehensive dictionary of semantic relationships, but it is a serious attempt to let users of photographs suggest the conditions for the meanings of the various relationship types situated in the discourse of archives. An attempt is made to systematically collect and analyze a corpus of relationship types and instances from catalogers' descriptions and the verbatim communications between curators and image searchers. These efforts result in three contributions: (1) a *Corpus of Relationships* documenting the predicted relationship instances occurring in image descriptions, (2) a coding instruction book and forms for parsing semantic relations from unstructured natural language image descriptions, and (3) an ontology of the relationship types that emerge from this analysis. The underlying motivation for this study is to determine what categories of relationships are most commonly expressed and what categories of these relationships are indicated as most appropriate to represent the image content of photographs. These categories can provide the basis for expanded semantic relationship development in indexing languages and tools for automated image retrieval systems.

Historically, catalogers have hoped that photographs and other visual objects could be understood and described in certain general bibliographic terms. More recently it has become clear that different document formats and modes of communication require separate analysis. While pursuing the current strategy defers consideration of how a universal approach can be applied to all document formats and whether there are overarching principles describing a broader range of relationships beyond those discovered in this study, the hope is that a more certain understanding of relationships in the domain of image descriptions will ultimately improve successful general analysis.

Finally, language, to some degree, serves as its own metalanguage (Preston, 2004). This becomes evident when catalogers refer to thesauri and subject heading lists to better understand and expand upon subject terms. There are occasions when the meaning of semantic relations must also be explained, but there is no systematically collected corpus of semantic relations available for examination. Here then is a semantic gap that this dissertation proposal intends to begin addressing.

1.4 CONCEPTUAL FRAMEWORK FOR UNDERSTANDING RELATIONS

An important question to address at the outset is whether the entities joined by relationships are terms, concepts in the mind, picture elements in the image content of photographs, or concrete objects. It became evident during the preliminary investigation reported in chapter three that image searchers, curators, and catalogers express relationships textually. This suggests that humans draw relationships among linguistic entities even though the photograph in most cases is atextual—that is, consisting of picture elements rather than words written or printed. There are some relationships, however, that do not exist in linguistics. There does not appear any doubt in Figure 23 that the spatial relationship in the caption, “Desk under the window,” is not locating the word “desk” under the word “window.” The statement is referring to spatial entities that are pictured in the image content of the photograph. It is taken for granted that the meaning of “desk” and “window” are fundamentally a visual point of view. This viewpoint depends on the assumption that “desk” and “window” map onto the picture elements or visual signs representing the conceptual units “desk” and “window.”

The conceptual framework adopted in this dissertation for understanding relations in visual information descriptions is built around five layers of depiction. These layers of depiction are a heuristic device to prompt analysis and categorization of relation types and are concerned with both the production of images and image descriptions.

The five layers or domains of meaning and their corresponding contexts are:

1. A *referent* is a thing (concrete object) or experience that exists in real-world space and temporality. For example, a person who poses for a portrait photographer is a *referent*. Referents exist prior to now, prior to the existence of the photographic image.

2. A *depicted referent* is a photographic image—a three-dimensional scene transferred onto a two-dimensional light-sensitive surface. As William Mitchell once described it, "it is a direct physical imprint, like a fingerprint left at the scene of a crime or lipstick traces on your collar" (Mitchell, 1992, P. 24). An analog photograph depicts the referent by means of chemical processes on the surfaces of film and photographic paper. A digital photograph relies on a Charged Coupled Device (CCD) to translate light patterns into a series of digital codes.
3. An *object showing depiction* is a physical entity—an illuminated, tangible object that displays or projects the image. It is the image carrier. For example, a photographic print is an object that shows a depicted referent. This distinction is useful for making references to the location of a picture element on the surface of a print or CRT display. For example, directing a viewer's attention to picture elements situated in the lower left-hand corner of a screen or top edge of a print.
4. The *linguistic domain*. Ritchen (1990) states that photography "constitutes a rich and variegated language, capable like other languages of subtlety, ambiguity, revelation and distortion"(p. 1). The meaning that photographic images invoke can be verbalized, written down, and recorded as textual descriptions. It is through language that humans express the structures they project onto images, for example, the notion of relationships holding among entities. The image description, therefore, is an interface between human agents and the meaning of photographic images.
5. The *conceptual* domain is a subpropositional mental representation layer (Margolis & Laurence, 1999, p.4). Concepts are mental information, ideas that have a classificatory function, gathering together things with shared properties.

1.5 ASSUMPTIONS AND LIMITATIONS

The following assumptions and limitations are recognized regarding the scope of this study and the nature of meaning in images.

1. First and foremost, this study does not aim at completeness. For one thing, the meaning of photographic images is polysemous and potentially unbounded. Furthermore, this study rejects the monosemy model sometimes adopted by catalogers. This is the notion that a single, unambiguous image description can be achieved and that it is possible for a cataloger to maintain an objective position.

2. Semantic patterns and differences in meaning among individual lexical items can differ across languages. The dissertation proposed here is concerned only with English expressions of relationships.

3. The discursive space of a photographic library as an institution and physical collection is dynamic, complex, and also unbounded. It extends beyond the physical boundaries of catalogs and library holdings to include communication and exchanges with museums, archives, school classrooms, journal articles, monographs, and online image collections. The analysis proposed in this study, however, attempts to set boundaries and focuses its attention on the data set defined in Ch. 3, § *3.2.4.1 Definitions*.

4. No attempt is made to be explicit about how concepts and relationships between concepts are mentally stored and represented or what word meanings are. There is a wide range of views on these subjects that span a wide range of disciplinary interests and a unified approach seems elusive. This problem is considered to be outside the direct concerns of LIS. In general, however, this study adopts the view put forth by M. Lynne Murphy (2003) that concepts represent knowledge of the world and since language is part of the world, these representations include language. Therefore, this investigation proceeds under the assumption that the sources of knowledge about the structure of relationships in photographic images are the same or similar to representations in linguistic expressions.

5. The semantic properties of words are not the only factors at work in judging relationships among image attributes. There are other factors at work in relating image attributes and judging semantic relations among image attributes, including context and word senses.

Jackendoff and Landau (1992) and Jackendoff (1996) argue convincingly that there are some aspects of vision that cannot be expressed linguistically, just as some aspects of language do not translate into visual elements. So additionally, it is not a concern in this study the degree to which visual representations correspond with their linguistic counterparts or which aspects of relationships in visual representation can be expressed in language and which cannot.

6. This is not a proposal to produce a new indexing language or to even extend existing thesauri such as *Art and Architecture Thesaurus (AAT)* (J. Paul Getty Trust, 2000) or *Library of Congress Thesaurus for Graphic Materials (TGM)* (Library of Congress, 1995). The procedure instead will be to examine some of the ways humans express relationships in the institutional discourse of photograph libraries—particularly the Pittsburgh Photographic Library—and examine the ways in which image searchers, curators, and catalogers make use of relationships during various describing tasks.

7. The analysis will attempt to reveal the full range of semantic relations present in the text contained in the sample data set. To some degree a preexisting awareness and understanding of the context and role of the Pittsburgh Photographic Library on the researcher's part allows for pragmatic inferences to take place. This type of inference will be made whenever possible. For example, if P-1859 is handwritten in the border of a piece of correspondence and it can be determined with confidence that P-1859 represents a photograph's accession code, then the accession code may be included in the formal representation of the associated relationships. For instance, instead of stating [1.1] in the Relationship Corpus, knowing that P-1859 represents an accession code would more likely result in two propositions [1.2] and [1.3].

[1.1] 'photograph' <u>has subject</u> 'street railway'

[1.2] 'photograph' <u>has accession code</u> "P-1859"

[1.3] "P-1859" <u>has subject</u> 'street railway'

8. Finally, excluding large, wall-sized tableau forms, the photograph can be observed from a single viewpoint.[2] In this sense, the photograph space of interest in this study is a small

---

[2] Art photography characterized by French critic Jean-François Chevrier as the "tableau form" is large in size and intended to hang on the wall and viewed like paintings. The artists Jeff Wall,

space. Naturally this definition depends on the viewer context. A photograph might be considered large-scale when viewed at high magnification or through a small hole. However, this study is concerned with people engaged in everyday viewing activities—the kinds that take place in archives, museums, galleries, and other memory institutions. The observations made available by the corpus data set (called *incidents*) engage the researcher in common sense observations of relationships between picture elements in image content expressed verbally.

1.6 CONCLUSIONS

Knowledge of relationships in the field of LIS is basic to daily activities such as organizing, storing, and retrieving photographic prints, negatives, and digital image files, classifying and describing photographs, and also reading and understanding image searchers' queries. Thus, to understand image content and make image retrieval more effective, it is important to understand how people code and decode relationships in images, how catalogers describe the implicit relationships they see in photographic images, and how curators and reference librarians navigate around relationships presented in image searchers' queries. With at least a reasonably close correspondence between image content, representations of the visual environment in language, cataloger's interpretations and descriptions, and the structuring and processing of relationships in automated information systems, it is reasonable to expect an image searcher will find more effectively and with greater accuracy the images she needs. This is especially true in an environment where digital image archives are growing at exponential rates.

Thomas Ruff, and Jean-Marc Bustamante exemplify this new style. (See Fried, M. (2008). Why photography matters as art as never before. New Haven, Conn.: Yale University Press.

The dissertation addresses the relationship problem by examining the properties of relationships expressed in English by image users searching for images. It takes into account the detailed analysis of queries made by searchers that express not only image attributes, but also relationships among attributes and relationships that emerge in the viewing context, for instance, when catalogers describe photographs.

The three contributions promised by this proposal also serve as the stopping criteria for the dissertation. These include the coding instruction book and forms designed for extracting semantic relations from unstructured natural language image descriptions, a corpus of relationship instances, and an ontology of the kinds of relations that are embodied in image descriptions.

Finally, writing about any subject is easier when there is a clear view of the underlying motivations and the endpoint toward which ideas are aimed. On the one hand, bibliographic database records, finding aids, and controlled vocabularies have had a longstanding tradition in LIS. On the other hand, researchers have nearly exhausted their options for improving image retrieval by focusing efforts on controlled vocabularies and categories of image attributes alone. Therefore, while the dissertation does not attempt computational implementation of a new kind of descriptive machinery, purposeful steps are taken towards building a corpus data set of relationship instances and types that naturally leads to the next step possible—ontology-based representation of image content in knowledge-based systems such as Scone or in information retrieval systems that are capable of reading and interpreting relationships among image attributes.

In relation to this goal, the current project can be considered an extension of the researcher's investigation of the semantic archives model which began in "The archival photograph and its meaning: Formalisms for modeling images" (Benson, 2009), and continued in "Capturing relationship in image descriptions: a preliminary investigation." (Benson, 2011a) and "Ontophoto and the role of ontology in organizing knowledge" (Benson, 2011b).

2.0 REVIEW OF THE LITERATURE

This review attempts to lay down the interesting intellectual activity evident in the literature that addresses relationships and their role in organizing information, especially visual information expressed in image descriptions. It is, in Mitchell's (1994) terms, an examination of the ways in which image content in photographs may be turned into language. It is concerned with the fundamental concept of relationships, negotiating through both the generalities and specifics as they may pertain to how people describe and represent relationships in image descriptions within the context of library information systems.

There is a significant literature devoted to naming and categorizing attributes and visual primitives, but few empirical studies try to account for relationships in image descriptions, so the potential for LIS contributions to this debate could be substantial.

This fundamental gap in the literature can be attributed to the influence of well-established bibliographic standards and the traditional approach to information seeking problems as indexing and subject analysis problems. Michael Krause illustrates this point when he describes how indexers locate the meaning of a photograph within a catalog by breaking apart its meaning and assigning the photograph one word or very few words as authorized headings, also known as *entry points* (Krause, 1988).

2.1 MODES OF DESCRIPTION

One way of viewing relationships as they are applied in LIS is according to modes of description—various systems adopted for describing and representing information. The following four categories are proposed as a framework for guiding and limiting this discussion of the literature: (1) natural language descriptions, (2) machine-readable descriptions, (3) phenomenal descriptions, and (4) conceptual descriptions.

The review begins with a brief definition of what is meant by natural language descriptions and their relationship with machine-readable descriptions. The image descriptions examined in this research project exist entirely in the realm of natural language descriptions—humans expressing image content in everyday language. Next, machine-readable descriptions are defined and image content and relationships within this context are examined in six studies. The review then turns to phenomenal descriptions focusing on research concerned with image searcher behavior. Conceptual descriptions are not considered in this review, but they do play an important role in representing visual information. A significant and distinguishing characteristic of image descriptions in this class are their ability to represent the meaning underlying terms. They are symbolic representations that function in knowledge-based systems instead of databases. Furthermore, concepts and relationships in conceptual descriptions support reasoning. For example, Fahlman describes using transitivity as an axiom of his PART-OF relation and uses transitivity extensively for reasoning in Scone (Fahlman, 1979, p. 114-118).[3]

---

[3] Scone is an open-source knowledge-base system currently under development in the Language Technologies Institute of Carnegie Mellon University. Scone supports representation, searching and limited forms of common sense reasoning, all features that are useful for representing and using knowledge relating to photographs. Situated at the heart of the Scone Project is Scott

2.1.1 Natural language description

Several natural language (ordinary language) approaches have been suggested for accessing meaningful information in photographs. The handwritten inscriptions illustrated in the right half of Figure 3 show how photographer Raymond Baird uses handwritten natural language to describe the subject of a photograph he took on March 18, 1943. On a preprinted form provided by the *Pittsburgh Sun-Telegraph*, Baird describes Florence Smith, the first woman employed in the roundhouse at 190-39th Street in Pittsburgh. Later, a picture cataloger adds subject headings to the back of the print: “Women Workers” and another subject heading in red pencil, “Industries.”[4]

Fahlman, its creator, as well as graduate students who are members of the Scone Research Group at Carnegie Mellon University.

[4] The Pittsburgh Photographic Library at the Carnegie Library of Pittsburgh still maintains a manual picture card-catalog. The P&AGA Catalog (Popular & Applied Graphic Arts catalog) is also still in use at the Library of Congress. Most of the data and some of the images in the P&AGA catalog have been digitized and are accessible through the Library of Congress online catalog. The online counterparts to the paper cards can be viewed at http://www.loc.gov/rr/print/catalog.html. Click on the blue button and scroll down to #51 on the list of collections. Librarians are still directed to the catalog through old references and the catalog provides useful historical insights on where images were filed earlier. Barbara Natanson, Head Reference Section Prints and Photographs Division, Library of Congress, states, “over the years, staff included references to hard-to-find related images in it that aren’t captured elsewhere. It is also a regular stop on my reading room orientation because it demonstrates that long before the age of computers, Library staff knew that people seeking pictures need to see an image early in their searching process.” [E-mail message to researcher, November 13, 2009.]

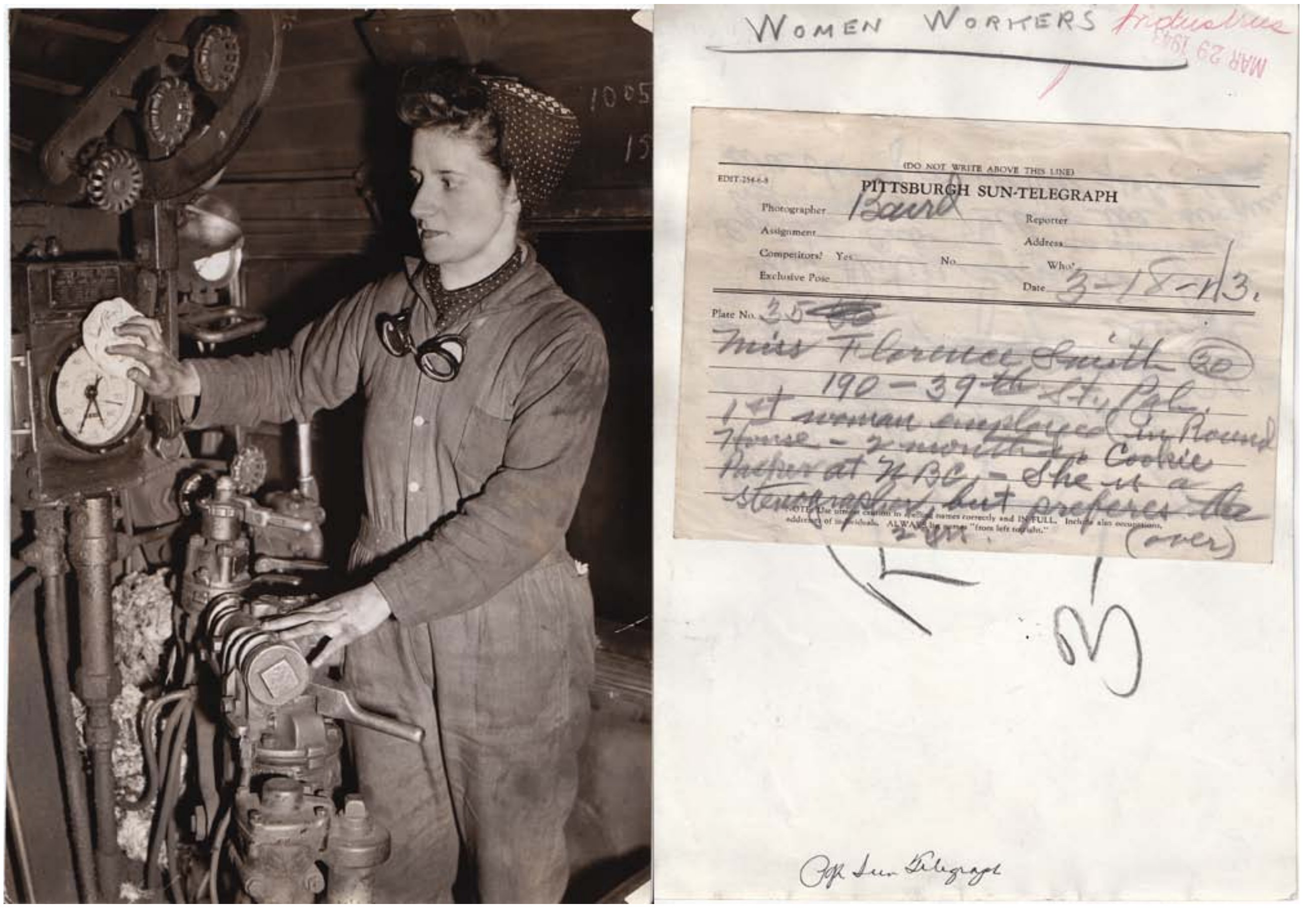
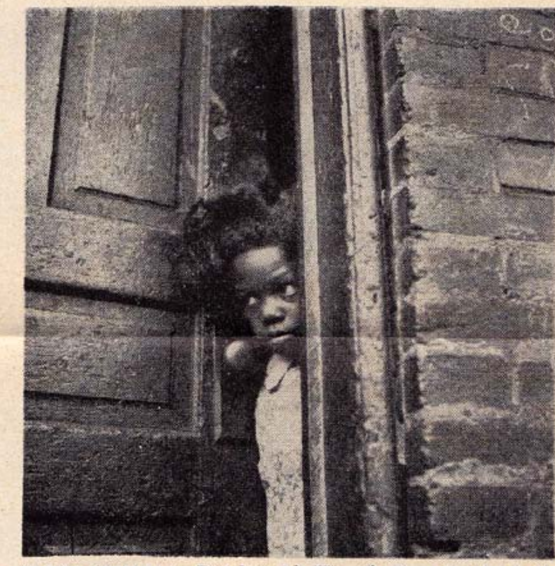


Figure 3: [Raymond] Baird. Women Workers, 1943.] Gelatin silver print, 8 x 10 in. (left). Backside of print (right). (Reprinted with permission.) Carnegie Library of Pittsburgh. All Rights Reserved. Unauthorized reproduction or usage prohibited.

There are several influential works that describe image meaning through other natural language means outside the domain of LIS. These include critical analyses of photographs (Shore, 1998; Szarkowski, 1966), historical perspectives (Batchen, 1997; Batchen, 2001), theory of image (Flusser, 1983; Mitchell, 2005), photography practices (Adams, 1983), and visual ethnography (Taylor, 1994). Roland Barthes suggests accounting for a photograph's meaning in terms of messages: "a linguistic message, a coded iconic message and a non-coded iconic message" (Barthes, 1977, p. 25). Picture captions like those commonly found in newspapers and magazines exemplify what Barthes' means by *linguistic messages*. The caption accompanying

Esther Bubley's photograph in Figure 4 is an example of the *linguistic message* ("Speaking of Pictures," n.d.).



SLUM CHILD in Pittsburgh's Wood's Run district half-opens the door to peer at case worker and photographer.

Figure 4: Photograph by Esther Bubley, Pittsburgh Photographic Library, original negative no. 17. (Reprinted with permission.) Carnegie Library of Pittsburgh. All Rights Reserved. Unauthorized reproduction or usage prohibited.

The drawing in Figure 5 serves as an example of Barthes' *iconic message without a code*. This drawing, taken from an image searcher's query sent to the Pittsburgh Photographic Library is considered a non-coded iconic message because it is an image that does not reproduce everything visible.

Other writers have suggested category systems for understanding photographs. For example, Terry Barrett suggests a category system based on linguistic statements, basing visual analysis on aesthetic form and content. Barrett recommends placing photographs in one or more categories comparable to "descriptive, interpretive, explanatory, ethically evaluative, aesthetically evaluative, and theoretical statements in language" (1986, p. 55). Moholy-Nagy turns to vision and ways of seeing as a means of expressing in natural language the meaning of photographs. She describes eight varieties of photographic seeing, for example, comparing the photogram to *abstract seeing*, reportage to *exact seeing*, and snapshots to *rapid seeing* (1991, p. 52).

Natural language descriptions of photographs in these instances do not assume the role played by descriptions in machine-readable cataloging, better known as MARC. (From this point forward, the phrase ‘machine-readable’ refers to bibliographic and other information that is structured, represented, and communicated in MARC formats.) The content of machine-readable records is a mix of natural language expressions and machine-readable tags.

2.1.2 Machine-readable description

The machine-readable mode of describing photographs is concerned with structured data in records motivated by a desire to make photographs and photographic image data accessible and retrievable in online library systems. MARC is an indexed system, searchable, and brittle in comparison to human thought. This mode of description is *readable* by machines and humans, but the predicates and word senses, or concepts and relationships among concepts that may be represented in database records are only meaningful to humans.

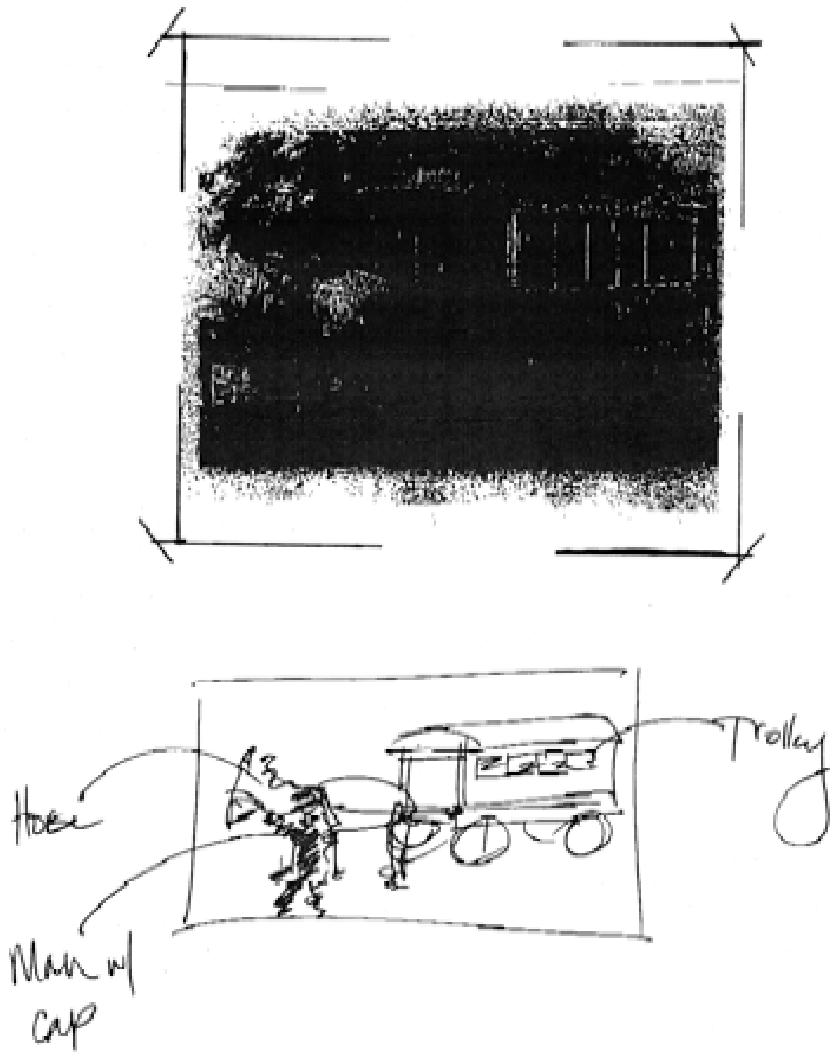


Figure 5: Image searcher's correspondence sent to the curator of the Pittsburgh Photographic Library illustrative of Barthes' non-coded iconic message. The message includes an illegible photocopy of a known PPL photograph accompanied by a line drawing and words

Machine-readable descriptions typically involve formalized standards that generally fall into one of three categories: data structure, data content, and data values standards.[5] When

[5] This terminology is borrowed from a framework proposed by David Bearman for classifying standards and levels of archival description, which was presented at the first meeting of the

photographs are described in machine-readable mode, information is structured around authorized headings lists. For example, the Library of Congress Subject Headings; controlled vocabularies such as the Library of Congress *Thesaurus for Graphic Materials*; and metadata standards such as the Dublin Core Metadata Initiative. Librarians adopt traditional subject cataloging, descriptive cataloging, and finding aids practices when describing image content in this manner.

The main focus of research in this category of description has been on bibliographic entities, but there are some studies that engage with relationships in general and with image content specifically. The following sections examine Farradane's relational indexing (1960s-1980s); Shatford Layne's relationship attributes and Svenonius's scientific model of aboutness—both published in 1994; Green's entity relationship model published in 2001, which touches on the role of relationships in data value standards; Enser's generic-specific continuum published in 2008; and finally, Jaimes and Chang (2000), developers of a conceptual subject indexing model.

Working Group on Standards for Archival Description (WGSAD) who in turn developed a three-dimensional matrix that viewed standards in terms of strength of standard, developer of standard, and level of description. The matrix presents four levels of description. The broadest level is Information Systems standards. The three most directly concerned with describing photographs are data structure, content, and value standards. Data value standards are closely associated with how information is represented in systems. Bearman describes them as "lists or tables of terms , names, alphanumeric codes, or other specific entities that are acceptable for entry in a particular data element." These include subject heading lists, thesauri, and other controlled vocabularies. See David Bearman,"Strategy for Development and Implementation of Archival Description Standards," in *Toward International Descriptive Standards for Archives*; Papers presented at the ICA Invitational Meeting of Experts on Descriptive Standards, National Archives of Canada 4-7 October 1988 (Munich, K.G.Saur, 1993), 161-171. The matrix can be viewed at the Society of American Archivists website, "Standards for Archival Description," http://www.archivists.org/catalog/stds99/intro.html#6.

2.1.2.1 Farradane's relational indexing

Farradane identifies and classifies nine main categories of relationships that he thinks effectively represent meaning among all concepts for indexing purposes. He calls his system Relational Indexing (RI) defining it as "a means of expressing relations on a basis of the mechanisms of thought, to be converted directly into indexing notation…the framework of possible relations is limited and invariant" (Farradane, 1980, p. 268). Farradane's schema of nine relationships and their corresponding codes are presented in Table 2. Figure 6 illustrates Farradane's application of RI to a statement.

Table 2: Farradane's schema of nine relational indexing categories.

Discriminatory mechanism	Associative mechanisms		
	Awareness	Temporary association	Fixed association
Concurrent conceptualization	1 /θ Concurrence	4 /* Self-activity	7 /; Association
Nondistinct conceptualization	2 /= Equivalence	5 /+ Dimensional	8 /(Appurtenance
Distinct conceptualization	3 /) Distinctness	6 /− Action	9 /: Functional dependence (causation)

Farradane is directly concerned with the study of semantic relations as a mechanism for improving document retrieval, arguing that encompassing semantic relations between concepts would enhance current keyword models. To briefly explain how his representation system works, three RI categories are defined and then applied to the statement in Figure 6.

The statement:

"A proposal that copying for research should not be an infringement of copyright of documents from projects supported by the Government of the U.S.A." is expressed in Relational Indexing as follows:

```
          Projects   /: Documents /-   Copying   /; Proposal
            /;            /;             /=
         Supporting    Copyright /-  Infringement
            /;
U.S.A. /( Government
```

Figure 6: Farradane's translation of a natural language statement into RI representation.

Looking at the first line of this statement reading from left to right, Farradane utilizes three relations: *Functional Dependency*, which is the "relation of one thing causing or producing something;" *Action* relation, which "is used for any thing or operation acting on, or affecting, another thing or action;" and *Association*, which "expresses various forms of association, which may be unspecified." (Farradane, 1980a, pp. 271-272). Based on Farradane's definitions, the first four terms presented in Figure 6 may be translated as follows:

[2.1] Documents are the outcome of projects = Projects /: Documents

[2.2] Documents are the object of a copying action = Documents /- Copying

[2.3] There is an unspecified association between the terms "copying" and "proposal" which could be interpreted as taxonomic, as in 'a copying proposal is a type of proposal' = Copying /; Proposal

Farradane bases his relational analysis on psychological principles arguing that humans construct knowledge independent of natural language and that meaning lies in the relations between concepts rather than in the language used to express ideas. He had a difficult time convincing other researchers of his day (Gerald Salton, Karen Sparck Jones, Stephen Robertson, et al.) that RI was a more effective information retrieval tool than the alternative keyword indexing. To prove the validity of his approach, Farradane was faced with the task of building large-scale RI test collections to test against databases using keyword indexing and he had to demonstrate that the test results justified the extra intellectual effort demanded by RI. In 1986, for example, Brooks recalls an instance where he observed Farradane "patiently compiling a database of 1,000 papers in psychology indexed by RI and that took him six months of tenacious effort" (1986, p. 17).

In the mid-1980s, in light of computers being successfully programmed to search for words and phrases, Farradane remained persistent in hand-indexing documents, limiting relationships to nine types, and insisting they were derived from psychological principles, which Brooks criticized, asking "how could one program psychological principles?" (Brooks, 1986, p. 17). This review finds no evidence of Farradane's influence in any later research examined here.

2.1.2.2 Shatford Layne's relationship analysis

Shatford Layne—best known for her writings on subject analysis (Shatford, 1986)— briefly touches upon the notion of relationships in the context of images in 1994, when she describes four types of attributes that should be recognized when cataloging images. She introduces the notion of *relationship* as a type of image attribute suggesting, "Images can be related to, or associated with, other images, or textual works, or even objects" (p. 585). This confuses the term *attribute*, which describes monadic, one-term, subject-predicate structures, with relationships, which describe associations among two or more items.

Shatford Layne offers only hints regarding what types of relationships might exist in the domain of images in her sample propositions listed below in examples [2.4]-[2.8]. She recommends that it is important to not only identify the objects being related, but also indicate the nature of the relationships involved, asserting there are relationships among entities in propositions [2.4] and [2.5], but does not state explicitly what those relationships are, which she designates using a series of question marks.

[2.4] preliminary drawings *????* finished painting

[2.5] architectural plans *????* image of finished building

[2.6] image illustration *appears in* textual work.

[2.7] image *depicts something described in* textual work.

[2.8] textual work *describe* image.

In these examples, it appears she is not concerned with relationships that might occur within the content of the visual image. The relationships Shatford Layne attempts establishing connect together objects that are members of collections. She draws relationships between images and the bibliographic works within which images are contained. The term ‘image’ is treated as a primitive term in the proposition and it helps explain the relationship but the image itself is not explained or analyzed. In the following section Svenonius contributes to the discussion on image attributes framed as roles and the problematic nature of image descriptions in subject analysis.

2.1.2.3 Svenonius’s scientific aboutness model

In the same year that Shatford Layne introduces relationships as an image attribute, Svenonius looks closely at visual information and the extent to which it can be described in linguistic terms. She offers as an example the photograph *Migrant Mother*. Svenonius states, “There are no words for what is expressed. What is expressed cannot be spoken of; it cannot be referred to using language; it cannot be named and cannot be indexed by index terms" (Svenonius, 1994, p.603). Her work on subject analysis of image content is thus concerned with the problem of translation—decoding visual information and making ontological commitments in a domain Svenonius calls "wordless media."

She approaches the problem of subject in pictures by examining the issue of *aboutness*, beginning with the logical assertion that "Snow is white." The term "snow" is the subject and it denotes a concept or thing. The predicate "is white" describes the concept or thing. Svenonius then extends the notion of subject to the level of documents stating that a document about snow consists of many propositions about snow. Continuing along this line, the topic of snow eventually develops into a body of literature about snow, a branch of knowledge, "a *subject* that is studied" (Svenonius, 1994, p. 601).

During the course of her discussion, Svenonius reconfirms the traditional viewpoint that entities in the bibliographic domain are dichotomously classified as either books or nonbooks. The general assumption is that members of these classes are clearly distinguishable possessing unique properties and that all materials are members of one class or the other. It is not difficult, however, to find exceptions that belong to neither category, such as talking books, wordless picture books, and photographs of text. A more interesting and obvious question is why re-present "snow is white" as a single index term "snow." The meaning underlying the original assertion might also be about "white snow" or that the state of snow is white. Here the important distinction is that Svenonius postulates a model of subject indexing, one she refers to as the *scientific aboutness model*, by placing subjects in the context of grammar. She asserts that an indexer who applies this model would describe a document using the index term "snow" if and only if the document contained information about snow.

A visual researcher or picture cataloger is more likely to describe a photograph as being a picture of *children playing in the snow*, *white snowdrifts*, *snow-covered mountains,* or *dirty snow piled high along streets in Boston*, but it is not likely that anyone would describe a photograph with the caption: “snow is white.” This brings to light an important distinction between how humans describe image content using concepts and semantic relationships and the purposes underlying thesauri classification, which is to normalize and extend the descriptions of documents by using descriptors. As Svenonius describes it, “a query using the term “Snow” put to a document retrieval system retrieves all documents indexed by the descriptor “Snow.”

Svenonius turns to philosopher Suzanne Langer (1948) for an explanation of visual symbolism in hopes of explaining how the scientific model of *aboutness* might be extended to pictures, and whether visual information is really about the same thing as written texts, though the subject happens to be expressed visually. Langer argues that the elements in a photograph, “are not units with independent meanings” representing the elements that language represents (Langer, 1957, p. 94). “They are a thousand times more numerous” (Langer, 1957, p. 95). She points to light and shadow as examples of picture elements without names capable of displaying gradations that cannot be enumerated. Yet contrary to Langer’s assumptions about picture elements, artists describe at great length the use of light and shadow in photographs.

Throughout this paper Svenonius brings into her argument Shatford’s well-known distinctions of *ofness* and *aboutness* as tools for uncovering nontraditional indexable attributes in images (Shatford, 1986). Shatford, in discussing subject access to pictures, is best known for extending Erwin Panofsky’s iconological method to categorization of the subjects of images. Based on Panofsky’s three levels of meaning—pre-iconographical, iconographical and iconological—Shatford (1986) lays the groundwork for classifying the meaning of pictures as *generic of*, *specific of*, and *about*. Shatford argues that subject indexing can be performed at the first two levels of meaning. That is, at the pre-iconographical, a picture's ofness is what it depicts and its aboutness is what it expresses. At the iconographical level ofness is what a picture represents and aboutness is what it symbolizes. At the highest level of meaning, the iconological level, Shatford believes distinctions cannot be drawn between ofness and aboutness and subject indexing is not possible.

Svenonius began her argument thinking that there was only one problem to investigate, namely determining subjects for nonbook materials that defy subject indexing because they are visual in nature. Her conclusion, however, points to another fundamental issue beyond the traditional book/nonbook dichotomy. It is the unique nature of the image medium in terms of its role. When the medium plays a descriptive or documentary role, it is more likely to be a candidate for effective subject indexing. In other words, "Subject indexing presupposes a referential or propositional use of language" (Svenonius, 1994, p. 605).

2.1.2.4 Green's Entity Relationship Model

Green (2001) provides a practical place to begin a discussion of the literature as it pertains to relationships in LIS and their role in defining data value standards. In LIS there are a limited number of relationships and entities they join together through associations that are named, defined, categorized, and given some kind of organizational structure and symbolic representation. To begin understanding what these are, Green proposes a system for specifying relationships—a means of explaining their semantics—by designating three conditions: (1) what entities are bound by the relationship, (2) what an entity's role is within the relationship, and (3) a determination of whether the relationship is abstract or concrete. Green's specification encompasses subject relationships, properties of relationships and entities, and instances, thus forming an entity-relationship model from which a simple ontology emerges shown in Figure 7.

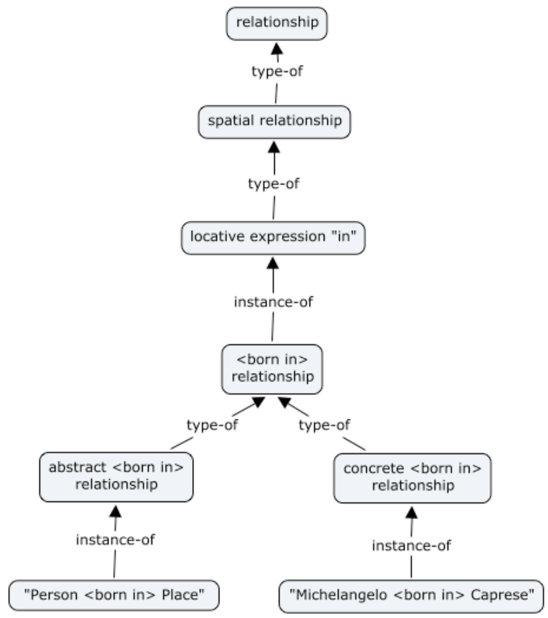


Figure 7: Diagrammatic view illustrating Green's implicit <born in> relationship.

In Green's formulation, while the relationship *born-in* is not explicitly defined, she argues that the specification model demonstrates that knowledge of the entities involved provides clues as to the nature of the semantic relationship linking entities together. In other words, the nature of the entities joined by the *born-in* relationship determines the relationship's classification as either abstract or concrete. The semantics of *born-in*, therefore, are not the variables of interest here as much as it is the entities joined together by *born-in*. Green presents three propositions, or knowledge statements, attempting to illustrate the roles played by entities in specifying semantic relationships:

[2.9] Person <born in> Place

[2.10] Michelangelo <born in> Caprese

[2.11] Raphael <born in> Urbino

Green classifies proposition [2.9] as an *abstract <born in> relationship* because it joins together entity classes—the set of all persons with the set of all places. The unique nature of the entity, therefore, is that it represents a class of things rather than a named individual. She classifies propositions [2.10] and [2.11] as *concrete <born-in> relationships* because they form associations between proper names, or what she calls "specific entities" (Green, 2001, 3).

Green further refines semantic relationships by classifying them according to an open or closed class distinction. She draws a correspondence between parts of speech and class distinction, extending it to relationships constituting open or closed classes. Her proposed open and closed class distinctions are illustrated in Table 3. A closed class can be enumerated whereas an open class cannot. In other words, it is not likely that someone will invent a new preposition, but it is not unusual to see new nouns and adjectives entering our language.

Table 3: Open and closed class distinction applied to relationships

Parts of Speech	Action	Conceptual Component	Type of Class	Linguistic Relationships	LIS Relationships
Noun	tends to name	Entities	open class		
Pronoun	tends to name	Entities	closed class		
Adjective	Tend to name	Attributes	open class		
Adverb	Tend to name	Attributes	open class		
Verbs	generally expresses	Relationships	open class	Expressed through syntagmatic Relationships	associative
prepositions	generally expresses	Relationships	closed class	expressed through paradigmatic relationships(hyponymy, meronymy, synonymy, antonymy)	Hierarchical and equivalence
Conjunctions	generally expresses	Relationships	closed class		

This important observation regarding the character of closed class relationships helps define the nature of spatial relationships in photographs, which are closely related to prepositions. Prepositions are linguistic parts of speech shown in Table 2 to be a closed class of words. This has important implications for the analytic constructs applied in the pilot study described in chapter four and subsequent research carried out in this dissertation. It provides for an interesting space to investigate knowing that in terms of linguistic expressions, at least, the set of possibilities is most likely finite.

Returning to Green's earlier distinctions between *abstract* and *concrete* relationships, she states that in the context of LIS, "data modeling of bibliographic entity classes operates with abstract relationships, but catalogs record concrete relationships" (Green 2001, 3-4). Historically, the photograph in library information systems has been categorized along with books as being a type of bibliographic entity. To extend the categories of concrete and abstract relationships to machine-readable records describing photographs ultimately returns to the problem of representation. Does a photograph represent and convey information in the same way as, for instance, a book, and is the assertion true that abstract relationships represented in bibliographic entity classes are distinct and separate from concrete relationships represented in catalog records?

Turning to Sowa (2000) for an answer, he notes that abstract forms may be embodied in physical entities and that the same abstract form may be embodied in many different physical objects, especially in library information systems. Take for example Dorothea Lange's photograph "Migrant Mother" housed at the Library of Congress and instead of using the spatial relationship "in" to relate two entities, we use the predicate "photograph" to relate Lange to the photograph *Migrant Mother*, as in "Lange photographed *Migrant Mother*.

Applying Green's formulation, an abstract relationship would be expressed in proposition [2.12] and a concrete relationship would be expressed in [2.13].

[2.12] Person <photograph> Person

[2.13] Dorothea Lange <photograph> "Migrant Mother"

Applying Sowa's analysis, the name "Migrant Mother" could refer to an abstract form conceived by Dorothea Lange, or it could refer to its embodiment in the film negative, the film copy negative, or the photographic print. Once the name is coded in MARC and the image is scanned for processing in the Library's OPAC (Online Public Access Catalog), the number of "Migrant Mothers"—both physical and abstract—multiply. The point Sowa is making is that the same abstract form may exist in multiple physical embodiments and that distinctions must be drawn not only between abstract and concrete relationships, but between the many abstract forms that could be represented in different ways. These distinctions are especially important in computational systems where ontological commitments must be made clear, for example, in knowledge-based systems.

Green turns next to the semantics of relationships describing four ways in which they can be understood. Ranging from implicit to explicit, they are:

1. The enumerations of participant or participant types (entity types) imply the semantic relationship. Green compares this to "see also" or RT (related term) in thesauri.

2. The semantics of a relationship type holds by convention. She points to Ranganathan's PMEST formula as an example.[6]
3. The terms used in natural language enable a person to access an understanding of the semantic relationship. A person through personal experience, for instance, understands the semantics of the relationship "part of."
4. The most explicit way in which the semantics of a relationship may be understood is by definition.

Green acknowledges, however, that explicitly stated relationships in knowledge organization have limited application based on end user needs. The limit of their usefulness is sometimes based on the claim that end users do not understand the notations used for expressing standard subject relationships. Added to this, it is no simple matter to specify the precise nature and semantics of relationships or enumerate the types of entities that participate in or imply certain kinds of relationships.

2.1.2.5 Enser's generic-specific continuum

Enser (2008) proposes doing away with the generic-specific dichotomy that is central to Shatford's (1986) subject oriented organization scheme for pictures, Armitage and Enser's mode/facet matrix (1997), and Jaimes and Chang's Pyramid (2000). In Enser's formulation he moves from a dichotomous model to a generic-specific continuum. His model utilizes Shatford Layne's (1994) composite model of "subject," made up of four attributes: object, spatial,

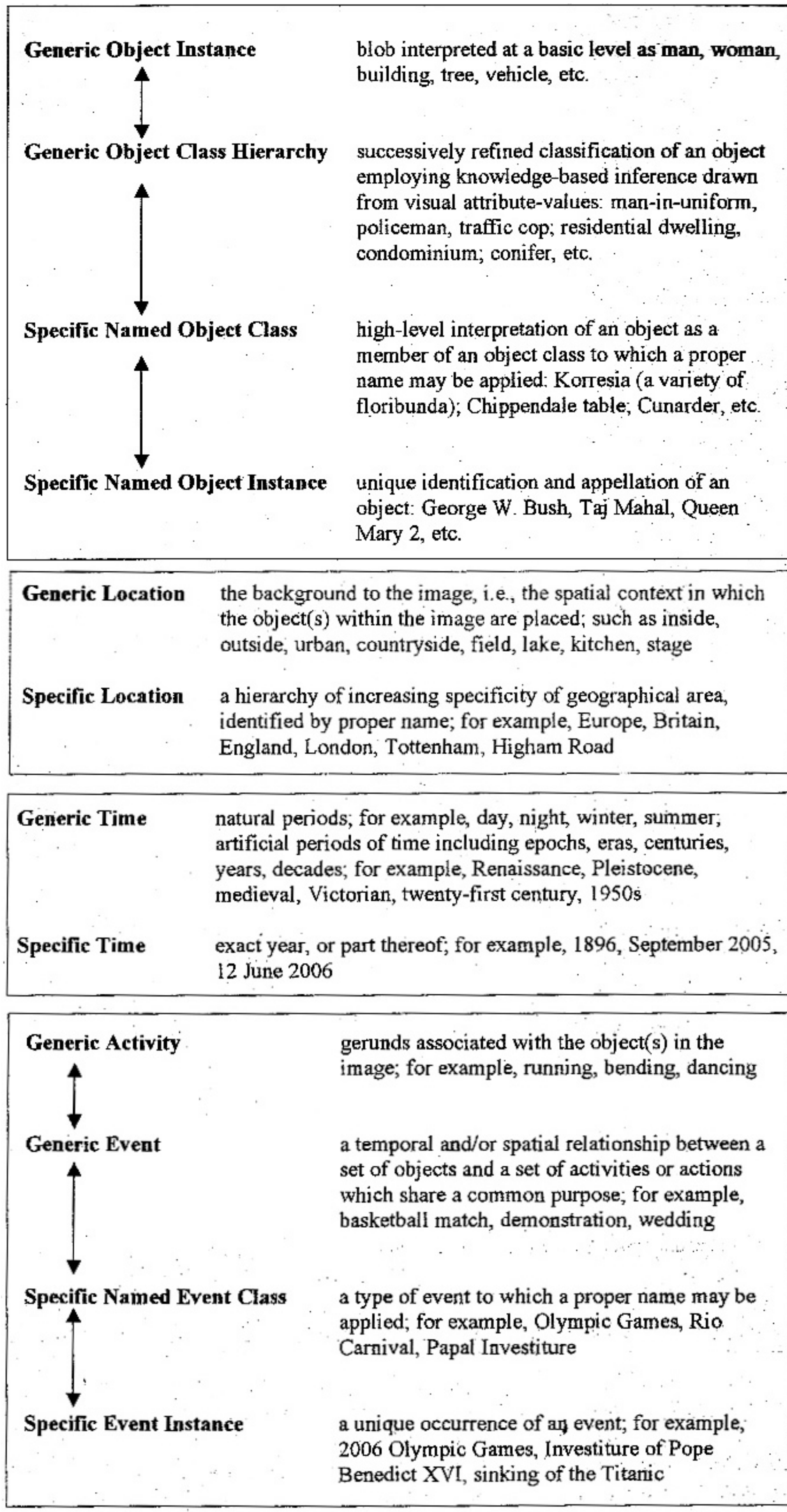
[6] PMEST stands for five facets used in faceted classification. These include Personality (topic), Matter (form), Energy (process), Space (physical location), and Time (date). Green argues that the conventions of Colon Classification systematically determine the semantic relationship.

temporal, and activity/event. Enser argues that his continuum model accommodates a hierarchy of gradations. Modeling the *object* attribute of subject, as shown in Figure 8, Enser uses bi-directional arrows to represent the notion of a continuum. However, it appears that what was formerly a two-category model of generic and specific has evolved into a four-category model: (1) Generic Object Instance, (2) Generic Object Class Hierarchy, (3) Specific Named Object Class, and (4) Specific Named Object Instance.

The four labeled-categories raise the question, is Enser's model ontologically neutral and did he intend to accommodate all possible relationship types in images? If not, what are the limitations of this model?

It appears in Figure 8 that he makes it easier to conceptualize hierarchical relationships among objects than to conceptualize other entities and their relationships. For example, making explicit the position of natural lighting during a shoot in relation to the objects and their shadows pictured in a photograph would require representing entities outside the frame of the photographic image. This falls outside the parameters of objecthood in image content. Furthermore, there can be hierarchical relationships at the "Specific Named Object Class" level just as he has shown they are possible at the "Generic Object Class Hierarchy." Absence of the term "hierarchy" in his model seems to imply classes of things at the "Specific Named" level are not hierarchical.

He informally defines what he means by "generic location" and "specific location." Locative expressions in his model, however, are limited to geographic locations and places and do not include other spatial expressions such as paths. For instance, "bear running around a tree."



Generic Object Instance	blob interpreted at a basic level as man, woman, building, tree, vehicle, etc.
Generic Object Class Hierarchy	successively refined classification of an object employing knowledge-based inference drawn from visual attribute-values: man-in-uniform, policeman, traffic cop; residential dwelling, condominium; conifer, etc.
Specific Named Object Class	high-level interpretation of an object as a member of an object class to which a proper name may be applied: Korresia (a variety of floribunda); Chippendale table; Cunarder, etc.
Specific Named Object Instance	unique identification and appellation of an object: George W. Bush, Taj Mahal, Queen Mary 2, etc.

Generic Location	the background to the image, i.e., the spatial context in which the object(s) within the image are placed; such as inside, outside, urban, countryside, field, lake, kitchen, stage
Specific Location	a hierarchy of increasing specificity of geographical area, identified by proper name; for example, Europe, Britain, England, London, Tottenham, Higham Road

Generic Time	natural periods; for example, day, night, winter, summer; artificial periods of time including epochs, eras, centuries, years, decades; for example, Renaissance, Pleistocene, medieval, Victorian, twenty-first century, 1950s
Specific Time	exact year, or part thereof; for example, 1896, September 2005, 12 June 2006

Generic Activity	gerunds associated with the object(s) in the image; for example, running, bending, dancing
Generic Event	a temporal and/or spatial relationship between a set of objects and a set of activities or actions which share a common purpose; for example, basketball match, demonstration, wedding
Specific Named Event Class	a type of event to which a proper name may be applied; for example, Olympic Games, Rio Carnival, Papal Investiture
Specific Event Instance	a unique occurrence of an event; for example, 2006 Olympic Games, Investiture of Pope Benedict XVI, sinking of the Titanic

Figure 8: Enser's generic-specific continuum model.

Jaimes and Chang (2000) are interested in capturing the meaning of visual and non-visual information for the purpose of developing a conceptual indexing structure to assist in describing images in digital image libraries. Over the course of their paper they present a variety of representational views and frameworks for describing image attributes and relationships. In their attempt to provide a unified, interdisciplinary approach they generate a set of possible lenses through which to view meaning in image content. A small ontology of image content relationship types and instances of relationships evolve from their discussion.

They begin with their motivating question, which is, given an image, the problem is determining what to index and how to carry it out. They start by carving up the world into general concepts and visual concepts and then ask what meaning emerges along these dimensions given an image of “a portrait of a man wearing a suit.” They answer, “suit, man, physical size, physical weight, hair color,” and so on. It is worth noting that the semantic relationships at the sentential level are not preserved in the sample list of index terms. In other words, it is not stated whether the man is wearing the suit, the suit is hanging in his closet, or that it is at the cleaners. No definition is offered for general or visual concepts, but in a later example of describing a ball, the authors state that its color and shape are visual concepts and its weight and material are general concepts.

The current study attempts to address ways of gaining knowledge about such entities by taking into account their corresponding relations. In other words, to make sense of what it means to say “suit,” “man,” “hair color,” when describing image content these entities are to be understood as classes and roles and that each instance of a class in an image description stands in an instance-level relationship to some corresponding instance.

The systematic indexing model Jaimes and Chang propose actually consists of four separate schemata. The central schema, shown in Figure 9, is a pyramidal structure built up from ten levels and to the right of the pyramid they categorize five types of relationships. The top four levels in the pyramid are labeled *syntax* and *percept*. The lower six levels are labeled *semantic* and *concept*. They define *percept* as “what our senses perceive—in the visual system it is light” (p. [3]). Examples of percepts include texture and color. This coincides with Brentano’s ontological category of *inherence* that identifies entities by their intrinsic qualities, such as shape and color, and Peirce’s category of Firstness, “the conception of being or existing independent of anything else” including relationships with other entities. (Sowa, 2000, pp. 60). The concept levels “refer to an abstract or generic idea generalized from particular instances” (p. [3]). This level of meaning depends on background knowledge and interpretation.

Intersecting with the percept and concept are the notions of *syntax* and *semantics*. Jaimes and Chang describe *syntax* as “the way visual elements are arranged” and *semantics* as “the meaning of these elements and of their arrangement” (Jaimes & Chang, 2000).

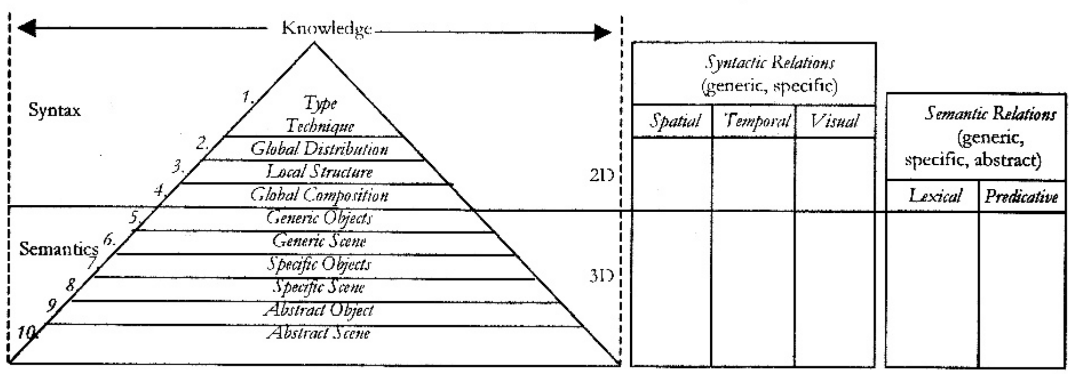


Figure 9: Jaimes and Chang’s pyramidal indexing structure.

The researchers make the claim that the pyramidal model has a direct impact on what searchers look for and how searchers go about finding it. The pyramid’s shape symbolizes degrees of knowledge required to carryout indexing. Concepts situated near the top of the pyramid require less general or specific world knowledge than those at the bottom. Their semantic levels align with Panofsky's modes of image content, adding to a growing succession of image behavior researchers who adopt Panofsky's modes to image indexing and classification.

The researchers state that they are interested in relations among images elements, but frame their discussion around relationships that occur within the pyramidal model in Figure 9 where two broad categories of relationships emerge: *syntactic* and *semantic*. Syntactic relationships occur between image elements at any level in the model. Semantic relationships occur only in levels five through ten. Five instances of relationships are represented in Figure 10. To assist in clarifying their syntax, they are re-presented in [2.26] through [2.29].

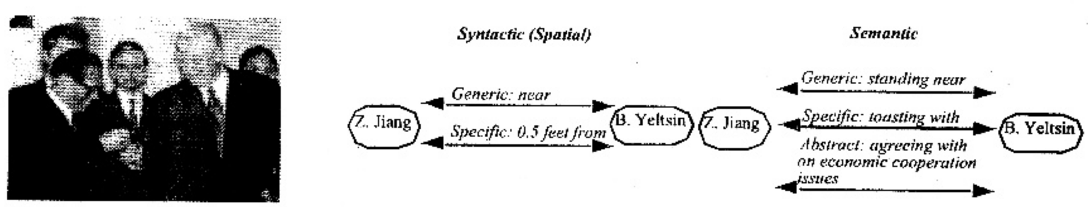


Figure 10: Jaimes and Chang illustrate their applications of syntactic and semantic relationships.

Syntactic (Spatial)

[2.25] Z. Jiang *Generic: near* B. Yeltsin

[2.26] Z. Jiang *Specific: 0.5 feet from* B. Yeltsin

Semantic

[2.27] Z. Jiang *Generic: standing near* B. Yeltsin

[2.28] Z. Jiang *Specific: toasting with* B. Yeltsin

[2.29] Z. Jiang *Abstract: agreeing with on economic cooperation issues* B. Yeltsin

During the course of their discussion, a simple ontology of relationships and instances of relationships emerge, although making it explicit as shown in Figure 11 makes it clear that there are some modeling problems. What is important, though, is seeing the various ways Jaimes and Chang begin to examine this aspect of image description.

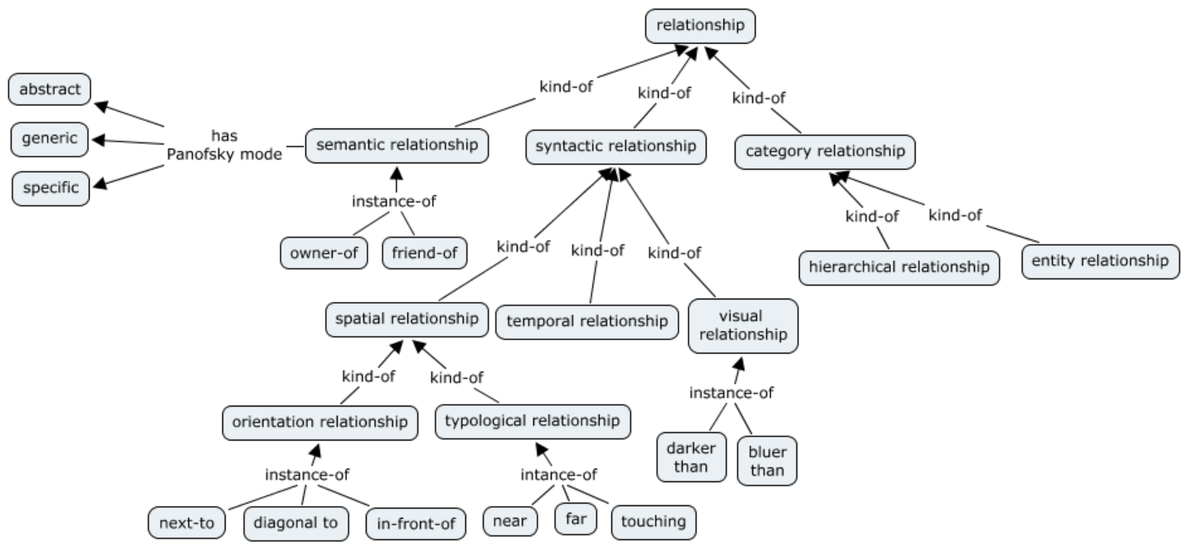


Figure 11: Simple ontology illustrating Jaimes and Chang's proposed relationship types and instances.

One modeling problem that surfaces, for instance, is that entities "friend-of" and "owner-of" are considered instances of "semantic relationship" or possibly subtypes of that class. These would traditionally perform as functions or attributes in an information system. Consider a photograph showing two friends named John and Mary. The caption might read, "John with his friend Mary" and a more formal representation could be written as a function [2.30] or expressed as a binary predicate [2.31].

[2.30] *John(friend_of(Mary))*

[2.31] *has_friend(John, Mary)*

Furthermore, the category labeled “visual relationship” may be too general and ambiguous to count as a relationship type and deserves clarification, as do all the other nodes and links in the model. The diagrammatic view in Figure 11 attempts to illustrate what the authors are explaining in the text of their paper and it makes explicit the flaws that are less easily detected in the context of a natural language description. More importantly, a visual representation makes it easier to detect what interesting ideas about relationships are being stated here and which ones can be carried forward into future research on relationships expressing image content. The current study proposes building an ontology of relationships that will be conceived in this manner. That is, as a graph-theoretical structure consisting of terms that form nodes that are linked together by means of edges called relations.

In 2001, Jörgensen, Jaimes, Benitez & Chang test the Pyramid’s conceptual structure and usefulness as a tool for capturing attribute classification. Their research questions, while interesting, focus not on relationships but rather on how well the Pyramid assists student indexers classify terms describing image attributes. In Appendix A of the study “relationships” is listed as an attribute and two terms “brothers” and “romance” are given as examples of relationships. Other than this hint at connections or associations that may exist between members of the class “persons,” there is no other reference to relationships. This study and other work in this area remains focused on indexing image attributes (Fidel, Hahn, Rasmussen, & Smith, 1994; Shatford Layne, 1994), the range of attributes used for describing images (Jörgensen 1998), and analysis of image search logs measuring user behavior based on number of sessions, searches per session, terms per query, and so on (Jörgensen, & Jörgensen,, 2005). Recognition and analysis of relationships among attributes, for the most part, remains an elusive subject.

2.1.3 Phenomenal descriptions

Phenomenal descriptions evolve during research studies examining image user behavior (Jörgensen, 2003) and during analysis of concepts and relationships included in controlled vocabularies (Bean, 1996) or data sets extracted from online image databases (Tribble, 2010). This is a less direct link to machine-readable processes. Researchers develop categories of concepts and relationship types and apply various systems of representing these entities within the context of their research papers, which may or may not coincide with existing bibliographic machine-readable standards. Phenomenal descriptions are often times more formal than natural language, but not always. More importantly, this section examines research on the nature of description within the framework of image searchers' questions.

2.1.3.1 Enser's unique/non-unique dichotomy

Enser (1993) examines the nature of user demand for visual information by closely looking at the form and content of user requests received by the Hulton Deutsch Collection Limited. The unique character of this research project is the author's recognition that designing effective retrieval systems relies in part on understanding the nature of image searchers' queries—a belief that lies at the heart of the current dissertation proposal.

The Hulton Collection contains over ten million images ranging in format from photographic prints and negatives to cartoons, maps, and engravings. It is important to note that the majority of queries comprising the data set for this study are telephone queries where a Hulton picture researcher mediates and elicits a subject statement and records the client's requests on an "Internal Enquiry Form." It is not clear what part of the requests reflect the client's own words and what part, if any, is interjected or interpreted by the picture researcher.

As part of this study, Enser also tries to assess the extent to which the Collection's *Gibbs-Smith* and *Keystone* classification schemata could represent the level of detail found in the uniquely defined subject requests. The *Gibbs-Smith* schema is explored in detail in section *2.2.3.2 Enser & McGregor's Hulton Collections analysis*. Enser's research is important to the current study because it examines the semantic content of a set of image requests that contain complex, natural language statements expressing image attributes and relationships among attributes.

Enser brings together a test collection of 1,000 client requests that are selected and stratified according to six user types including "other." A total of 2,722 individual image requests were extracted from the 1,000 request forms. Enser attempts to analyze the test collection by classifying image content according to two of Erwin Panofsky's levels of understanding and interpretation: (1) *preiconographical description*, which Panofsky describes as the first level of interpretation that a viewer apprehends "by identifying pure *forms*, that is: certain configurations of line and colour, or certain peculiarly shaped lumps of bronze or stone, as representations of natural *objects* such as human beings, animals, plants, houses, tools and so forth" (Panofsky, 1939, p. 5). Enser equates this with a level of meaning found in generic subject queries, and (2) *iconographical description*, which Panofsky describes as a level of interpretation gained through knowledge of literary sources and practical experience. For example, ascribing politeness or a polite greeting to the lifting of a hat (Panofsky, 1939, p.4). Enser has difficulty applying this dichotomous classification system, however, acknowledging that while the initial request might be stated as a generic topic—pre-iconic subject—the picture that is finally selected and retrieved for the client has iconic properties associated with it. He illustrates this problem with a generic request for a pre-iconographic image of "the first microscope." Enser explains that in order to find this image, the picture researcher performs interpretation, factoring in other unique properties not explicitly stated by the client. The retrieved image, he argues, is essentially an image that is iconic in nature and concludes that Panofsky's preiconographical and iconographical classifications used for fine art do not work effectively in the general commercial environment of the Hulton collections.

Enser sets out devising a new image property called *uniqueness*, which he defines as "a request for the visual representation of an entity, the desired, particular occurrence of which can be differentiated from every other occurrence of the same entity type" (Enser, 1993, p. 29). He settles on four categories of image requests: (1) non-unique, (2) non-unique refined, (3) unique, and (4) unique refined, and adds sub-categories of time, location, action, event and technical specification. For example, the request "shell shock" would be classified as a unique subject and "shell shock after First World War" as a unique subject with the refiners of time period added. The query "5-6 year old boy trampolining, in mid-air, in silhouette" is considered by Enser to be a non-unique subject, a boy, refined by age, event, and technical specification (in silhouette).

Enser discovers that refiners play an important role in characterizing visual information. In 34% of the unique and non-unique queries, the target entity is expressed in the context of a given time period or era. This has implications for accommodating the capture and representation of temporal relationships in the current proposed study. Enser also discovers that 69% of all requests sought unique entities associated with people, objects, locations or events, especially in the case of requests coming from newspaper and magazine publishers. An example of a request in this category is "crying, distress, must be over 16, good focus on individual." Given the extent to which refiners are added to unique and non-unique subjects, Enser concludes that clients rely heavily on the picture researchers' roles as intermediary and that the Gibbs-Smith scheme "can function only as a blunt pointer to regions of the Hulton collections where pertinent material might be co-located" (Enser, 1993, p. 35).

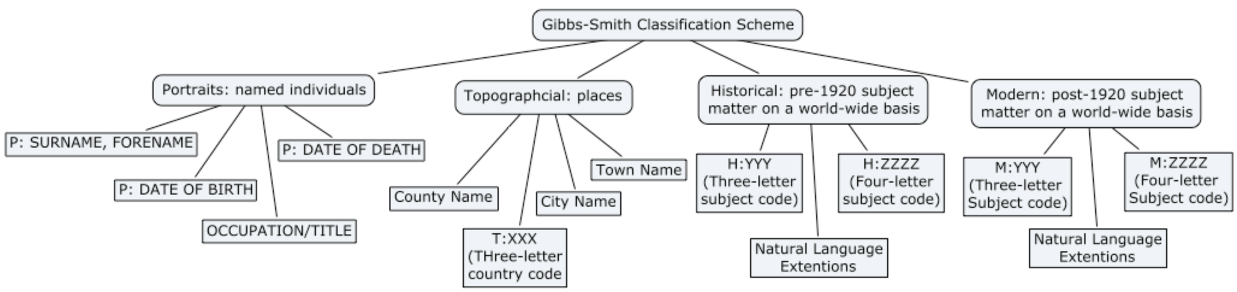
2.1.3.2 Enser & McGregor's Hulton Collections analysis

In Enser and McGregor (1993) the *Gibbs-Smith* and *Keystone* classification schemata are introduced—subject classifications developed for use in cataloging the pictures in the Hulton image collections. This formal research report is a more detailed analysis of the research project introduced in the preceding section and it is reviewed here to see what, if any, relationships are revealed in Hulton's image classification systems. Focus is placed on Gibbs-Smith classification, "the more extensive and comprehensive of the two schemes" (p. 3). This report also includes valuable illustrations showing samples of prints from the collections, request forms, a cross reference sheet, over four hundred image searcher's queries, and a "pictorial request" similar to the one discovered in the sample data set for the current study pictured earlier in Figure 5.

The scope of the Hulton picture collections spans the picture contents of the Hulton Group's periodicals, the best known being *Picture Post.* The predicted client base and the potential range of picture topics are unlimited: "every 'picturable' subject and activity on earth…throughout history to the present day" (Gibbs-Smith, 1950). The original intentions were that the classification scheme would work in partnership with a well-informed picture researcher knowledgeable of the collections and that successful retrieval would engage both directed searching and browsing.

Enser and McGregor describe the Gibbs-Smith scheme as hierarchically structured dividing knowledge into four broad categories. They present it "formally" as a set of terms representing classes and subclasses. The diagrammatic view presented in Figure 12 is a literal re-presentation of the narrative explanation illustrated here as a concept map with nodes and unlabeled links arranged hierarchically in a top-down fashion. Viewing the schema in this manner makes clear some of the strengths and weaknesses of the model.

Figure 12: Diagrammatic view of the Gibbs-Smith classification scheme.



Three of the more interesting observations are presented here.

1. The relationships among entities are not made explicit (the edges linking nodes are not labeled) and do not appear to reflect a hierarchical framework as suggested by the authors. This less formal structure invites a broad interpretation of headings and may explain in part how the staff succeeded in manually cataloging some ten million images—the estimated size of the collection in 1993.
2. Every top-level node could represent two or more concepts and relationships, or predicates and words senses. The resulting ambiguity may make processing and subject analysis less cumbersome and time consuming.
3. A second-level node or subclass appears to represent attributes or *roles* associated with members of their superclasses. This reflects a general viewpoint expressed throughout the

literature that at this level of modeling—natural language narrative—it is not necessary to be explicit or precise in distinguishing between roles and relationships.

4. A distinction is not drawn between a person's name (a string) and the person (a human being).

The model supports three or more levels of specificity. As the level of specificity goes up, more attributes are associated with the broad class. For example, the three levels of specificity for the "Portraits" class are illustrated below.

- Specificity Level 1: P:SURNAME, FORENAME
- Specificity Level 2: P:SURNAME, FORENAME: DATE OF BIRTH, DATE OF DEATH
- Specificity Level 3: P:SURNAME, FORENAME: DATE OF BIRTH, DATE OF DEATH: OCCUPATION/TITLE

A photograph the authors draw from class M (Modern) is classified to specificity level 4 as: M:MOT:BUSE:COACHES:PASSENGERS. This can be translated as meaning a subject classification of "Modern: Motor transport: Busses: Coaches: Passengers."

Other evidence of semantic relationships in this system can be found in the cross-referencing. Enser and McGregor describe three relations that guide users to other relevant parts of the collection: (1) Here, (2) Away, and (3) See also. These are roughly equivalent to the traditional thesauri relations Broad, Narrow, and Related terms.

2.1.3.3 Enser, Sandom, & Lewis

In Enser, Sandom, & Lewis (2005), the interest shifts to the semantic gap as it pertains to still images. Their research concerns designing and building image retrieval systems that bridge this gap.[7] The authors focus their attention on surveying still image types, image users, and image metadata with hopes of providing insight into scope and significance of the semantic gap. They are concerned with what they view as image retrieval in "real-world" applications.

Their project is important because the research design of the current dissertation proposal mirrors certain components Enser, Sandom, and Lewis incorporate into their study. One of these components is the attention they give to the ontological question: What image types exist in the domain of still images and how do they relate to one another? They develop a simple taxonomy of images illustrated in Figure 13 that includes a list of informal definitions. Their emphasis on "real-world" scenarios and the operations side of image retrieval systems underpins their research efforts as well as the current study's goals. They provide an opportunity to see first-hand the visual researchers written query and corresponding image and textual record that is subsequently retrieved. One notable difference, however, concerns aspects of representation. Enser, Sandom, and Lewis only provide implicit evidence of relationship types in image descriptions, but the goal of the study now under consideration is to represent and structure concepts and relationship information in a consistent and semi-formal manner so it can be systematically reexamined and reused in the future as a corpus data set in knowledge base and information retrieval research.

---

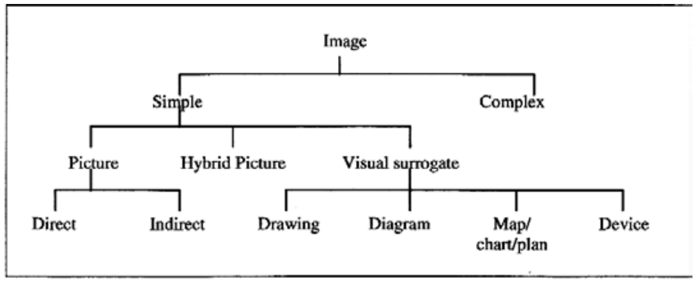
[7] In traditional content-based image retrieval (CBIR), the *semantic gap* refers to regions of information that lie between the low-level, automatically extracted image primitives such as shape, color, and texture, and the higher-level processes that humans apply during searching and describing tasks such as interpretation or that requires recognizing or labeling abstract entities such as point of view and mood. (*See* Yang, H-C. & Lee, C-H. (2008). Image semantics discovery from web pages for semantic-based image retrieval using self-organizing maps. *Expert Systems with Applications 34*: 266–279.)

Enser, Sandom, and Lewis ground their arguments in the belief that, “the retrieval utility of visual images is generally realized in terms of their inferred semantic content” (Enser, Sandom, & Lewis, 2005. p. 177). They claim inferential reasoning arises from semiotic distinctions drawn between denotation and connotation of image content. The content-based image retrieval (CBIR) community has attached the label “semantic image retrieval” to the formulation and resolution of information needs that engage this process. The authors conclude, from research conducted by Enser (1995), Armitage & Enser (1997), Ornager (1995), and Enser & Sandom (2002), that, “*identification* is dependent upon prior existence—and knowledge by the user—of a defining linguistic label” (p.180). In other words, people search for images that depict named objects and events. The authors add that certain image components may be recognizable by shape alone--for example, an image of a refrigerator--while certain other attributes may rely on textual annotations--such as limiting search results to refrigerators manufactured in the 1950s.

Other image attributes that reside in textual metadata accompanying images are context and significance of image features. Their point here is that there are processes taking place during image search and retrieval that engage a person’s higher-level cognitive processes, “completely removed from CBIR functionality” (p.180). The authors conclude that this provides evidence for a need to develop systems that support both text-based and CBIR indexing. It is worth noting that systems that rely on full-text, keyword, and subject indexing do not move us closer to the automation of inferential reasoning. Reasoning over image metadata, a process the authors agree is an essential part of image description and retrieval, can only be realized once image metadata is formalized to a level of representation and structure which allows machines to effectively process its meaning.

The authors' taxonomy shown in Figure 13 reveals a simple ontology that, without close scrutiny, appears to be hierarchically structured. Upon closer examination, however, the ambiguous node labels and lack of labels on the edges results in uncertainty and concepts may require refinement. For example, to say "A *Simple* is a type of *Image"* sounds nonsensical, just as nonsensical as saying "A *Picture* is a type of *Simple*" and "A *Direct* is a type of *Picture*," and so on.

A brief look at Enser (2008) shows him extending this model three years later, describing a visual information domain—still and moving images—as a dichotomous community with one group working in image retrieval and another engaged in curatorial image management. The former group is made up of researchers and the later practitioner. Enser is concerned that there is limited communication between these two groups and that image retrieval researchers know very little about image searchers' needs or the logistics of managing picture collections. This results in procedures and practices that are technologically feasible, but that may serve little useful purpose in practice.



The following definitions have been used in the construction of this taxonomy:

Image	a two-dimensional visual artefact.
Simple Image	an undifferentiated image.
Complex Image	an image which comprises a set of simple images.
Picture	a scenic or otherwise integrated assembly of visual features.
Hybrid Picture	a picture with integral text.
Visual surrogate	non-scenic, definitional visual artefact.
Direct Picture	a picture, the features of which can be captured and/or viewed within the human visible spectrum.
Indirect Picture	a picture, the features of which must be captured and/or viewed by means of equipment which extends viewability beyond the human visible spectrum.
Drawing	an accurate representation (possibly to scale) of an object, typical applications being in architecture and engineering.
Diagram	a representation of the form, function or workings of an object or process.
Map/chart/plan	a representation (possibly to scale) of spatial data, typical applications being in geography, astronomy, and meteorology.
Device	a symbol or set of symbols which uniquely identifies an entity, e.g., trademark, logo, emblem, fingerprint.

Figure 13: Enser, Sandom, & Lewis's taxonomy of still images, 2005.

Enser frames this problem by trying to illustrate the principles and practices of image retrieval from the researcher and practitioner's perspectives. This is an important work because it continues to build on his assumptions about the world of still image formats and his development of a simple ontology organized around a taxonomic framework. He takes the position that the *direct picture* is the most common image form researchers are concerned with in indexing and retrieval research and upon which image retrieval activities can be conducted, defining *direct picture* as “a picture, the features of which can be captured and/or viewed within the normal human visible spectrum” (Enser, 2008, p. 5).

2.1.3.4 Keister

Keister (1994) writes about a retrospective analysis and re-analysis of users' queries at the Prints and Photographs Collection of the National Library of Medicine (NLM). The dataset consists of the reference query log for the year 1984 and 291 additional queries from 1991. Keister does not disclose the total number of queries she analyzed.

Keister recognizes the rich meaning embedded in image searchers' queries, but stops short of identifying and discussing the meanings of these relationships. She does draw an important distinction among users from the health profession, museum, and art community noting, for instance, that "picture professionals (still picture researchers, TV, film, or media personnel) think visually and use art and/or graphics jargon, e.g., 'an action shot of George Papanicolau, has to be horizontal and color' " (1994, p. 9). She infers a relationship between types of image searchers and the language of the query. Furthermore, while this query represents only one example, it serves as evidence that an image searcher expresses more than just image attributes when submitting a request. This example introduces, for example, a *functional expression* "action shot of" that when applied to the name "George Papanicolau" denotes a new single entity. Thus, applying "action shot of" to the name "George Papanicolau" we get the phrase, "action shot of George Papanicolau," which, in a digital image archive, could automatically designate some unique individual photograph. Admittedly, interpreting this as a functional expression may not seem unusual coming from a librarian in 2010, but when Keister published this paper in 1994, LIS image research was not concerned with this level of semantic analysis.

It is worth noting that according to Keister, the queries in this study are reconstructed from memory by staff members' "cryptic notes." This makes it difficult to distinguish a staff members' interpretations from the searchers' original language. Additionally, it is not clear whether the three *query portrayals and dialogs*—mediated reference interviews included in this paper –are verbatim transcriptions or memories of conversations recalled by staff members. In spite of these shortcomings, on the surface what is most striking about these dialogs, in addition to the predictable problem of conveying visual information in words, is the rich use of semantic relationships to link together entities. For example, one person purportedly asks for an image of "poor people, especially children, maybe on a city street, lame or crippled, with canes" (1994, p. 11). Like the earlier example, the searcher relies on a rich set of semantic relationships that goes beyond listing single keywords.

Keister introduces a new term describing image queries that attempt to "construct images with words." She labels this type of query as an *image construct query,* possibly recognizing the existence of a unit of expression that is semantically richer than a simple unordered set of keywords. She provides an example of an *image construct query* using a well-known image in their collection by Benjamin Rush titled "Tranquilizing Chair." Keister claims that image searchers requesting this image generally ask for, "the man sitting in the chair with a box on his head" (1994, p. 13). She reports that one-third to one-half of all image requests in her library end up being this type of query. She offers other examples including *people racing in wheel chairs* and *surgeons standing.*

Keister recognizes aspects of the semantic relationship problem recognizing that searchers cannot find pictures using their descriptions if the words they are using are not accessible in the catalog record. Her solution includes cataloging images at the item level (versus finding aid descriptions at the collection level) and including a surrogate image in the catalog record. The fundamental record, however, remains a database record structured around MARC with controlled vocabulary content.

2.1.3.5 Bean

While at the Department of Medical Informatics, Columbia University, Bean (1996) tried to determine the character, structure, and organization of implicit associative relationships among entities relevant to clinical anatomy. Her research is relevant to the current study because she begins adopting analytic tools for determining the nature of relationships in more specific terms than those expressed later by Greisdorf, H. & O'Connor (2002). Her approach can inform the domain analysis of image content in photographs.

Bean examines a total of 256 term pairs that she organizes into five broad clusters or semantic categories of terms: (1) Procedures, (2) (Other) Anatomic Entities, (3) Functions, (4) Disruptions to Functions, and (5) Chemical Agents. Bean's description of a semantic relationship can be illustrated as shown in Figure 14.

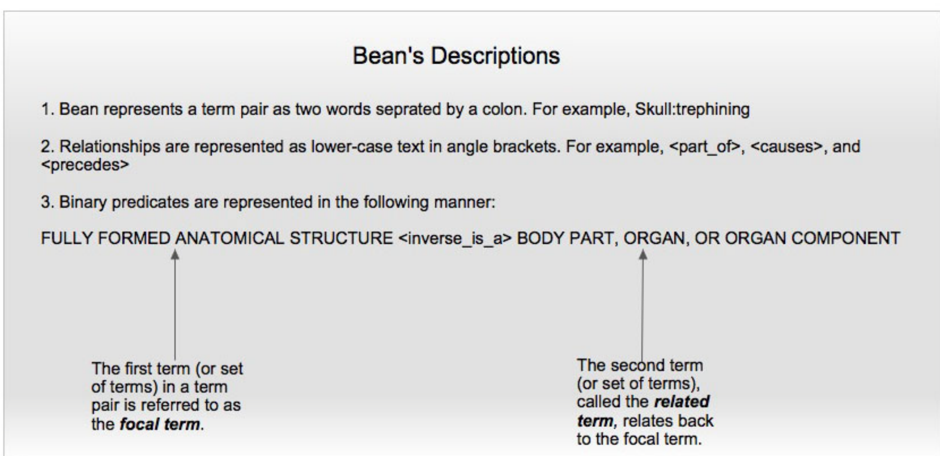


Figure 14: Bean's descriptions of entities and relationships in the Medical Subject Headings (MeSH).

Bean's investigation shows that the most common semantic category was neo-classical compounds joined with a combining form that designates a therapeutic, preventative, or diagnostic procedure. For example, *Bronchi:bronchoscopy*. This suggests a need for a system of representation and structure that supports both entities and non-hierarchical relationships among entities. In Bean's terminology, this example illustrates a semantic category of procedures that relate terms naming a particular procedure (*bronchoscopy)* performed on a focal anatomic entity (*bronchi*). Her goal is to identify the associative relationship joining these two terms.

By applying Bean's analysis to the photograph description in Figure 15, similar associations among term pairs can be defined. For example,

[2.14] *package_of_rice:conveyer*

[2.15] *package_of_rice:fiber_carton*

[2.16] *State_rice_mill:Abbeville_Louisiana*

Figure 15: Russell, L. (1938c). *Removing packages of rice from conveyer to place them in fiber cartons. State rice mill, Abbeville, Louisiana.* (Library of Congress, Prints & Photographs Division, FSA/OWI Collection, [LC-USF34-031383-D (b&w film neg.)]

In some cases the related term (second term) is not known. Bean predicts that this happens when the procedure is specific only to a particular anatomic entity, for instance, *gonads: castration*. In other words, *castration* is a procedure specific to the focal term in this term pair. Herskovits (1986) maintains us that general and specific world knowledge about a particular domain or context—in this instance, knowledge of medical subject headings and procedures—enables making pragmatic inferences, adding facts to a description that go beyond the original utterance. In the case of the archival photograph pictured in Figure 15, enough contextual information is given to infer from the description that it is a *woman* who is <removing> the packages and <placing> the packages. Bean considers relationships such as <removing> and <placing> as *low-level reality*. That is, they are instances of applying general principles to specific instantiations. Although she does not use the term *ontology* anywhere in her paper, her *high-level* model would in effect be concerned with ontological relationships or relationships that can be generalized across any domain. The relations <removing> and <placing> might then be classified instances of case relationship types. Bean does not, however, make these distinctions explicit in her research.

Bean observes that in the contemporaneous UMLS (1994), relationships between anatomical entities and procedural entities are not allowed. She suggests that this is the case because procedures act by affecting the functioning of anatomic structures—something that cannot be expressed within the confines of a binary relationship, the only relationship that is allowed in the UMLS. Complex assertions consisting of multiple concepts, therefore, can only be expressed as a series of binary relationships, which in turn obscures any direct relationship between procedure and anatomy. Bean represents a binary relationship between a therapeutic procedure and an organ by chaining together four binary relationships as illustrated in [2.17].

[2.17]THERAPEUTIC OR PREVENTIVE PROCEDURE <treats>
PATHOLOGIC FUNCTION

PATHOLOGIC FUNCTION <is_a> BIOLOGIC FUNCTION

BIOLOGIC FUCNTION <has_location> FULLY FORMED
ANATOMICAL STRUCTURE

FULLY FORMED ANATOMICAL STRUCTURE <inverse_is_a>
BODY PART, ORGAN, OR ORGAN COMPONENT

Bean argues that this representation ignores the reality that many procedures are direct actions on an anatomic entity, which in turn leads to a change in function. The UMLS(1996) does allow a direct relationship between these two entities, but only as a locative expression as shown in [2.18].

[2.18] THERAPEUTIC OR PREVENTITIVE PROCEDURE
<has_location> BODY PART, ORGAN, OR ORGAN COMPONENT

According to Bean, this suggests a need for a set of distinctive relationships that reflect precise actions of procedures on anatomical entities. Her ontology of relationships is presented diagrammatically in Figure 16. Bean's semantically overloaded class designations—made more complex by the use of compound terms—would only function within the context of the MeSH considering that her main desire is to eventually pursue using a simple pattern-making program to classify relationships based on string matching features of terms.

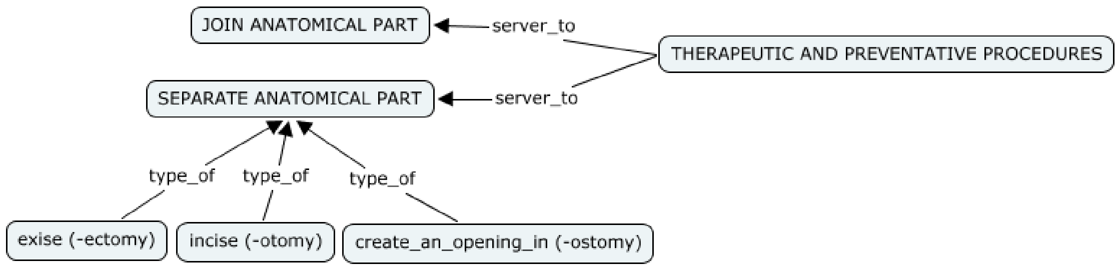


Figure 16: A simple ontology of Bean's procedure relationships.

2.1.3.6 Armitage and Enser (1997)

Armitage and Enser (1997) provide an analysis of a cross section of query types collected from seven libraries whose archival holdings in still and moving images cover a wide variety of topics. Their goals are to advance the understanding of visual information needs and to inform the interface designs developed for accessing image collections. This research is an extension to earlier research by Enser and McGregor (1992) and Enser (1993)—studies concerned with analysis of information need in the visual domain.

In this study, Armitage and Enser choose a sample of 1,749 images in consultation with library staff claiming that the sample represents "the usual flow of requests." The authors examine image descriptions in the context of image queries and organize the data very broadly according to query type and image content type and then more specifically they code image content by subject. In their analysis they label four types of image queries: (1) image content (for example, 'I want a picture of so and so'), (2) identification/attribution/provenance checking, (3) accessibility to work and ownership/viewing availability, and (4) miscellaneous. The last category presents interesting possibilities for semantic analysis of the outliers that they consider to be, among other things, "unusable queries" and "requests for administrative procedures" (p. 288).

An implicit categorization of queries emerges that divides questions into (1) mediated, recorded questions and (2) unmediated questions. The former, it appears, are questions jotted down by reference librarians during reference transactions. The later, extracted from image request files in public libraries, are queries expressed by image searchers in letters and/or faxes sent by image searchers. The authors present several examples of mediated and unmediated queries describing images. A sampling is shown in Figure 17.

Mediated (National Film and Television Archive):
1. "1937 television recording of Rothschild sale."
2. "Urban life."
3. "man/women working in nuclear power stations + wearing protective clothing."

Unmediated (Birmingham Central Library):
1. "...[to] trace a photograph of Corporal Martin Langfield, a bugler in the above charge, ...depicted in one of the several books written on the battle."
2. "...a photograph of Mr Hugh Gilzean-Reid—MP in 1885 for Aston Manor."

Figure 17: Samples of mediated and unmediated image searchers' queries describing images.

The authors present a second system that classifies queries according to "image content," which presents a problem in that searchers are not describing images per se, except in the case of known-image requests. Their four categories include (1) By named artist, (2) Known items, (3) Unique subjects, and (4) Non-unique subjects.

Armitage and Enser build a model for analyzing levels of meaning in image content based on Panofsky's modes of image analysis (Panofsky, 1962) and facet analysis introduced by Markey (1983) and refined by Shatford (1986). The focus here is on the utility of their mode/facet matrix and the extent to which it meets their promise of representing and characterizing both queries and image content. The matrix is presented in Figure 18. It functions by generating twelve categories of subject content by combining one of Shatford's facets (who, what, where, when) with one of Panofsky's modes (iconography, pre-iconography, iconology). The result is assigned a code consisting of a single letter and number.

	Iconography (Specifics)	Pre-iconography (Generics)	Iconology (Abstracts)
Who?	individually named person, group, thing **(S1)**	kind of person or thing **(G1)**	mythical or fictitious being **(A1)**
What?	individually named event, action **(S2)**	kind of event, action, condition **(G2)**	emotion or abstraction **(A2)**
Where?	individually named geographical location **(S3)**	kind of place: geographical, architectural **(G3)**	place symbolised **(A3)**
When?	linear time: date or period **(S4)**	cyclical time: season, time of day **(G4)**	emotion, abstraction symbolised by time **(A4)**

Figure 18: Armitage and Enser's matrix for coding categories of subjects in image content

A theme throughout this research paper is the problem of interpreting concepts consistently and determining in what subject category to place borderline and ambiguous concepts, compound terms, and entire sentences. This paper provides an opportunity to explore these problems— gradations of variability and deviating from the set of conditions placed on the denotation of a class of subjects. While not explicitly stated in these terms, the problem surfaces while they struggle to pigeonhole words and phrases into rigid categories.

The issue of finding the right category in which to place a thing aligns closely with Dorr's special composition question, "Under what circumstances do several things compose something?" (Dorr, 2005, p. 234). The 'several things' in this instance could be the multiple facets present in a query and the entity that is composed of these parts could be the query's subject category. For example, Armitage and Enser claim a hierarchical relationship between unique and non-unique subjects in the visual domain. As they state it, “an entity can always be interpreted into an hierarchy of related super-concepts and sub-concepts” (p. 290). Their concern is at what level in the hierarchy does an entity cease being unique and instead become non-unique. This is a common problem in dichotomous systems of classification—a problem explored earlier in Enser’s (2008) continuum of *general* to *specific* in section 2.2.2.5. In Dorr’s terms, the question posed here could be framed as: ‘Under what circumstances is there a subject category having each of several facets as component parts, every facet of which is related to one of them?’

The special composition question can be applied to a specific example where Armitage and Enser try to code the query, “cheetahs running on a greyhound course in Haringey in 1932.” They conclude that the subject is S2, that is, a *specific individually named event*. Firstly, they answer the compositional question in part by saying an *individually named entity* does not necessarily have to consist of a proper noun—a named individual. It can be a generic reference. Secondly, the essence, or sense of the query can be conveyed by a single facet or notion of *an event* without making the remaining concepts and relationships in the query explicit in the coding. In their terms, the S2 category in this case is “less cumbersome” than a G1 + G2 + S3 + S4 coding (generic cheetah, greyhound, and course + generic event + a named geographical location: Haringey + date 1932).

A useful example to illustrate the problems Armitage and Enser face in applying the mode/facet matrix is illustrated in the single facet versus multi-facet image description. This can also be described as the geographical location problem. Consider the three queries listed in Table 4.

Table 4: Three queries applying combinations of Armitage and Enser's mode/facet codes.

Query	Code (see Figure 15)
[2.19] carnivals	G2
[2.20] Rio carnivals	S3 + G2
[2.21] Rio Carnival, 1986	S2 + S4

Query [2.19] is a single facet generic query; query [2.20] is a multi-facet query consisting of a unique, named geographic location plus a non-unique subject; query [2.21] is a multi-faceted query consisting of a unique, named geographic location plus a specific date. The question they raise is how to classify [2.20] when its component parts are both a specific geographic location and non-unique event. Their solution is to introduce the concept of *refinement*, that some subjects can be non-unique, but refined with a modifier or "refiner." Thus, "carnival" is non-unique (G2) refined by location "Rio," which is unique (S3). Based on the implicit relationships in the Panofsky-Shatford mode/facet matrix, the image content described in Table 4 can be diagrammatically modeled as shown in Figure 19.

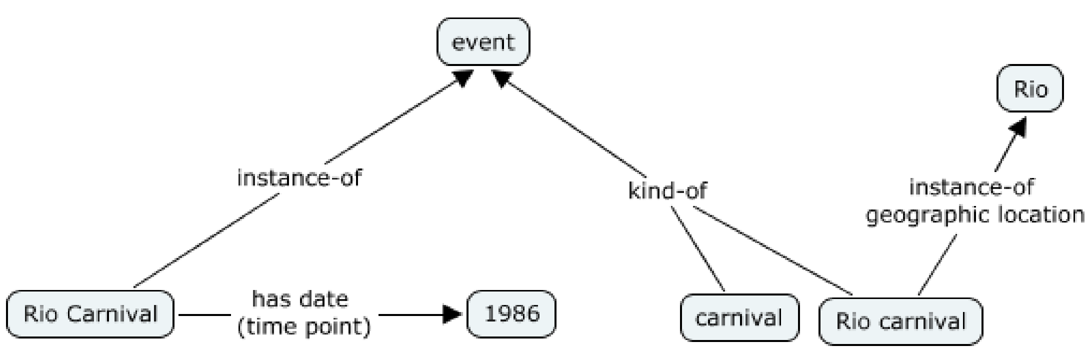


Figure 19: Armitage and Enser's model of three events (carnivals, Rio carnivals, and Rio Carnival, 1986).

Viewed from the perspective of Figure 19, concepts appear tangled with one another in unexpected ways. A “Rio carnival,” for instance, is coded as S3 + G2, which means it is a kind of event and it is an instance of a geographic location. Ontologically, however, “Rio carnival’s” most fundamental relationship is its relationship to carnival. That is, a “Rio carnival” is a kind-of “carnival,” but this viewpoint cannot be expressed or encoded in the mode/facet matrix of subject classification. This is not a criticism of the matrix model. The model was not intended to express image content at this level of granularity. An attempt should be made in the dissertation proposal detailed here, however, to attempt extending Armitage and Enser’s analytic approach. For instance, if the concept “carnival” is expressed as having a date and location function (that is, date-of and location-of), it could serve as a template for explaining any number of events expressed in image descriptions comprised of an event name, and event location, and an event date. A possible logical representation for this expressed as closely to first order as this researcher can make it, is expressed in [2.22].

[2.22] $\exists x,y\ \text{carnival}(x) \wedge \text{event}(y) \wedge \text{Rio}(\text{location-of}(x)) \wedge 1986(\text{date-of}(x)) \wedge x=y$

The authors describe another geographical location problem that exists in a query that asks for an image showing “scenes of London.” Their informal definition of a named geographic location states that a geographic location is, by its very nature, unique and specific “because it occupies different space to any other named location” (p. 294). In what ways would New York City and the Big Apple—two different named geographic locations—occupy different spaces?

What makes the Ozark Plateau a uniquely different space from the Ozark Mountains? In both of these examples different names share the same denotation, but different senses.[8]

One final problem that arose that is of interest to the current project is how to express attributes of the image carrier included in the image description. The authors use the mode/facet matrix for this purpose by stretching the meaning of Panofsky's modes to include the physical manifestation of works of art. If an image search asks for a painting of a well-known person, for example, Armitage and Enser code the image of the individual as S1 (individually named person) and the format as a "sub-selection" G1 (generic kind of thing).

2.1.3.7 Collins

Collins (1998) is concerned with effectively indexing visual materials. She argues that little is known about patrons' visual information needs in historical photographic collections. This leads her to investigate whether there is a need to index more image attributes at the pre-iconographic level and whether some attributes can be considered more important than others. Another overarching question is whether detailed item level descriptions of photographs are better than collection level descriptions. She addresses these concerns by focusing on these specific research questions:

1. What percentage of queries involves terms describing generic subject matter?

[8] This is a well-known problem, the name/referent distinction. Gottlob Frege argued that names express a denotation and they express a sense and that words can have different senses, but the same denotation or conditions of reference. A classic example presented by Frege is the distinction between "Hesperus," the ancient Greek name for the evening star, and "Phosphorus," the ancient Greek name for the morning star. Two different names with two different senses—independent astronomical facts that mean two different things—however, but both share the same condition of reference. Both refer to the planet Venus.

2. Are there attributes of images that are not being indexed but should be useful for retrieving images?

3. Which categories of terms are most commonly used in queries?

Collins considers pre-iconographical descriptions as, “identification of forms as representations of objects and event,” adding that it can be factual or expressional (p.38). This level of description is considered the primary subject matter level of description (Keefe, 1990; Markey, 1988; Shatford, 1986). Collins assumes that visual researchers’ queries can be decoded in Panofsky’s analytic framework. Shatford argued this point using Panofsky’s theory as a basis for subject analysis during the description process. Before proceeding further, two critical points deserve close attention: (1) There is an important difference between focusing analysis on the image content of photographs and focusing analysis on the expressions humans use to describe images. The former is concerned with picture elements and entails interpretation on the part of the researcher. The later focuses on descriptions—linguistic expressions humans use to represent image content. (2) The “subject” is treated as a primitive in image indexing, but intuition alone suggests that the index term or “subject” of a photograph can be complex and not easily reduced to a controlled vocabulary term. Even the simplest term engages some level of pragmatic inference. That is, humans rely on general and specific world knowledge and knowledge about communication to infer more facts than what is apparent in the utterance of a single index term.

To support her argument for more primary subject matter, or pre-iconographical description, Collins says catalogs usually describe an image’s secondary subject matter. This corresponds to Panofsky’s iconographic level of analysis. To support this argument, Collins offers us a scenario where an image is described at the secondary subject matter level, which she claims requires a higher level of interpretation. For example, the caption, “Maj. Jesse Marcel holding debris from Roswell crash.” This, she proposes, prevents an image search from searching at the primary pre-iconographic level on generic phrases such as “people holding debris” and “men in uniform”—“information that both specialists and nonspecialists can use to gain access to the collection” (p. 40).

This reinforces the point raised earlier, that the complexity of a primary subject is underestimated and considered a generic primitive. Yet, Collins’ own example of “people holding debris” reinforces her earlier definition that primary subjects describe objects and events. “Person holding debris” is not an index term in the LCSH, AAT, TGM. It is an example of two entities being joined together by the relationship “holding,” which in an instance of a type of case relationship. In formulating her analysis, Collins establishes a dichotomy of image users: specialists and nonspecialists. That the former benefits from secondary subject matter and latter from primary subject matter (Collins, 1998, p. 41; Markey, 1988, p. 167).

2.1.3.8 Jorgensen’s examination of user behavior

Corinne Jörgensen maintains, “the problem involved in access to images is the question of how to index them” (Jörgensen , 1998, p. 162). Her research is devoted to establishing classes of image attributes based on user behavior in carrying out three categories of tasks she calls describing, sorting and concept searching (2003, p.203). Analysis of her data reveals twelve distinct classes of image attributes with the ‘object’ class being the most prevalent level of image description in all tasks.

Jörgensen argues that a major component missing from all the representation techniques and models discussed in her book *Image Retrieval* (2003), “is the notion of relationships, or interaction, among image elements (Jörgensen, 2003, p. 257). Jörgensen is referring to problems surrounding the nature and structure of the relationships among individual image attributes and classes of attributes. She returns to the original content produced by participants in her study to determine conceptual relationships, adopting Graesser and Goodman's (1985) concepts of text analysis, which argues "implicit knowledge structures contribute as much to understanding as does explicit information" and proceeds to describe two relationship types that emerge during her dissertation research (Jörgensen, 1995, p. 144).

Participants in her study were asked to view and describe six projected images within the three categories of descriptive viewing, descriptive searching and descriptive memory. Descriptions ranged in length from "a low of two terms to a high of sixty terms per image description" (Jörgensen, 1995, p. 145). She briefly examines occurrences of hierarchical relationships measuring what percentage of terms are "basic level," subordinate, and superordinate. As an example, if "gun" represents a basic level term, then "riffle" represents a subordinate term and "weapons" a superordinate term. She determines, for example, that in the descriptive viewing task 76.6% of the terms are basic level, 11.8% are superordinate, and 11.6% are subordinate.

In the conceptual search task, participants were shown the same six images as those used during the other tasks, but this time are asked to describe in words the search statement they would model to retrieve each image. In her analysis of this task it is revealed that participants describe images in a narrative style using a wide variety of relationships as shown in [2.23]—[2.25]. There was no attempt made, however, to label or categorize what relationship instances were expressed.

[2.23] Cartoon or drawing of whale surging out of body of water. Nightscape with abundant stars and land in the distance.

[2.24] I need a picture of a traditional Japanese woman looking out on the night sky through a small window. The background is all white and a red vine or floral design across the picture.

[2.25] Need picture of a boy riding a rocking horse through outer space.

Jörgensen makes reference to another relationship she calls an “object-figure/action structure” defined as “a main figure or object and an action performed by or upon the main figure or object”(Jörgensen , 1995, p.173). She notes that 64% of participants’ image descriptions contain an object-figure/action structure. This relationship is characterized by an object or figure term followed by a verb, in turn followed by one or more nouns or prepositional phrases. Beyond these two types of relationships, Jörgensen’s discussion of relationships is implicit and blends into the general landscape of her research. For example, she describes images as having focal areas and names the primary focal area the “central object,” “main subject,” or “figure” and the area that remains in the background as the “context” or “ground” (Jörgensen, 1995, p. 159). She brings these two entities into a relationship with one another, but does not state explicitly what that relationship is or how it is expressed in the participants’ descriptions.

2.1.3.9 Greisdorf and O’Connor

The last review in the phenomenal category of descriptions looks closely at Greisdorf and O’Connor (2002) who are concerned with what image viewers think about when they evaluate images. They base their study on three assumptions. Firstly, that prior research indicates there are seven descriptive categories of image attributes that describe viewers’ percepts. These include color, shape, texture, object, location, action, and/or affect. Secondly, they support a *hierarchy of perception* model for image search, retrieval, and evaluation. This model, first introduced by Eakins and Graham (1999), organizes aspects of human perception on three levels:

(1) primitive features (e.g., color, shape, and texture), (2) objects (e.g., person/thing, place/location, and action), and (3) inductive interpretation (e.g., symbolic value, prototypical displacement, and emotional cue). The authors argue that image searchers' interpret and evaluate images along these three dimensions. Thirdly, they perceive image retrieval as a process of reconceptualizations. The query context is where concepts are initially embodied and the viewing context is when concepts emerge. This extends earlier analysis (Greisdorf, 2000) that sought to develop a general-purpose three-step process for evaluating retrieved text that leads from determining whether a document is on topic, to comprehension, to determining whether the document can be used.

Nineteen participants were given twenty-seven single, pre-selected query terms and ten pre-selected, grayscale images drawn from the NOAA government database. Terms were selected based on the seven attribute categories mentioned earlier. Participants were also given the opportunity to develop their own descriptors, which were examined in relation to the seven attribute categories to determine whether other categories emerge. The researchers were interested to learn which categories of terms were used most often and which categories seem to be unimportant. They were particularly interested in learning to what extent participants tapped into Level 3 in the hierarchal levels of perception. That is, do image seekers look for entities that are not objects pictured in the content of the image?

Their results suggest that affective/emotion-based query terms are an important descriptive category in image retrieval, positing that, "These adumbrative, impressionistic and abstractionist concepts that relate viewer to image need to be captured with some type of retrieval mechanism in order to enhance retrieval effectiveness for system users" (Greisdorf & O'Connor, 2002, p.11). In other words, if image seekers are interested in, for example, the emotions an image expresses, then these aspects of image need to be indexed along with the objective descriptions.

The difficulty that arises in discussing the analysis of what image viewers think about is the effects of presenting participants with pre-selected terms and images. The researchers argue that user self-reports tested *a posteriori* showed that the users were just as likely to match the pre-selected words to the pre-selected images as not, so the chosen words, images, and attribute categories did not unduly bias the selection process (p.13). The reasons for their results are to some extent apparent given the limited character of the pre-selected materials. For instance, a participant describes one image of a large body of water resembling a lake surrounded by a forest using the descriptors *green*, *boat*, and *fishing* even though the image is grayscale and neither boats nor fishing were present in the scene. The choice of these descriptors, the authors argue, supports the notion that image seekers look for things not present in images. This leads to their central argument that “Information sought in the form of an image may not always be a retrieval process, so much as a conveyance process” (Greisdorf & O’Connor, 2002, p.19). In contrast to textual indexing, they suggest classifying and categorizing images according to metaphoric, analogical, metonymic, or synecdochic relationships. It is important, in their view, that these aspects of an image be conveyed to image searchers who are uncertain about the actual topic they seek.

In a cursory discussion attempting to delineate these different figures of speech and relate them to images, the authors explain that metaphoric relationships are used when images or picture elements within an image stand for something else and analogy conveys something about the story an image tells. There are countless examples of photographs picturing businessmen wearing suits and carrying an attaché while climbing a ladder--a picture that could be interpreted as a metaphor for *climbing the ladder of success*. The authors admit that this meaning could be represented by existing retrieval systems through captions attributed to photographs by the artist, but claim that the artist's labeling may not match the searcher's needs.

They define *metonymy* as "a figure of speech that consists of using the name of one object or concept for that of another to which it is related" (pp. 20-21), such as "lend me your ear" to mean "pay attention." A more useful metonymic relationship in the domain of photographs might be a picture of Washington, D.C., described as a metonym for the U.S. Government. Subtypes of metonymic relationships could include cause and effect, container and contents, possessor and the thing possessed, and an occupation and its sign. In the example given earlier, Greisdorf and O'Connor identified the terms *boat* and *fishing* as metonymic characteristics of the image picturing a large body of water.

Closely related to metonymy are synecdochic relationships, which involve using a part to represent a whole as in "red coat" standing in for "soldier in the British Army." The authors also consider the participants in their study assigning color attributes to black and white images as evidence of synecdoche.

2.2 DISCUSSION

This chapter reviewed relevant literature on relationships in the context of three categories of image description: natural language, machine-readable, and phenomenal. The literature demonstrates that categorical relationships are critical aspects of the expressions used in describing and identifying objects and relationships among objects in documents including image descriptions. Researchers, however, are generally not concerned with how relationship information can be expressed in precise terms. For example, credit can be given to Farradane for attempting to make clear his assumptions about concepts and relationships, developing a wider range of relations than previous research in LIS had done. His family of nine relations, however, is not an empirically based account of the relations expressed by searchers. Furthermore, his definitions are informal and sometimes ambiguous, yet they are intended to represent meaning among all concepts. There is no evidence suggesting that Farradane was aware of or interested in building on the relationship classification work of his contemporaries in other disciplines, for example, Cruse (1979), Flavell and Flavell (1959), Lyons (1977), Miller (1969), Perfetti (1967), Riegel and Riegel (1963), Rosch & Mervis (1975), and others.

The traditional description models focus more on image attributes as a means of representing semantic information about image content. Shatford Layne's treatment of relationships as an image attribute represents an early attempt at making relationships in the domain of image cataloging explicit in machine-readable records. While her examples do not achieve the level of refinement that other attributes are given in machine-readable records, they do present interesting observations about relationships that can obtain between images in different formats and between image and text.

Svenonius focuses on the propositional use of language for describing image. Her conclusions raise questions about image attributes that go beyond subject, author, and title. She explores, for instance, the nature of *aboutness* and its relationship with subject, concluding that it is a property of the subject attribute.

In general, Svenonius invites representation models that go beyond the facts presented in the original utterance, models that include inference. In the *Afterword* to her book (Svenonius, 2000), for example, she describes intelligent search engines with inference abilities as, “harboring the seeds of progress” (p. 197). This, too, is relevant to the current study of relationship types where the eventual goal is to inform ‘mechanically embodied intelligent processes’ (Smith, 2004). This is similar to Svenonius’s descriptions of automated subject determination where machines analyze words and word phrases based on concepts rather than words, and systems that utilize *is-a* hierarchies to deduce specific subjects from general topics (Svenonius, 2000, p. 49).

How many relations must be accounted for is unclear, yet Green contends that the number and complexity of relationship types makes it an impractical task expecting information professionals to consistently and effectively apply relationships. The current study tests this contention. Green does not discount the potential for future applications of relationships to organizing knowledge, however. She acknowledges there is a disconnection between humans’ intuitive understandings of relationships—an effortless process—and the problems of representation and processing relationships in a computational environment. Considering the quantity of information available to searchers, the rate at which it is growing, Green holds out hope that formalized representations of relationships in computational systems are “perhaps our best hope for infusing higher quality into our retrieval systems” (Green, 2001, 14). She notes that relationships can both aid in finding relevant materials and as a filter identifying less useful materials. Green’s most provocative assertion is one familiar in the field of AI and concerns the notion of reasoning, that it is through reasoning over relationships types that the discovery process enables searchers to find information that would otherwise go unnoticed.

Studies by Enser recognize the importance of using image searchers’ queries to learn about how image are described and how this should determine indexing and subject classification in information systems. However, existing frameworks still lack the expressive power to represent the rich semantic relations that exist in visual information and image descriptions. In spite of these limitations, Enser's continuum does make explicit Shatford Layne's earlier call for a finer-grained approach to subject indexing. Enser hints at making several relationships explicit, especially in the semantic field of spatial expressions. The current study intends to devote considerable attention to this category of relationships because their semantics hold the key to understanding a wide variety of other semantic fields.

Turning the attention towards phenomenal description and the nature of queries, researchers begin to ground their understanding of relationships in the language of image searchers. Enser (1993) performs the first major study examining nearly 3,000 requests received by the Hulton Deutsch Collection. There is a rich set of sample queries that express image attributes and relationships included in the appendix of this paper and in Enser and McGregor (1993). The only relationships that are analyzed, however, are those that are detected between user type and the request categories "unique," or "non-unique." No analysis is given on the relationships among concepts present in queries.

The research reported in Enser and McGregor (1993) makes evident a rich variety of relationships present in the Gibbs-Smith classification scheme and cataloging practices of the Holton Collections. While only the three cross-reference relations are made explicit and the rest remain implicit, this study provides an early view of the role relationships play in what is claimed to be Europe's largest picture library. The report includes several hundred queries gathered together by the researchers. Whether these queries are the original words of searchers or formulations constructed by Hulton's picture researchers, they demonstrate the presence of a variety of relationships among image attributes.

Keister introduces a model that employs a combination of controlled vocabulary keyword searching on the first pass and then visual browsing of surrogate images on the retrieved record set to see which images convey “most effectively the desired message,” recognizing the problem of conveying image content in words. She introduces the notion of an “image construct query,” which views queries as attempts to build images with words. Most importantly, Keister, like Enser, provides an empirically based account of image searchers’ reliance on semantic relationships when constructing queries.

Bean concludes that beyond relationship recognition, it is also important to determine how relationships work. In other words, determine what properties relationships possess, such as transitivity, symmetry, and reflexivity, “because of the increased storage and retrieval efficiency that can be gained by applying subsumption and transitive ‘inheritance’ principles” (Bean, 1996, pp. 85-86).

Armitage and Enser (1997) provide an empirical study that gathers together a semantically rich data set expressing both image attributes and relationships among attributes. They introduce the mode/facet matrix as a function that generates codes for explaining categories of image content, claiming that there are only twelve possible codes that are available for classifying all image content. While relationship types are not explicitly examined, they exist implicitly in the matrix and they included a generous sample of mediated and unmediated queries from all seven libraries in the appendix. The mode/facet matrix could provide a valuable tool for generating descriptions in the context of composing catalog descriptions of photographs—that is, serve as a kind of idea generator.

Finally, Greisdorf and O’Connor argue that by describing image content within a framework of ‘figures of speech’ traditional indexing systems can be extended by generating an array of associations and meanings between entities not visible in the image. The researchers' goal, in effect, is to represent in image descriptions what image searchers "see." They claim that by applying these relationship types to search and indexing systems searchers can achieve more meaningful and relevant search results.

It remains that little is known about the nature and scope of relationships expressed in the contexts of describing, searching, and retrieving photographs either in manual or automated information systems, or the intellectual problems posed by these activities and the resulting need to express or discuss different kinds of objects and processes. It seems likely that a complete understanding of relationships represented in image descriptions must include the activities of catalogers, image searchers, and curators situated in the social milieu of archives. It is the goal of this study, therefore, to address this problem by investigating how humans express relationships among image attributes in the *context of archival discourse*.

3.0 RESEARCH DESIGN AND METHODOLOGY

The literature review presents current understandings of relationships in image content from the standpoint of description. When researchers describe image content in photographs, reference is made to image attributes holding relationships with other image attributes. To represent this knowledge in computational systems—the ultimate goal of image search and retrieval systems and computational ontology—the mechanisms that govern the expression of relationships and attributes in language must be identified.

Chapter three outlines the research methodology used for collecting, managing, and analyzing qualitative data. Qualitative data in this study means text—occurrences in language of facts taking place in the institutional discourse of the Pittsburgh Photographic Library that identify relationship types and instances of relationships among image attributes. Text occurs as photograph descriptions recorded on catalog cards, in letters, and in email. Text also means narratives and correspondence—narratives about research projects, narratives about memories of scenes, and business correspondence from publishers and authors.

Tesch (1990) distinguishes between a *linguistic tradition* that treats text as the object of analysis and a *sociological tradition* that treats text as a window through which human experience can be viewed. The methodology applied in this study is concerned with both viewpoints. As explained in section *3.3.2 Context*, which frames this study within the *context of archival discourse*, there are two viewpoints explored in this project. One is concerned with nonlinguistic entities, the milieu of institutions, collections, facilities, donors and curators. This study is also concerned with the semi-formal linguistic analysis of semantic relationships. For example, which word in the phrase, “cows in a barn” is the *located entity;* which word is the *reference entity*; and which word assumes the role played by *location* in the semantic field of spatial relationships.

The chapter is divided into the following sections: Section 3.1 lays out the fundamental research questions and related, more specific questions. Section 3.2 presents a detailed overview of the methodological approach applied in this study. It introduces the Pittsburgh Photographic Library and explains the *context of archival discourse*, the units of analysis, sampling method, and analysis methods. Section 3.3 discusses the unique nature of the data collected for analysis and Sections 3.4 and 3.5 address validity and reliability.

3.1 RESEARCH QUESTIONS

The manifest behavior produced by photographs is to solve subject classification and description problems for catalogers and visual searching needs by users. The central object of interest in this study is the image content and the structure of descriptions that manifests relationship types and instances of relationships. This leads to the fundamental question

What kinds of relationships do humans express when describing image content?

Individuals building ontologies, knowledge-based systems, and image retrieval systems, as well as individuals working in question-answer systems, machine translation, and natural language processing need the answers to this question. However, as the literature review demonstrates, very little is known about the nature and role of relationships in image description. Ultimately stakeholders are faced with the representation problem, representing image content and the behavior and belief brought to viewing, interpreting, and describing.

Therefore, the goal of this research is to compile a corpus of relationship instances generated during various description tasks carried out by image searchers, curators, and catalogers. This corpus will include the component parts of relationships suggesting their roles within each instance as well as an informal ontology of relationship types.

The two overarching questions and sub-questions that are examined in this study include:

(1) What kinds of relationships do humans express when describing image content in photographs?

a) What relationships do Pittsburgh Photographic Library catalogers instantiate in their photograph descriptions?
b) What relationships are instantiated in the content of queries written by image searchers using the Pittsburgh Photographic Library?
c) What relationships do Pittsburgh Photographic Library curators express in their written correspondence with image searchers?
d) What are the similarities and differences in the use of relationships in describing tasks by Pittsburgh Photographic Library searchers, catalogers, and curators?

(2) What relationship types or *categories* emerge from the corpus of instances?

a) How can relationship instances expressed during description tasks be classified into types?

b) What properties do relationship types typically embody in visual information description tasks?

c) What is the frequency of occurrence of relationship instances within each category?

3.2 METHOD

The Method section explains the critical components of the overall research design. Humans have a general understanding of relationships when they occur in everyday language—for instance, in the picture caption described earlier that accompanies the photogenic drawing in Talbot's *Pencil of Nature*—without being able to pinpoint their occurrences or explain in precise terms what *type* a particular relationship is and what it means. Therefore, this study begins with the basic problem of discovering what humans express when describing image content in photographs. The next two sections introduce the research site and the meaning of context as it is applied in this study. This is followed by explanations of the units of analysis and methods of data collection and analysis.

3.2.1 The Pittsburgh Photographic Library

The Pittsburgh Photographic Library (PPL) originated in 1950 under the direction of Roy Emerson Stryker—economist, historian, and photograph documentalist best known for his large-scale photographic project with the Farm Security Administration during the Great Depression in the United States. R. H. Fitzgerald, the chancellor of the University of Pittsburgh at that time, provided partial financial support for the project as well as darkroom facilities and office space on the thirty-sixth floor of the University's Cathedral of Learning ("Proposal," 1949). The PPL came into existence during a time when Pittsburgh was undergoing a significant rebuilding program. When Fitzgerald officially announced the library's opening he noted, "We are glad pictures can be taken to record these changes, so Pittsburghers today and tomorrow may see the progress made in the city" ("Pitt Plans," 1950). During the years 1950-1955, Stryker and the photographers assigned to the PPL project captured thousands of images of the building renaissance, social, and cultural activities going on in Pittsburgh at that time. The early provision of photographs for use in newspaper and magazine stories, talks, pamphlets and other publications remains a central mission of the PPL today.

In 1960, the PPL collection moved from its original location in the Cathedral of Learning at the University of Pittsburgh to its current location at the Carnegie Library of Pittsburgh. The pictures that Fitzgerald originally saw as vehicles for recording change now form the archives' core collection, which is comprised of 12,669 medium-format negatives and 3,899 35-millimeter negative strips with four to six frames per strips. There are several smaller collections that have been added over the years, including the work of a *Pittsburgh Sun-Telegraph* photographer, Frank E. Bingaman, a collection consisting of over 1,000 photographs taken during the first quarter of the twentieth century; the Abram Brown Collection that includes local scenes, especially of the Pennsylvania Railroad for which he worked; and the Stefan Lorant Collection that consists of copywork of the images which appeared in his book: *Pittsburgh—the Story of an American City* (Lorant, 1964). Another body of Stryker's work held by the Pittsburgh Photographic Library is the collection of copywork that Stryker assembled for *A Pittsburgh Album*, originally published in 1959 as part of the city's 200th anniversary celebration.

There is no clear record or complete inventory showing just how many named collections exist within the PPL, but an estimate was made in March 2000, that the total holdings for the Library numbered 57,008 prints, 58,292 negatives, 1,234 slides, 310 lantern slides and 13,000 contact prints ("Pittsburgh Photographic," 2000).

Gilbert Pietrzak, the current curator for the collection, is responsible for appraisal, cataloging, outreach, preservation, rights management, and reference services. The Carnegie Library Pennsylvania Department--the local history department of the Library—provides additional assistance. Use of the Library has been substantial over the years. Prints from their collections appear in books (Schulz, 1999; Stryker & Seidenberg, 1959), newspapers (Thompson, 1959), film ("An Evening," 2003), and journals (Benson, 2010).

The historic collections and administrative files maintained by the PPL are likely to match those typically found in other picture archives. Like other institutions, the PPL's administrative records come into existence to serve practical and immediate purposes. Current year's records are kept in the curator's office and back files are stored in the archives. Contents of these files include requests for permission to publish photographs, requests for biographical information, and correspondence concerning donations, along with memos and other papers documenting daily operations. What are of primary interest to the current study are the longstanding historic files relating to image searchers' requests. Considering the amount of detail associated with each image request—a set of papers stapled or paper-clipped together as a record of what transpired during individual transactions—there is reason to believe the PPL is a prime source for supporting a study examining image searching in the same archives over a period of time.

3.2.2 Context

The word "context" in LIS is ambiguous and requires an explanation to clarify its use in this study. In one sense of the word, its meaning concerns the linguistic text that stands in for visual information; it is the discourse that surrounds the image searcher's query, the curator's response, and the cataloger's description. It is about the language inscribed on the fronts and backs of individual photographic prints; the text in a donor's letter describing a photograph's provenance; and the subject-headings list used for classifying and filing photographs by subject.

In another sense, context in this study concerns a nonlinguistic entity—the environment or domain of interest—the milieu that includes the institution "Pittsburgh Photographic Library," and the artifacts it collects including photographic prints and negatives; the picture card-catalog records; the archival space with its storage facilities and physical systems for labeling and filing negatives and prints; and the library staff, persons using the collections, donors, and photographers.

These two senses of the word are referred to as the *context of archival discourse* and it is within this context that the central focus of this study is situated. Adopting this definition of context, however, still leaves open the possibility of an infinite number of complimentary approaches for describing an infinite variety of phenomena within the *context of archival discourse*. Further distinctions are made regarding the central theme of this investigation by making explicit the method of generating descriptions of the phenomenon of interest. The following sections explain the analytic framework and units of analysis utilized in this study.

3.2.3 Content analysis

The primary task is to identify the relationships expressed in the context of the Pittsburgh Photographic Library (PPL) reference transaction. In this study the collection of papers associated with a single transaction is an *incident*. The transaction event begins when a visual researcher submits a written query to the photograph curator of the PPL.

Image searchers' mediated requests are not directly observed, so the written requests and correspondence between curator and researcher invite content analysis, as do the visually- and textually-rich catalog card records. This study adopts Klaus Krippendorff's definition of content analysis: "Content analysis is a research technique for making replicable and valid inferences from texts (or other meaningful matter) to the contexts of their use" (2004, p. 18). Krippendorff asserts that texts attain significance in the context of their use. Significance in this study is measured along various dimensions, including contents, interpretations, and meanings. Contexts are constructed for analyzing recording units using the theoretical frameworks provided by Cooper's analysis of English locative prepositions (1968), Veda Storey's analysis of case relationships (1993), IDEF5's class inclusion relationship library (Perakath, 1994), Winston, Chaffin, and Herrmann's taxonomy of part-whole relations (1987), and Wand, Storey, and Weber's ontological analysis of relationships (1999). These theoretical constructs play a role in constructing the environment in which analysis takes place. They help explain how relationships are determined based on particular words that are used to infer the relationships.

There are two general types of content analysis applied in this study. In one, the text in the reference transaction is segmented into its basic meaning components: semantic units, factual assertions, relationship instances, and words, all of which are defined in section *3.3.4.1*

Definitions. In the other, meaning is found in large blocks of text to aid in distinguishing between searchers' queries, curators' correspondence, and catalogers' descriptions.

Content analysis is used to classify the relationships according to their meanings. The analytic process is both quantitative and qualitative. Counting and making explicit the times certain types of relationships occur measures frequencies. Determining what kinds of relationships are captured in the corpus is a qualitative process. There are two broad categories of relationships being observed: ontological and instanced. *Ontological* analysis is used to guide decisions about what relationship types exit in reality. These are the *ontological* relationships that constitute broader categories or *types* in which instances can be classified. The semantic field of "spatial expressions" is an example of ontological relationships. *Instances* of relationships constitute relationships that obtain between entities in the domain of interest—the context of photograph libraries. For example, analysis of the picture caption "Peaches in bowl" would reveal an *instance* of the relationship type *"spatial*." The expression describes a locative entity "peach," a spatial relationship instance "in," and a reference object "bowl."

It is usual in content analysis to refer to other sources for theoretic constructs and detailed descriptions to ensure the consistent and coherent account of relationships. There is not a single comprehensive account of semantic relation types in English, so it has proved necessary to bring together many sources to attempt providing this framework. Nevertheless, the set of relationships is selective rather than exhaustive. These are explained in the section *3.3.7 Ontological analysis*.

Krippendorff maintains that the chief advantage of the unstructuredness of content analysis data is that it preserves the conceptions of the data's sources, which structured methods largely ignore" (Krippendorff, 2004, 41). A query sent to a photograph curator, the analytic context of the current study, is a social process transacted through a meaningful, written request for photographs. Analysis within this context explains how these texts came to be, what they mean, and what they can do. There are other equally valid contexts and other readings in which these same texts could function as well, for example, readings by photograph historians or analysis by visual artists.

3.2.4 Units of analysis

There is an order of succession in the reference transactions analyzed in this study: Query submitted by searcher; response from curator and mediation during retrieval. Depending on individual circumstances, communications may continue between curator and searcher until the desired images are retrieved. Krippendorff describes narratives like this as streams of text that interact with each other (Krippendorff, 2004). The structures of interest in this study of relationship types and instances of relationships involve several streams of text manifested in vocabularies of words, prepositional phrases, sentential constructions, and larger textual units. The largest textual units of concern here are *bilateral incidents* and *unilateral incidents*. These and the other units of analysis adopted in this study are distinguished for separate recording and coding.

3.2.4.1 Definitions

This section defines the key terms used in this study for describing units of analysis.

Sample unit: the units that are distinguished for selective inclusion in the analysis. The sample unit in this study is the *Incident.*

Incident: a collection or set of documents comprising a reference transaction. In this study an incident typically consists of a set of paper documents stapled or paper-clipped together, communications from a searcher, hand-written, typed, or sent via email, and a written response from a photograph curator. In some cases the curator responding to the query makes copies of picture card catalog records that include images and descriptions and these become part of the incident.

Bilateral and Unilateral Incidents: There are two types of incidents: (1) unilateral and (2) bilateral. The *unilateral* incident consists of an image searcher's written query and may or may not include a record of the curator's written response. If the curator's response is included in the incident, there is no visual evidence of the photographs or descriptions selected to satisfy the searcher's request. The *bilateral* incident includes the image searcher's written request, correspondence from the curator, and copies of the photographs and cataloger's descriptions retrieved and given to the searcher in response to his/her query.

Recording unit: the unit distinguished for separate recording or coding. A unit that is separately described or categorized and contained within the sampling unit, coinciding with, but not exceeding the sampling unit. There are two main recording units in this study: the Semantic Unit and the Factual Assertion. There are numerous finer-grained recording units that mark attributes associated with Factual Assertions. These are defined in the Coding Instruction Book.

Semantic unit: a course unit of analysis. A discourse structure that segments the original description into sentences that describe a person, object, event, or state of affairs represented in a photograph description. Semantic units are direct quotes from incidents. The language of semantic units is unique in the context of archival photographs. That is, they express ideas unique to photographs typically found in archives. The length of semantic units range from minimal descriptions, sometimes consisting of single words or dates, representing subjects and capture dates to complete sentences and paragraphs.

A sentence or paragraph may consist of several individual instances of semantic units. Herskovits provides a useful model for determining the lower limits of size as it pertains to spatially related expressions (Herskovits, 1986, 7). These are explained in the Coding Instruction Book as they pertain to locative expressions.

Factual Assertion: a fine-grained statement re-presenting a basic sentence structure found in a *semantic unit*. A *factual assertion* typically represents a predicate with two arguments, for example, "Dorothea Lange employed by Roy Emerson Stryker" or "John standing near river." In this study the term *predicate* is used as a synonym of the term *relationship.* If a factual assertion has only one argument, for example, "subject fishing" or "capture date 1936," it is considered an attribute. Pragmatic inference guides attempts to expand these unary predicates into binary predicates. For instance, in the context of a given incident, "subject fishing" may be expanded to become "photograph subject-of fishing" and "capture date 1936" maybe expanded to "photograph capture-date-of 1936."

3.2.5 Data Selection

This section explains two different data selection procedures. Statistical random sampling is applied to select a representative sample of reference incidents and corpus construction is applied as the principle for data collection within individual reference incidents.

3.2.5.1 Sampling

The Pittsburgh Photographic Library maintains a paper-based correspondence file going back to 1963 that contains 1,673 documents. The correspondence addresses a variety of topics including general reference, donations, licensing rights and permission to publish, and questions from individuals seeking photographs. There are a total of 180 photograph requests, which form the Corpus Data Set for the current study. Breakdowns for these 180 requests are shown in Table 5.

Table 5: Breakdown showing the total number documents in the Pittsburgh Photographic Library administration files representing reference questions and the subtotal relating to photograph requests.

Year	**Total Documents**	**Total Photo Requests**	**Bilateral Incident**	**Unilateral Incident**
1963-1979	15	0	0	0
1980-1989	209	14	2	12
1990-1999	48	3	1	2
2000	0	0	0	0
2001	85	24	15	9
2002	138	73	25	48
2003	15	0	0	0
2004	230	0	0	0
2005	308	23	2	21
2006	144	0	0	0
2007	141	15	0	15
2008	166	10	1	9
2009	174	18	0	18
TOTAL	1673	180	46	134
%		100%	26%	74%

The Corpus Data Set of 180 documents is divided into the two relevant subgroups: *unilateral* and *bilateral incidents*. Of the 180 incidents, 26%, or 46, are *bilateral*; 74%, or 134, are *unilateral*. Using a table of random numbers, a sample is randomly selected based on 25% of the total, or 45 incidents. The variable of interest is incident type, so to achieve proportional representation, 74% of the sample (33 incidents) is drawn from the pool of unilateral incidents and 26% (12 incidents) is drawn from the pool of bilateral incidents. From this sample of 45, random selection is applied once more to achieve a proportional representation, 20% of 45 incidents, or nine, for the pilot study. The remaining sample to be analyzed as part of the dissertation research consists of 36 incidents, or 20% of the total population of 180 incidents. The procedure described here is illustrated in Figure 20.

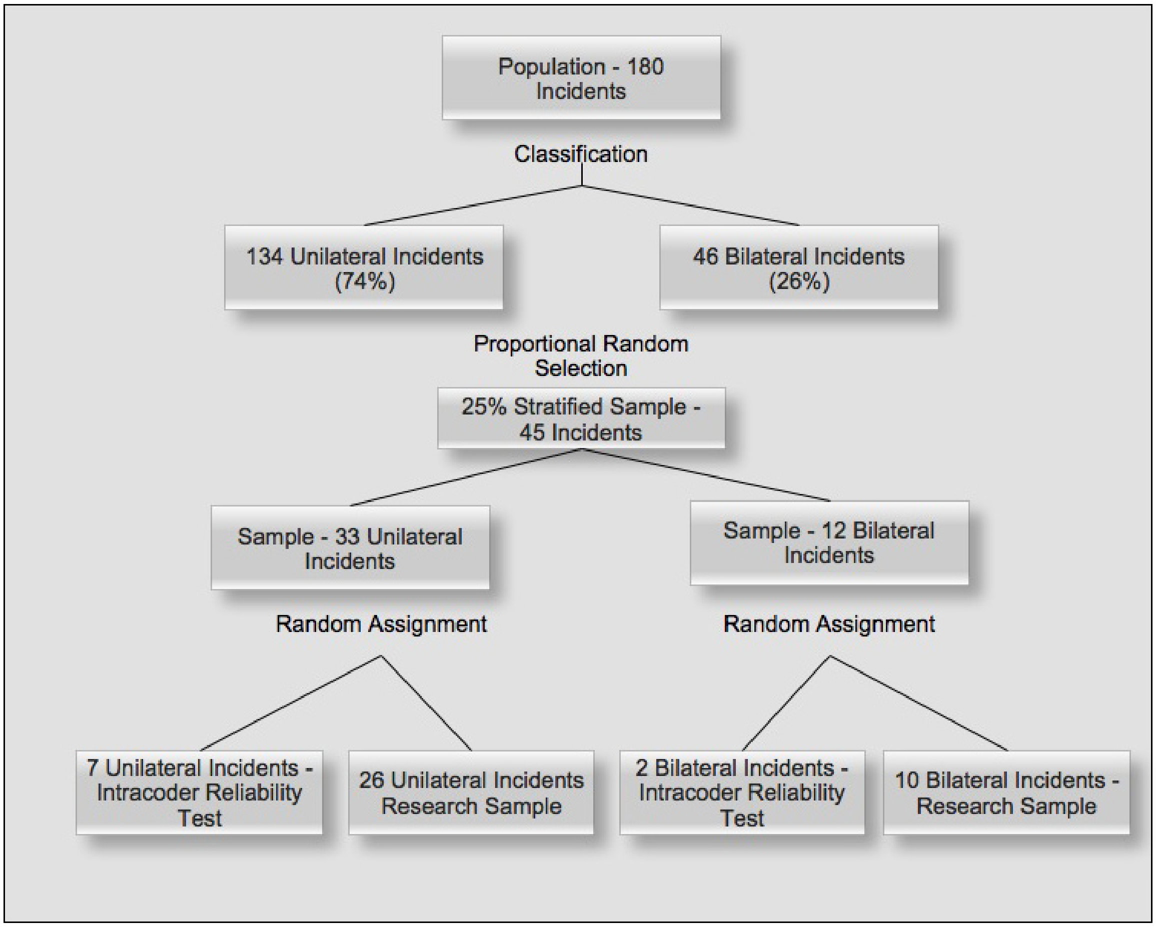


Figure 20: Procedure for selecting a stratified sample based on incident type.

3.2.6 Corpus Construction

A linguistic corpus is "The body of written or spoken material upon which linguistic analysis is based" (*Oxford English Dictionary Online*, March 2011). The corpus being compiled in this study, referred to as the *Relationship Corpus*, is designed for a narrowly defined purpose: to aid in the analysis of relationships expressed as part of the discourse in the PPL. The corpus is manually constructed and proceeds through three stages of parsing.

Stage One

In the first stage of parsing a careful reading is made of the entire content of the incident. A determination is made regarding what portion of text describes image content and those texts are recorded as *semantic units* and assign *semantic unit numbers (SEMNL#)*. Codes are explained in the Coding Instruction Book, Appendix A. To maintain homogeneity among groups (catalogers, searchers, curators) the text is separated from these groups into different corpora for comparison.

Stage Two

The second stage of parsing involves analyzing the corpus into *factual assertions* representing relationships. Each assertion is assigned a *factual assertion number (FACTS#)*, which corresponds with the specific *semantic unit* from which the assertion is derived.

As an example, Figure 21 shows a sample of semantic units and corresponding factual assertions developed during the pilot study described in chapter four. This second stage of parsing is designed to capture a comprehensive range of relationship instances.

Column 4	Column 5	Column 7	Column 8
013SEMNL10	So, what I'm looking for is more recent--high-quality images of Michelle Madoff (founder of GASP and later city councilwoman);	013FACTS1032	subject_of("Michelle Madoff",photograph)
		013FACTS1033	founder_of("Michelle Madoff","GASP")
		013FACTS1056	has_agent(found,"Michelle Madoff")
		013FACTS1057	has_theme(found,"GASP")
		013FACTS1034	occupation_of("Michelle Madoff",'city councilwoman')
		013FACTS1035	synonym_of(image,photograph)
013SEMNL11	Images of the Allegheny County Air Pollution Control Board Variance Board meetings in the Gold Room of the City/County Building;	013FACTS1136	subject_of("Allegheny County Air Pollution Control Board Variance Board meetings",photograph)
		013FACTS1137	in(meetings,"Gold Room")
		013FACTS1138	synonym_of(image,photograph)
		013FACTS1139	part_of("Gold Room","City/County Building")
		013FACTS1140	part_of('city building','county building')

Figure 21: Sample from pilot study Relationship Corpus illustrating original semantic units (Column 5) and factual assertions (Column 8) derived from the original statement.

As the analysis progresses forward moving from natural language semantic units to more explicit variants of description, a distinction must be drawn between named entities (e.g., *Andrew Carnegie* is a person and *Pittsburgh* is a place) and *universals* or *types*, which can be said of many things (e.g., *music*, *steel mill*, *city*, and *striker*). The distinction between the name of the individual "Andrew Carnegie" and the referent or projection in the real world of an individual who goes by that name requires extending the layers of meanings to include text.

When restructuring semantic units into more precise factual assertions, proper nouns and phrases containing proper nouns, such as *Andrew Carnegie*, dates, accession codes, and any other named entities, are distinguished by placing them in double quotes. For instance, *death_date_of("December 18, 1882","James Laughlin")* and *subject_of("James Laughlin",photograph)*. When universal or kinds, such as music, railroad, and worker, occur in a phrase, the phrase is placed in single quotes. For example, *subject_of('suffragettes petitioning',photograph)*.

The content analyst also resolves anaphoric references (personal pronouns). For example, in the semantic unit

"Andrew Carnegie's anteroom leading from his Library to Office. His home at 2 East 91st Street, New York."

"His" is anaphoric (it refers to Andrew Carnegie).

<u>Stage Three</u>

The third and final stage of parsing involves analyzing the factual assertions to determine what types of relationships are present and what are their component parts (themes, reference objects, places, paths, subjects, temporal units, and so on).

Schmied's (1990) *corpus theoretical paradox* is addressed by adopting a cyclical process approach. In general, Schmied's paradox is a *chicken or the egg* dilemma. The Relationship Corpus resulting from this investigation should be representative of the language used in photograph libraries. The corpus should reflect all the variables that determine variations in that community. The problem, however, is that in order to accomplish this task the researcher must first determine these variables empirically through analysis of results from a representative corpus. So the questions are where to begin and when to stop.

The method for addressing this paradox adopts Biber's view that corpus design is a cyclical process (Biber, 1993, 256). In the current study it began with a pilot empirical investigation and analysis outlined in chapter four, which was conducted during May through September 2010. The cyclical improvements made during the pilot study were documented, extending and refining the scope of variables until no additional varieties could be detected. The *stopping criteria* as described by Bauer and Gaskell (2000, pp. 32-34) determines that once saturation is reached it is time to stop the cycling process. Saturation occurred in the pilot study when adding further relationship types and instances and the variables defining them made only a small difference with regard to additional representations of semantic units and factual assertions.

Figure 22 depicts the basic components of the three stages described earlier that are involved in constructing a Corpus of Relationships and ontology of relationship types.

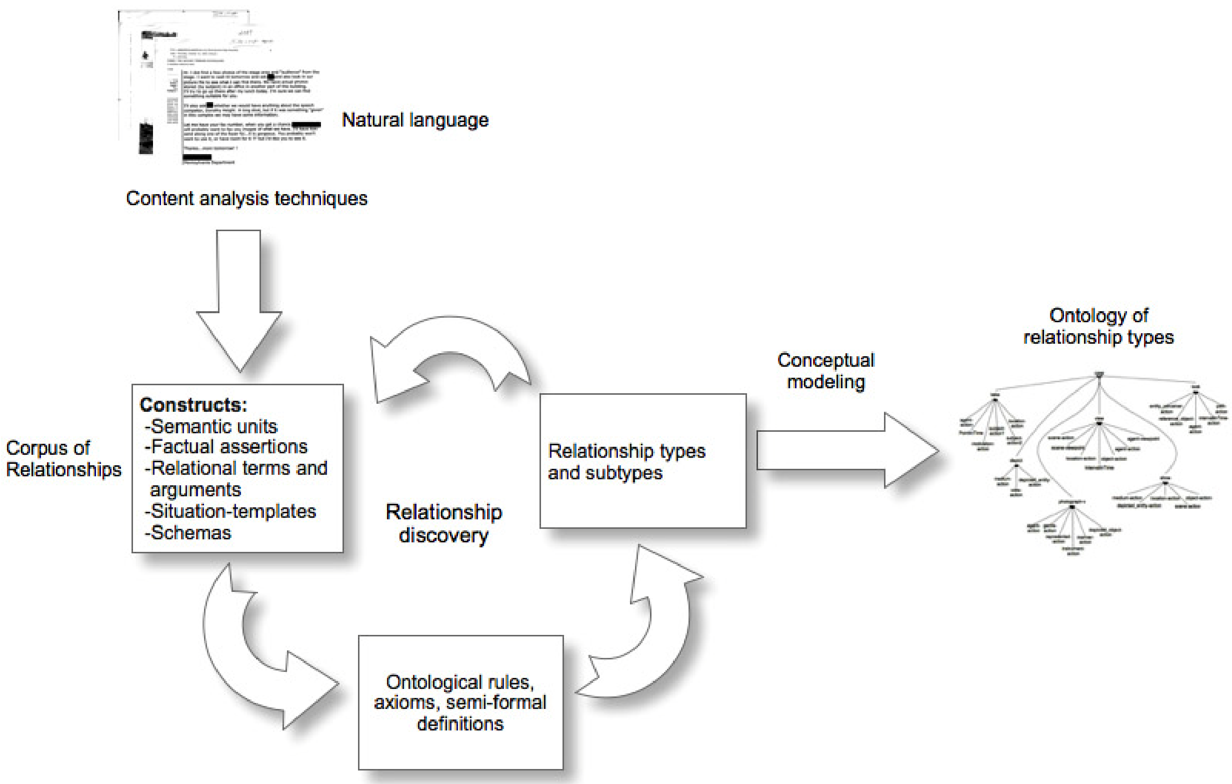


Figure 22: Architecture of the Corpus and Ontology Construction

3.2.7 Ontological analysis

Ontological analysis of relationships guides the coding activities during content analysis and the subsequent modeling of an ontology or *family tree* of relationship types and instances. The analysis thus far has been exercised on the ordinary English-speaker's use of relationship types emerging in a pilot study (introduced in Section 3.5). Through an iterative process of analyzing incident content and evaluating the work of researchers engaged in defining the relationship construct, the following categories of relationships emerged as being predictors of the types of relationships expressed in the pilot study sample: *attribution, case, inclusion, meronymic, spatial, synonym* and *temporal.* The current study intends to begin building on this ontological framework.

The goal in this study is not to restate what theorists have already explained in detail in their respective papers and monographs. A brief description of each relationship type examined in this study is included in the next chapter. It is helpful from a methodological standpoint, however, to list in Table 6 all the resources consulted and the relationship types of interest in each given resource.

3.3 UNIQUE NATURE OF THE DATA

Analyzing curators' administrative records documenting reference transactions offers a unique opportunity to perform a content analytic examination of the discourse surrounding image descriptions. The data is unique because the PPL incidents are written evidence of events that occur in the context of ordinary cultural performance--that is, the conditions under which image behavior normally occurs (Frake, 2009, p. 133).

All of the text in this study is free-flowing discourse (Denzin & Lincoln, 1994, p. 769). That is to say, none of the material is generated from systematic elicitation, for example, from interviews or questionnaires. The exception may be the catalogers' descriptions, which result from self-imposed elicitations of facts—an interrogation of the photograph. The cataloger asks, “What is the accession code, who is the creator, what are the subjects?” These are short, open-ended questions driven by a necessity to conform to local cataloging standards and practices.

The text examined in this research originates from three independent but interrelated sources: image searchers' queries, photograph curators' responses, and picture catalogers' photograph descriptions. Naturally, creation of a cataloger's description precedes an image searcher's query and the curator's description in correspondence is a response to the searcher's query. The query itself describes the image content of a known photograph or one yet to be discovered, a topic about which little is known.

Table 6. Resources consulted for ontological analysis of relationship types**.**

Code	Resource	Relationship Types
CHA1	Chaffin, R., & Herrmann, D. J. (1984). The similarity and diversity of semantic relations. *Memory & Cognition 12* (2), 134-141.	**II. Class Inclusion** **IV. Case Relationships** **VI. Meronymic Relationships:**
CHA2	Chaffin, R. , & Herrmann, D. J. (1987). Relation element theory: a new account of the representation and processing of semantic relations. In D. S. Gorfein & R. R. Hoffman (Eds.), *Memory and learning*, 221-245. Hillsdale, NJ: Lawrence Erlbaum.	
CHA3	Chaffin, R., & Herrmann, D. J. (1988). The nature of semantic relations: a comparison of two approaches. In M. W. Evens (Ed.), *Relational models of the lexicon*, 289-334. New York, NY: Cambridge University Press.	
GSC1	Cooper, G. S. (1968). *A semantic analysis of English locative prepositions* (Bolt, Beranek, & Newman, Report No. 1587). Springfield, VA: Clearing House for Federal, Scientific, and Technical Information.	**I. Spatial Relationships** (See Coding Instruction Book for details)
CBP1	Perakath C. B. et al. (1994). *IDEF5 Method Report*. Knowledge Based Systems, Inc.	**II. Class Inclusion** 1. Functional 2. State 3. Activity 4. Action 5. Perceptual **III. Temporal Relationships** 1. Time-point 2. Time-interval
VCS1	Story, V. C. (1993). Understanding semantic relationships. *VLDB Journal, 2,* 455-488.	**IV. Case Relationships** 1. Agent-action 2. Agent-instrument 3. Agent-object 4. Action-recipient 5. Action-instrument
WSW1	Wand, Y., Storey, V. & Weber, R. (1999). An ontological analysis of the relationship construct in conceptual modeling. *ACM Transactions on Database Systems, 24* (4), 494-528.	**V. Attribution**
WCH1	Winston, M., Chaffin, R., & Herrmann, D. (1987). A taxonomy of part-whole relations. *Cognitive Science*, 11, 417-444.	**VI. Meronymic Relationships:** 1. Component/Integral Object 2. Member/Collection 3. Portion/Mass 4.Stuff/Object 5. Feature/Activity 6. Place/Area

3.3.1 Catalogers' descriptions

Catalogers create one of the data sets examined in this study. Image descriptions, such as the one accompanying the contact print in Figure 23, are attempts by catalogers to predict what a researcher might want to know about the visual information in photographs.

To use Rothkegel, Wender, and Schumacher's (1998) distinction, the cataloger is describing two kinds of space: *concrete physical space* and *psychological space*—spatial concepts that are particularly relevant to human cognition. An infinite set of locative entities and relationships between them exist in the *physical space* of the Director's office shown in this image: ceiling lights and their position relative to objects in the room and the resulting shadows that are cast on the floor; four individuals sitting at desks, each in a different location, and Mrs. Ann C. Hall standing in the doorway.

During picture making, the 3-dimensional space of this scene is collapsed onto a flat, 2-dimensional space on the film's surface, which is later developed and projected onto the surface of a piece of photographic paper. This photograph exists in a different context and a different spatial environment than the physical space in which Lou Malkin shot the image. The picture that lies upon the surface of the print—a model of a pre-exiting 3-dimensional space—is perceived and described in *psychological space*. This is the channel through which the cataloger perceives space. It is the physical space of Director Martin's office as perceived by the cataloger.

Figure 23: Lou Malkin. *Carnegie Library – Director's Office*, December 17, 1973. From the Pittsburgh Photographic Library Picture Catalog. (Reprinted with permission.) Carnegie Library of Pittsburgh. All Rights Reserved. Unauthorized reproduction or usage prohibited.

A third spatial environment that is of particular interest to visual researchers exists outside the frame of the photograph. The camera's physical spatial position during the making of the photograph provides the image with a point-of-view (POV). POV transfers a spatial concept to a non-spatial domain outside the frame of the photograph, a dimension Habel and Eschenbach (1997) call *metaphorical space*. While POV is not made explicit in the photograph's description, it affects how the viewer reads and interprets which desk is located *first in the foreground*.

3.3.2 Researchers' correspondence

Visual researchers provide a second source of data examined in this study. The sample correspondence illustrated in Figure 24 illustrates the complex and semantically rich nature of visual researchers' queries. Like the cataloger, little is known about the levels of meaning expressed by the image searcher when describing relationships among image attributes in this context.

-----Original Message-----
From: [redacted]
Sent: Tuesday, January 29, 2002 12:14 PM
To: [redacted]
Subject: Information needed

Dear [redacted]

You may remember meeting me in Köln a few years ago when I was there wit
the Chicago Symphony Orchestra. Your brother introduced you to me and to
me of your position with the Carnegie Library of Pittsburgh. I mentioned
you that at one time I had contacted the library about obtaining a copy
photograph of the old Pittsburgh Orchestra and never received an answer.
told me to contact you and you would see what could be done about it. I
forget the date of the photo, but it was the full orchestra on stage.
Actually, if there exist more than the one full orchestra photo I would
interested to know that and would probably want to order a full copy of
each, if that service is available. Do let me know the details.

Another question came up. In looking through a complete run of the Music
Courier I found mention of a new monthly music periodical having been is
in Pittsburgh beginning in March, 1911, called "The Symphony." Was this
indeed published and do you have copies of it? Thank you for looking int
these matters for me. My best wishes to you for a healthy, happy, and
successful 2002!

Sincerely yours,
[redacted]

Figure 24: Sample correspondence from visual researcher requesting photographs from Pittsburgh Photographic Library.

3.3.3 Curators' mediation

Finally, there is the curator who mediates between the searcher's query and the cataloger's description and retrieves from the collection what s/he deems relevant to the search. This investigation attempts to make clear what types of semantic relationships curators decode and re-encode in their responses. Spatial expressions such as *right of* and *left of* are known to have indexical uses where the reader or hearer must determine the relation by considering the location of an implicit viewer or camera. Figure 25 illustrates an example of correspondence received by the curator and how the curator must determine a photographer's point of view placing a camera "on the stage area" viewing the "audience from the stage."

> Hi: I did find a few photos of the stage area and "audience" from the stage. I want to wait til tomorrow and ask [redacted] and also look in our picture file to see what I can find there. We have actual photos stored (by subject) in an office in another part of this building. I'll try to go up there after my lunch today. I'm sure we can find something suitable for you.
>
> I'll also ask [redacted] whether we would have anything about the speech competor, Dorothy Height. A long shot, but if it was something "give in this complex we may have some information.
>
> Let me have your fax number, when you get a chance. [redacted] will probably want to fax you images of what we have. I'll have him send along one of the foyer fyi...it is gorgeous. You probably won't want to use it, or have room for it ?? but I'd like you to see it.
>
> Thanks...more tomorrow! !
>
> [redacted]
> Pennsylvania Department

Figure 25: Sample correspondence from Pittsburgh Photographic Library curator in response to image searcher's query.

3.4 VALIDITY

Validity is concerned with the correctness of the inferences that are made regarding relationships discovered in the incidents. An attempt is made in the pilot study to consistently and systematically measure the correct relationships occurring in sample incidents and consistently classify instances with their corresponding types. Various analytical constructs introduced in §*3.3.7 Ontological analysis* are employed in this study to assist in this process. The final Relationship Corpus can also be evaluated and scrutinized to determine whether the relationships contained therein are consistent with the ontological constructs listed earlier in Table 6, §*3.3.7 Ontological Analysis*. These constructs also form the indexical basis of the content analysis enabling the formation of one-to-one relationships—a mapping function—from the variable to the phenomena the relationships are intended to predict. There is no isomorphism to the relationship types described by the theorists and what catalogers see in photographs or what searchers describe. In other words, no inference is being made that catalogers, searchers, or curators have a certain interpretation of a photograph from the fact that s/he assigned a particular proposition to a category of relationship types.

3.5 RELIABILITY

A preliminary investigation (Benson, 2011a) was carried out during May through September 2010, to refine the analytic constructs for investigating the language of relationships. Analysis was completed on a random sample of nine incidents containing 164 observations or "factual assertions" expressing relationships describing image content. Various theoretical frameworks were applied including Cooper's English locative prepositions (1968), IDEF5 relation library (1994), and Storey's semantic relations (1992) measuring each unit in light of ontological principles and content analytic techniques advanced in recent literature.

A content analysis codebook and forms were completed and intra-coding reliability and instrument validity were tested using test/retest, which entailed measuring agreement on 2,788 ratings using Kappa statistics. The 2,788 ratings are derived from measuring matches and mismatches on 17 different variables for each of the 164 observations. The calculated intra-coder corelationship coefficient for all variables ranged between .394 and 1.0. The strength of agreement for Kappa coefficients in the range of .21-.40 is considered fair and .81-1.0=almost perfect (Sim & Wright, 2005, p. 7).

3.5.1 Coding instruction book

The initial coding instruction book and forms, developed during May-June 2010, are used for coding the sample data set in the pilot study's intra-coding reliability test in July 2010. The content analysis and coding adopts *a priori* coding. That is, a preliminary set of variables are established prior to data analysis based on relationship theory in linguistics, computer science, and Artificial Intelligence (Cooper, 1968; Crombie, 1985; Fahlman, 1979; Herskovits, 1986; Levinson, 1996; Li & Gleitman, 2002; Perakath, 1994; Whorf, 1940/1956). The object of the coding process is to quantify relationships reported to hold between concepts in written image queries, curator's correspondence, and in photograph card-catalog descriptions. Twenty-six variables were analyzed and recorded during the pilot study.

During the first test in July 2010, as coding proceeded through the nine sample incidents numerous revisions were made in the relationship theory resource list and the variables being coded. The final version of the *Coding Instruction Book* is included in Appendix A of this dissertation. Recommended training for future coders consists of the head coder going through the instruction book with all other coders line-by-line explaining each variable and its appropriate response patterns. When this initial training is complete, a test-retest is recommended to determine stability and inter-coding reliability. The pilot project reported in chapter four is an example of utilizing only one coder and using test-retest reliability model to determine intra-coder reliability.

Written permission was received from the administration of the Carnegie Public Library to make copies of and analyze data samples. The original samples are scanned and saved to disk as unmarked copies and printed out for coding. The data collected from the samples is organized and recorded in a single Excel workbook. Personal names are blackened out in the hard copy samples used for analysis. Digital and paper-based copies of all samples and spreadsheet analyses are retained for future use.

3.6 SUMMARY

Chapter three introduces the two overarching questions that are examined in this study. It provides a detailed outline of the research methodology to be used for collecting, managing, and analyzing the qualitative data obtained from the Pittsburgh Photographic Library corpus data. A methodological framework evolves that examines relationships in four ways: One considers relationships individually and independently from others in the study. Another views all of the relationships together within the context of a corpus. Another examines how different groups describe images. A fourth approach views image descriptions as representing objects, events, and states of affairs at various levels of meaning including pre-photographic referents, depicted referents, concepts, and text.

4.0 ANALYSIS OF THE DATA

In order to identify and analyze relationships that occur in image descriptions, it is necessary to first draw a sample of documented incidents as described in chapter three. This chapter describes in detail the sample taken from the Pittsburgh Photographic Library administrative files. The ontology of relations that emerges from the sample is presented in the next chapter. The analysis presented in this chapter is critical for two reasons. First, it is essential to discover the nature and characteristics of relationships humans express and the frequencies of their occurrence. This knowledge forms the basis of the study and informs the modeling of the ontology in the chapter that follows. Second, the analysis helps explain the different contexts in which descriptions are made and this provides a basis for comparing the viewpoints held by catalogers, curators, and image searchers.

From this point forward the following convention is followed: Examples of text taken from the data set and included in the remaining chapters of this dissertation are labeled with the corresponding alphanumeric codes used in the Relationship Corpus. Two different codes require explanation: Semantic Units are coded using the letters SEMNL and Factual Assertions are coded using the letters FACTS. The semantic unit code 011SEMNL40 is to be read: incident number 11, semantic unit number 40. The factual assertion code 034FACTS0309 is to be read: incident number 34, semantic unit number 03, factual assertion number 09.

Chapter four has three main sections following this introduction. The first section briefly states the role played by content analysis in coding and recording the data sample. In the next section, the limitations on depth of analysis are considered. Then in the main body of this chapter, the analysis of the data sample is presented.

4.1 CONTENT ANALYTIC APPROACH

All of the samples obtained from the Pittsburgh Photograph Library are analyzed and coded. Coding systematically identifies, captures, and records relationships that occur in image descriptions. More specifically, the object of the coding process is to systematically record relationship instances, quantify relationship types, and identify certain properties of relationships in English reported to hold between concepts in written image queries and responses and in corresponding photograph card-catalog descriptions. The goal is to conduct these activities in a systematic, principled manner, moving from natural language descriptions to formal assertions only after the relationships are precisely understood.

The Corpus, organized in an Excel workbook, is a record of the original natural language descriptions and a semi-formal accounting of the relationship instances predicted to obtain in these descriptions. The coding for each categorical variable is recorded in the content analysis forms. The Coding Instruction Book instructions are included in Appendix A.

4.2 RESTRICTIONS ON DEPTH OF ANALYSIS

The broad categories of relationship types identified in this study are complex and some are broken down into collections of subcategories. For example, this dissertation adopts *meronymy* or *part-whole* relationships as one of its major categories. Winston, Chaffin, and Herrmann (1987) break this category down into seven subcategories: component-integral, member-collection, portion-mass, stuff-object, feature-activity, and place-area. These seven relationships are further subdivided according to three finer distinctions called *relation elements* (i.e., functional, homeomerous, and separable). An instance of meronymy, therefore, is defined by location within this hierarchy of meronymic features and types. For example, the phrase *a lens is part of a camera* can be described along three dimensions: 1) The lens-camera relation is a *component-integral* part-whole relation where the lens has a *function* with respect to the camera. 2) The lens is not the same kind of thing as its whole, so it is said to be a *non-homeomerous* part of the camera. 3) With some cameras, the lens is *separable* from its whole. In addition to these three special features, this study adopts the relation element view inspired by Chaffin and Herrmann (1984), Stasio, Herrmann, and Chaffin (1985), and Chaffin and Herrmann (1987) as criteria for determining and classifying relationships.

4.2.1 Relation element view

Chaffin, Herrmann, and Stasio's *relation element view* holds that the realtor in a polyadic relationship is decomposable into smaller components and that it is these smaller components that enable analysis and judgments to be made about similarities and differences among relationship types. This is an extension of Winston, Chaffin, and Herrmann's (1987) three features. The relation element view describes and classifies 29 different relation elements under four broad headings: 1) Intentional force (e.g., denotative, connotative elements), 2) Propositional (e.g., attributive, social, and agentive elements), 3) Dimensional (e.g., dimension and unilateral position elements), and 4) Agreement (e.g., inclusion, overlap, and intersection elements).[9] A factual assertion that expresses synonymy, for example, is defined under the *relation element view* along three dimensions: as *intersection* (word$_1$ is semantically included in word$_2$), *general inclusion* (word$_1$ is included in word$_2$), and *bilateral inclusion* (word$_1$ = word$_2$).

Relation elements are utilized in this study only to the extent that they assist in recognizing and classifying relations in the data sample that are not easily classified using formal and semi-formal definitions of relationship types alone.

4.2.2 Primitives in relational assertions

A second concern regarding depth of analysis is word meaning in arguments. This study is primarily concerned with capturing, recording, and sorting relationship instances into families of relationships. It does not undertake explaining the full range of relationship instances and the roles and structures of their arguments. This means that for the purposes of this study and the data analysis carried out in chapter four, words and phrases representing arguments in the Corpus

[9] Relations are presented as word pairs in a sorting study of relational elements conducted by Chaffin and Herrmann (1987) Examples of word pairs include old-young, car-auto, animal-horse, and farmer-tractor. The subjects were given the task of sorting these word pairs into piles of relations that were the same or similar. Results showed that in all but one case subjects sorted on the basis of the similarities of relations, not on the basis of similarity among the related items. In other words, the word pair farmer-tractor was sorted on the subject's recognition of the case relation *agent-instrument* and not on the word senses associated with 'farmer' and 'tractor.' Furthermore, Chaffin and Herrmann's research provides convincing evidence that subjects make similarity judgments based on componential features of relations and not on a unitary relation marker.

of Relationships are accepted as primitive and domain specific. No systematic attempt is made to disambiguate their meaning or analyze their semantic roles, with the exception of case role analysis explained later in this chapter and minor interpretations of misspellings and abbreviations. Taking the long view—the endpoint at which this research is aimed, that is, ontology-based representation in knowledge-based systems—knowledge modelers must not only attend to the relationship, but also to the various ways in which arguments can be encoded, the semantic roles arguments play, and how differences between encodings influence informational emphasis. These concerns are examined only briefly in chapter four.

4.3 DESCRIPTION OF THE SAMPLE

The description of the sample proceeds as follows. The questions posed in the content analysis forms that pertain to natural language expressions and their reconstructions as factual assertions are examined one at a time. Each question is examined in a subsection of § 4.3. The first variables of interest are the two broad categories of incident type: unilateral and bilateral incidents. This is followed by an analysis of natural language semantic units. These are image-related, verbatim expressions extracted from the incidents. Factual assertions are considered next. The analysis includes a thorough description of what constitutes a well-formed relationship in this study and explains the role pragmatics plays in formulating assertions. The next variable of interest is called *recording unit type,* which refers to the different groups describing images, namely, catalogers, curators, and image searchers. This leads into the main body of analysis where 1,655 relationship instances are analyzed based on their groupings into *a priori* relationship types.

4.3.1 Incident types

The sample is divided into two broad categories of incident types: 1) unilateral incidents, and 2) bilateral incidents. Once again, a *unilateral* incident consists of an image searcher's written query and may or may not include a record of the curator's written response. A *bilateral* incident includes the image searcher's written request, correspondence from the curator, and copies of the picture card catalog records retrieved by the curator. As expected, based on the procedure for selecting a stratified sample explained in chapter three, it is indicated in Table 7 that there is a proportionately greater number of unilateral incidents than bilateral incidents. One interpretation of this imbalance is simply as recognition of the curator's idiosyncrasies: some transactions demand copies of every aspect of an incident, including copies of associated catalog records, and some do not. When the curator decided to keep a record of catalog records associated with a reference query, that transaction, by this study's definition, became a bilateral incident.

Table 7: Incident Type Frequency

Incident type	**Frequency count**	**Percent of sample**	**Cumulative Percent**
Unilateral	26	72.0	72.0
Bilateral	10	28.0	100.0
TOTALS	36	100.0	

A description of all sample incidents follows. The analysis is divided into six broad categories that were coded and recorded in the content analysis forms: 1) Natural Language Semantic Unit, 2) Factual Assertion Statement (Relationship Instance), 3) Recording Unit Type, 4) Relationship Type, 5) Arguments, and 6) Relational Terms exhibited in the sample incidents.

4.3.2 Natural language semantic units

The nature of semantic units among different describing entities is considered here. The discourse surrounding reference incidents is mostly about things unrelated to the image content of photographs. The content analyst makes a determination regarding what portion of the text describes image content and it is that portion of text that is captured and recorded in the Relationship Corpus as a semantic unit. Semantic units, therefore, are verbatim accounts of image descriptions occurring in the sample incidents. Semantic units provide a transparent view of the raw data extracted from catalog records and correspondence.

A semantic unit is a non-unique expression, typically a sentence or clause relating to the image content of a photograph. A semantic unit may contain one or more instances of relationships and properties expressed in natural language. Typically, a period or semicolon marks the end of a semantic unit, but there are times when an image searcher simply lists in columns topics of interest with no punctuation. There is not always a clear boundary in the curators' correspondence separating what could and could not be considered relevant to image content. For instance, a photograph's relation to other entities such as collections or a researcher's intended use for photographs can provide important clues about the subject of desired photographs.

A unique system of numbering used in the Relationship Corpus, explained at the beginning of this chapter, enables matching semantic units to corresponding re-presentations as relationship instances, also called *factual assertions*. The structure of some sentences and sentence fragments occurring in descriptions is complex enough to warrant translating into two or more relationship instances.

The analysis shows that catalogers describe images differently from researchers and curators. Catalogers are charged with describing photographs according to local procedures and oftentimes national and international description standards. Catalogers usually rely on controlled vocabularies for subject classification and indexing purposes. Everything a cataloger states in an individual item's catalog record is descriptive matter dedicated to describing a single photograph. For this reason, everything a cataloger writes is semantically relevant to this study and it is treated as single semantic unit whether the description is composed of one word, one sentence, or an entire paragraph.

Catalog records occur in the samples as photocopies of the original *picture* catalog cards similar to the sample illustrated in Figure 23. The semantic content of these records is natural language, relatively straightforward, consisting of short, typewritten blocks of text accompanied by a small, gelatin-silver contact print. The cataloger's writing style can be cursory and fragmented as demonstrated in the following extract from incident number 37. Only two facts are included in the description: presumably an approved subject heading "River Traffic" and the accession code.

River Traffic. P-4785

The process of parsing relationships from semantic units like this is discussed in greater detail in the next section, but for illustrative purposes, the following examples of factual assertions are included here. They demonstrate how even a minimal subject-only description like this can be parsed into a set of relationships between the subject heading "River Traffic" and the photograph's accession code "P-4785."

accession_code_of(“P-4785”,photograph)

subject_of(‘river traffic’,”P-4785”)

In other instances, the cataloger provides a more detailed account of the depicted referents, photographer’s name, capture date, and the accession code as illustrated in this example taken also from incident number 37.

River Traffic on the Monongahela River. Tall building on the right is the LeGrande Apartments with Christopher's Restaurant on the top floors. Monongahela Incline. Photo taken by Lou Malkin. July 1974. P-4673

There are no instances in the data sample where a cataloger neglected to include at minimum an accession code along with a subject heading or picture caption.

Curators and image searchers generate image descriptions during an exchange of ideas about images during reference transactions. In the current study this conversation takes place in correspondence. In general, correspondence typically contains image descriptions, inconsequential data, and references to one or more images. For example, in Figure 26, the image searcher confirms with the curator a future meeting time and place and briefly describes the addition of a new Legislative Archives to the Capitol. There are instances in the sample of handwritten notes made by curators on image searchers’ correspondence. When the notes are relevant to an image description, they are treated as image description entities in need of capture and explanation.

Curators' and image searchers' correspondence is parsed to identify only those linguistic expressions that pertain to images descriptions. Table 8 indicates that the 226 semantic units resulted from unilateral incidents and 181 semantic units resulted from bilateral incidents. These totals, when compared to the percentages of unilateral and bilateral incidents shown in Table 7, demonstrate that on average, a unilateral incident produces 8.69 semantic unites per incident where a bilateral incident produces an average of 18.1 semantic unites per incident. There are significantly more unilateral incidents than bilateral incidents examined in this study, but bilateral incidents generate more semantic units because of the catalog records' highly descriptive nature.

Table 8: Semantic Units by Incident Type

Semantic units	**Frequency count**	**Percent of sample**	**Cum. percent**
Unilateral incidents	226	56.0	56.0
Bilateral incidents	181	44.0	100.0
TOTAL	407	100.0	

The first unilateral incident examined, number 010, consists of a total of 23 semantic units abstracted from one researcher's request for 21 images by subject. There is only one document in this incident, a two-page letter from the researcher sent to the curator of the Pittsburgh Photographic Library. It is shown in its entirety in Figures 26 and 27.

7 BEERS

HOUSE OF REPRESENTATIVES
COMMONWEALTH OF PENNSYLVANIA
HARRISBURG

February 18, 1986

[redacted]brary
Carnegie Library of Pittsburgh

Dear [redacted]

I shall be at your office Feb. 27, Thursday, at about 1 p.m. and hope to take an hour of your time to identify possible pictures the Pennsylvania Legislature might wish to purchase for reproduction.

The General Assembly is finishing a $200 million addition to the 1906 Capitol and is including a 4,000 square foot Legislative Archives and Visitors Center. I left the newspaper business after 30 years on Jan. 1 to plan and direct this new center. I am the author of six books, and for my 1980 "Pennsylvania Politics," published by Penn State, I use Carnegie Library pictures and was highly satisfied with the courtesy of your staff.

A feature of the center will be a graphic display of hundreds of pictures, cartoons and illustrations of the history of the Pennsylvania Legislature of 300 years. That amounts to a history of Pennsylvania. The picture requests I made would indicate that.

I only want to get the numbers on prints, no purchase now. I look forward to seeing you Feb. 27.

Cordially,

[redacted]
Bipartisan Management Committee
614-A Main Capitol
Harrisburg Pa. 17120

(717) 787-2738

Figure 26: Incident 010, page one (cover letter) of researcher's correspondence

Pittsburgh Pictures request

[redacted]
Pennsylvania Legislature

Whiskey Rebellion
i.e. storming of Gen. Neville estate ... drawing   LIB. OF CONGRESS ?

Canalboats and canal system

First Phila-Pitts. train 1852

Lafayette Hall and early Republican conventions of mid-1850s
anything on Republican gatherings of that 1850-1860 era

Civil War
military and manufacturing preparations
i.e. Ft. Pitt Foundry and biggest gun in the world

1877 Railroad Riot and burning of Union Depot

Homestead Steel Works and 1892 Riot

1889 Johnstown Flood and Pittsburgh Flood
also, 1936 false alarm of dam breaking on Johnstown main street

Pittsburgh millionaires dining in 1901 at Schenley Hotel
also pix of famous ones: Carnegie, Heinz, Frick, Charles Schwab,
Thomas Mellon (as well as Andrew, Richard Beatty Mellon, Wm. Larimer Mellon,
Richard King Mellon etc.), Henry Phipps, Capt. Bill Jones, Henry Oliver,
Charles Hall, Arthur Davis, Capt. Alfred Hunt, William Thaw, etc.

politicians Chris Magee and William Flinn

Black Pittsburghers of any era voting or in political situations

child and woman labor of 1800s and early 1900s

Suffragettes
marching, petitioning or early League of Women Voters etc.

Honus Wagner, the great shortstop

World War I and World War II
soldiers marching, Liberty parades, wartime manufacturing, newspaper headlines

Father James R. Cox, 1932, mass rally at Pitt Stadium
or any pix of Cox's Army to Harrisburg and Washington

Great Depression ... Pittsburgh joblessness

1936 Flood scenes

Dr. Jonus E. Salk and his anti-polio vaccine

Pittsburgh Renaissance ... before and after in Smokey City

Anything at all on Pittsburgh legislative delegation
(traditionally the most cohesive and best-organized in Pa., perhaps
late 1940s shot with Mayor Lawrence and Gov. Jim Duff)

Figure 27: Incident 010, page two (subject list) of researcher's correspondence

Sometimes researchers conflate the purpose for requesting images with the subjects of the images being sought. In the third paragraph of the cover letter in incident 010, the researcher writes:

A feature of the center will be a graphic display of hundreds of pictures, cartoons and illustrations of the history of the Pennsylvania Legislature of 300 years.

The researcher is describing a feature of the Legislative Archives and Visitors Center, but pragmatics suggests that this is part of the subject request. Pragmatics is discussed in more detail below (see §4.3.2.2), but a brief explanation is useful here for clarifying the current example. Pragmatics is concerned with aspects of information conveyed through conventional linguistic forms of language in conjunction the context in which they are used. This enables the analyst, and it is reasonable to expect the curator as well, to identity an image subject request marked by the word "feature" in the sentence "A feature of the center will be…the history of the Pennsylvania Legislature…" While the image searcher does not explicitly ask the curator for images that fall under this broad subject heading, in the context of photograph archives it is reasonable to infer that this sentence expresses a broad subject request for images of the Pennsylvania Legislature.

The second semantic unit 010SEMNL02 recorded in this incident immediately follows the first in the original correspondence. It reinforces the notion that *subject* and its relation to *picture* can occur between sentences. The image searcher goes on to say

That amounts to a history of Pennsylvania. The picture requests I made would indicate that.

The inter-propositional relation of subject and picture arises from the phrase “history of Pennsylvania” in one sentence and “picture requests” in another sentence. As problematic as the relationship between the two forgoing sentences appears to be for this study, so, too, is the next set of lexical hurdles taken from the same request: ‘canalboats’ written as one word, an ambiguous abbreviation ‘Phila-Pitts,’ which the analyst takes to mean ‘from Philadelphia to Pittsburgh,’ and disambiguation problems such as ‘(military) and (manufacturing preparations)’ versus ‘(military preparations) and (manufacturing preparations).’

Canalboats [sic] and canal system

First Phila –Pitts [sic] train 1852

Lafayette Hall and early Republican conventions of mid-1850s
anything on Republican gatherings of that 1850-1860 era

Civil War
military and manufacturing preparations,
i.e., Ft. Pitt Foundry and biggest gun in the world

The first semantic unit 010SEMNL01 places in close semantic proximity the narrower subject of “the history of the Pennsylvania Legislature” with the broader topic “the history of Pennsylvania” in 010SEMNL02. This same juxtaposing of narrower and broader subjects occurs in semantic units extracted from cataloger’s descriptions. For example, in the following bilateral incident 045SEMNL83 a cataloger describes a photograph first with a general heading "industries" followed by the narrow subject "iron and steel" and then finishes with an individual instance "Isabella Furnace of the U.S. Steel Corporation."

Industries. Iron & steel. Isabella Furnace of the U.S. Steel Corporation at Etna, PA, April 1948. P-7080

This raises a concern for consistency in interpretation and how semantic units are parsed into factual assertions. Narrower and broader topics and their relations to each other imply hierarchical schemata and this can be inferred at the level of semantic unit. That is, the conversation carried on between image searcher, cataloger, and curator. After a given semantic unit is parsed into one or more short, precise assertions under question 6 in the content analysis form, implicit hierarchical relationships that might have been evident within a single semantic unit are not recognizable in multiple, semi-formal assertions.

The relationship at the instance level is generally not included in an ontology, so these interesting inter- and intra-propositional relations that exist in natural language may be lost in an ontology-based representation. It is a topic for future study.

On the subject of hierarchical relationships, the general rule applied in this study has been the following: Unless entities are explicitly associated through class inclusion relations by the image describer (that is, the describer states explicitly in a sentence that *X is_a Y*, or *X is a kind_of Y*, or *X is a type_of Y* as in "a hammer is a tool," factual assertions are not used in the relationship corpus to build extensive hierarchical schema. For example, if the term "illustration" occurs in an image searcher's query, the analyst does not then translate this into a set of inclusion relationships such as:

kind_of(illustration,'visual image')
kind_of('visual image',image)
kind_of(image,picture)
kind_of(picture,'mental representation')

Turning now to distinctions between catalogers and image searchers descriptions, what is most apparent is the precision and economy of language with which catalogers express details, the kinds of details included, and the amount of background information that the reader must posses to fully understand the meaning of, for instance, "April 1948" or the alphanumeric code "P-7080." While image searchers are describing both known and unknown images and curators are responding to researchers and mediating between catalogers and researchers, a cataloger's task is to describe known images, but they often include information that can only be understood within the context of cataloging rules and practices. The most important distinction between the two, however, is that catalogers are describing known entities while the researcher is often describing unknown, possibly non-existent entities. The importance of this distinction lies in the legitimacy of building factual assertions from searchers' descriptions and including them in an information system, or a Corpus of Relationships for that matter. Moving from a database and word-indexing model to a knowledge base and relationship model makes searchers' descriptions highly problematic statements for two reasons: 1) They are propositions usually describing *wished for* photographs, and 2) The expressions searchers use are composed of common nouns such as "photograph" and "pools"—terms that could apply to an infinite number of entities.

4.3.3 Factual assertions

Factual assertions are the most important lexical units examined in this study because they make explicit the relationship instance. Table 10 indicates that 407 semantic units resulted in 1,655 individual relationship instances or factual assertions. *Factual assertions* are rendered as precise statements extracted from the semi-formal expressions of relationships indicated in the original natural language statements. Here the discussion turns to the nature and characteristics of the relationship instances that emerge from the semantic units. The first task is to establish how it can be determined that a factual assertion is what it claims to be. That is, a relationship instance.

Table 9: Frequency of factual assertions based on type of describer

Recording unit type	**Frequency count**	**Percent of sample**	**Cum. percent**
Image searcher's query	897	54.0	54.0
Curator's correspondence	205	12.0	66.0
Cataloger's description	553	34.0	100.0
TOTAL	1,655	100.0	100.0

4.3.3.1 Hood's criteria of relation

This study adopts Hood's (2004) criteria of relation theory and her approach to understanding and specifying relation in the logical syntax of first-order logic. According to Hood's criteria of relation theory, "First it must account for the minimally dyadic, ordered character of relation. Second, it must account for some relational predicate apart from the two items that stand in relation to each other" (p.2). The content analysis form provides a way of specifying a dyadic predicate *R(a,b)* and thus the ability to determine whether the relation meets Hood's criteria. In the relational proposition *R(a,b), R* is the relation and *a* and *b* are the terms or arguments. Following Peirce's categories of predicates, a proposition with two arguments is called a *dyadic predicate* (Michael, 1974). Hood's criteria of relation, when applied to original natural language expressions are a necessary, although not a sufficient, condition for arriving at dyadic relations. Applying Hood's criteria to the reconstruction of semantic units into factual assertions, however, is a sufficient condition for formulating relations.

To recap the coding instructions, the syntax of relational propositions specify a relational predicate, two or more items involved in the relation, and their order. In the factual assertions, the relational items—called arg1, arg2, arg3—are separated by commas and enclosed in parentheses. The relationship term is furthest to the left before the first parenthesis. Evidence that the relationship criteria have been met comes in specifying the minimally dyadic, ordered pair of terms (Columns 11 and 12 in the content analysis form) and a term representing the relational predicate (Column 14).

The topics of pragmatic inference and how relationships are identified and parsed from semantic units are discussed next before proceeding with an analysis of factual assertions,.

4.3.3.2 Role of pragmatics

To effectively translate natural language reference incidents into more precise factual assertions, the content analyst must not only look for meaning in the literal content of what is uttered by catalogers, curators, and image researchers, but also make appropriate inferences that capture their intended meaning. *Inference* is a type of reasoning that manipulates known propositions to produce new ones (Levesque & Lakemeyer, 2000). The term *pragmatic inference* is used in this study to expand on this meaning, taking into consideration the context in which propositions are stated. In the domain of image descriptions, pragmatic inference enables not only the content analyst to infer meaning from background, contextual information, it allows curators and image researchers to identify aspects of information implicit in the context of the reference incident. For example, in the dialogue in (1) taken from incident number 15, the curator infers from the image searcher's request that that the subject of interest is "snowstorm of 1950."

> (1) **Image Searcher**: "I noticed that your organization sells a photograph by Harold Corsini (ID# 3087) of a tank in the snowstorm of 1950. I am wondering if there are other similar pictures taken at the time."
>
> **Curator**: "There are several other photos of the snow of 1950 in the Pittsburgh Photographic Library collection. I can send you photocopies of the other 'snow' photos if you provide me with a mailing address."

Nevertheless, asking for pictures similar to the "tank in the snowstorm of 1950" does not rule out the possibility of the curator searching for more photographs of tanks. In this particular case, a follow-up letter from the researcher cancels the curator's initial inference in (2) and the curator responds that he has gathered together more photographs of tanks and military equipment.

> (2) **Image Researcher**: "I would be interested in other photos taken of tanks and military equipment at other times, but I imagine that might be a pretty involved task."

> **Curator**: "Enclosed please find photocopies of photographs in the Carnegie Library's photo collection relating to tanks, military equipment."

This division between linguistically encoded meaning and inferences was made prominent by Grice (1957/1975). *Semantics* refers to the aspects of the interpretation that can be derived from the meanings of words and the structural relationships between them. In contrast, pragmatic inferences are meanings "based squarely on certain contextual assumptions" rather than on the meanings of words and sentences (Levinson, 1983, p.167).

Pragmatic inference also plays an important role for the content analyst during coding. For example, if in a given context the photograph is known, but not stated explicitly in the image description, the coder may code an attribute as a binary predicate. For example, if the photograph's name "Migrant Mother" is present in a cataloger's record, but only the capture date of "1936" and the photographer's name "Dorothea Lange" are mentioned in the image searcher's request, the coder can infer in the factual assertion that the date and photographer's name are associated with "Migrant Mother." The factual assertions could be written thusly:

- *capture_date_of("Migrant Mother","1936")*
- *photographer_of("Dorothea Lange","Migrant Mother")*

4.3.3.3 Identifying and parsing relationships

One of the unique features of this study is its focus on relationships that occur in real-world discourse. In the current context of archival discourse there are various levels of granularity on which relationships among entities can be examined. This study focuses on finer-grained relationships expressed within sentences and sentence fragments: predicates that are class terms—i.e., nouns, adjectives, or verbs designating classes of objects—and proper nouns that name individual objects. There are other relationships that can be envisioned existing between larger semantic units. For instance, relationships between propositions within a single piece of correspondence and between questions researchers ask and curators' responses; sets of complex relationships between photographs within named collection and between collections within archives. The possibilities appear limitless or nearly limitless. While the goal of this investigation is to focus on sentences and sentence fragments, inter-propositional relations and relational inference do influence the manner in which the relational predicates in discourse are analyzed. It seems there are good reasons not to ignore these influences along with what is implicit in background knowledge.

The solution offered here is to expand the scope of analysis to at least examine relationships that exist between propositions within the context of a single incident and how this influences the interpretation and re-presentation of propositions as factual assertions. A detailed study of the interpretive process within the broader context of the incident provides the necessary framework for explaining a pragmatic approach to the problem of the relationship in instances where, for example, only topics and proper names are stated explicitly and their relationship to images or to objects depicting images is implicit in background knowledge. The following example attempts to illustrate this analytic approach. Here is a typical existential clause found throughout the corpus. It is a sentence fragment that simply states that something exists—a subject without a predicate. The example comes from a letter sent by an image searcher to a curator.

[011SEMNL40] *Armor plate manufacturing*

In this example, two relationships are formed by building associations with entities in statements made earlier in the incident event. The semantic unit ‘armor plate manufacturing’ is part of a list of subjects organized by the image researcher under the broader heading “Industry near Pittsburgh, generic.” This is a rare case in the sample where a researcher uses headings to categorize topics being requested as image subjects. It is useful not to separate this knowledge—knowledge that is part of the broader context of the incident—from the individual topic ‘armor plate manufacturing’. Going beyond this, pragmatics enables inferences to be drawn between different documents within a reference incident and among different individuals participating in the incident.

In the case of this semantic unit, the topic “armor plate manufacturing” carries with it even more background knowledge than that associated with the heading “Industry near Pittsburgh, generic.” In the cover letter accompanying the list of requested topics, the researcher in this incident states, “I am working on picture research for a History Channel documentary about the Homestead Strike of 1892 in Homestead, Pennsylvania,” and later adds, “My producer and I have attached a list of subjects/images that we are interesting [sic] in pursuing for use in the documentary.” From these remarks a relational predicate can be inferred for the re-presentation of the original topics as a more refined relationship instance and set of factual assertions in the relationship corpus. First, the sentence fragments ‘armor plate manufacturing’ and ‘industry near Pittsburgh, generic’ are translated into a sentence with a subject-predicate form as shown here:

The researcher is interested in photographs documenting armor plate manufacturing near Pittsburgh.

The predicated item ‘photograph’ is said of its subject ‘armor plate manufacturing.’ The item that is predicated has to be some entity that many things can be said to be because the related entities are classes of objects. Thus, this content analysis adopts the term ‘photograph’ as opposed to a named individual manufacturer, which is an unknown entity at this point in time.

There are a variety of terms in the sample used as a predicated item representing the photograph. In the discourse surrounding this one incident alone the image researcher uses the terms ‘picture,’ ‘image,’ ‘still image,’ ‘visual record,’ ‘photos,’ and ‘prints’ to infer photographs. When no predicate is stated explicitly, as in the case of ‘armor plate manufacturing,’ the term ‘photograph’ is applied in the relationship corpus. Here the semantic sense of the term *predicate* is used, not its grammatical sense. A grammarian refers to anything that is not the subject as being the predicate. In a grammatical sense, the predicate in this example is “is interested in photographs documenting armor plate manufacturing near Pittsburgh.” From a semantic standpoint, however, the predicate is the verb ‘document.’ The factual assertion parsed from the above semantic unit is represented in prefix notation here:

document(photograph,’armor plate manufacturing’)

This formulation can be read as ‘there is a photograph documenting armor plate manufacturing.’ An abstract semantic construction begins to emerge when the nature of the arguments are examined more closely. The term located to the left of the comma—named argument one or arg1 for short—is a term representing an object that depicts a pre-photographic referent. (See §1.4 for a definition of ‘object depicting referent’ and other levels of meaning considered critical to this study.) The object depicting the referent is a ‘photograph’, which is understood to be a physical manifestation or photographic print in the context of the sample incidents. The term in the position of arg2, ‘armor plate manufacturing,’ represents the subject of documentation. In this reading of the existential clause, the photograph’s role as a *documenting entity* is stressed. The photograph is associated with the topic ‘armor plate manufacturing’ through a documenting relation.

Encoding this proposition in English in other ways can maintain the same core meaning, but reflect different informational emphasis. For example, ‘the researcher’ is the grammatical subject in an active construction, but ‘photographs’ could be made the subject in a passive construction. Here it is presented in natural language and then in prefix notation as a factual assertion.

Photographs documenting armor plate manufacturing near Pittsburgh are of interest to the researcher.

interesting_to(‘photograph documenting armor plate manufacturing near Pittsburgh’,researcher)

The *interesting_to* relationship expresses an association between the researcher and the photograph. The several terms included within arg1 itself may also be reformulated to shift informational emphasis as shown here:

interesting_to((document(photograph,'armor plate manufacturing near Pittsburgh)), researcher)

The second proposition is more complex, expressing a factual assertion within a factual assertion. The value of this formulation is in the association it establishes between the researcher and the relationship that exists between the photograph and subject. In the relationship Corpus both simple and compound assertions are made providing a rich set of semantic relations describing each incident. This particular proposition is given additional analysis in the Corpus by recognizing, for example, the *near* locative expression in the relationship instance *near(manufacturing,"Pittsburgh").*

When the analysis begins with only a basic existential clause such as 'armor plate manufacturing' and no predicate stated explicitly, there are many possible ways to translate the clause into a relationship instance. In the assertion *document(photograph,'armor plate manufacturing'),* the term "photograph' plays the role of a non-sentient causative that precludes the explicit or implicit involvement of an agent, in this case the photographer taking the photograph. In the proposition 'photographs document manufacturing,' the documenting process involves no conscious activity. However, to state 'a photographer documents armor plate manufacturing in pictures' expresses an active, conscious, human activity taking place. The interpretation in the next factual assertion represents this sentient entity—a photographer—carrying out the action of photographing.

document(photographer,'armor plate manufacturing')

In this example the *document* relation is an instance of *agent-action case relation*. In the ontological analysis developed in chapter five, when a topic is listed by itself absent of any predicate—what has been called an *existential clause* in current analysis--the fragment is restated as an *attribution subject_of* relation based on definitions of attribute by Guarino (1992), Sowa (2002), and Woods' linguistic test (1975). Later in chapter five a detailed argument is presented that attempts to explain how an attribute can function within a dyadic predicate.

In the meantime, to get through the analysis of attributes being used in relational propositions, it is useful to consider Guarino's definition of attribute as, "a concept which, in the domain of interest, has a unique relational interpretation, satisfying the Attribute Consistency Postulate." The Attribute Consistency Postulate states, "Any value of an attribute is also an instance of the concept corresponding to that attribute" (Guarino, 1992). The following linguistic interpretation offered by Woods states that an attribute name has to denote a concept. If *Y* fitting this expression cannot be found, *A* cannot be an attribute.

Y is a value of the attribute *A* of *X* if we can say that *Y* is a *A* of *X* (or the *A* of *X*).

Applying this to the earlier proposition results in the following linguistic interpretation:

'Armor plate manufacturing' *Y* is the value of the attribute *subject_of* photograph if we can say that 'armor plate manufacturing' is a *subject_of* photograph or *the subject_of* a photograph. Since the value 'armor plate manufacturing' does fit this expression as being a subject of a photograph, *subject_of* can be an attribute and represented as a factual assertion thusly:

subject_of('armor plate manufacturing',photograph)

Given that attributes are monadic predicates and relations are polyadic, a question emerges how to treat the relational interpretation of attributes with values? In other words, it is common for humans to say a photograph has a subject. This expresses a monadic predicate. To add to this proposition a named subject creates a relationship between the subject and the object in which it inheres. Guarino's *attribute consistency postulate* addresses this question by presenting attributes as concepts—monadic predicates—which are associated with binary relations. The postulate states, "Any value of an attribute is also an instance of the concept corresponding to that attribute" (Guarino, 1992, p.). Consider the attribute *subject_of*. The concept corresponding to the attribute is the concept of "subject." In the relational proposition *subject_of(industries,photograph),* according to the attribute consistency postulate, "industries" is the value of the attribute "subject" and an instance of the concept "subject."

To summarize, the *factual assertion* is an instance of relationship type and the main unit of analysis in this study. Relationships evolve from the natural language image descriptions that make up the sample for this study. The analysis shows there are 1,265 instantiations of relationships in the data set. There are cases where relationship instances cannot be modeled relying on only the explicit statements in the original sentences or sentence fragments. The analyst draws supplemental knowledge from other propositions within the same incident, through pragmatic inference and from general, implicit background knowledge. The same proposition may be instantiated different ways to shift informational emphasis of relationships nested within an overarching relationship.

This analysis now turns to a more thorough examination of the nature and semantic role of items making up relationships in the factual assertions. A distinction is drawn between classes of things and individual things that are members of classes. Predicates and the properties of things are discussed in preparation for the ontological analysis that is required in chapter 5.

4.3.4 Record unit type

This study adopts the working assumption that the production of image descriptions is a social practice, as opposed to considering memory institutions' catalogs as being contexts in which image searchers come to view descriptions as if they are predetermined reflections of a prior set of institutional determinations made by catalogers. Thus, when examining how humans express relationships describing images, this study considers the record unit type, which refers to three categories of describers: catalogers, curators, and image searchers. As indicated in Table 10, it is not surprising to see catalogers' descriptions resulting in a fewer number of semantic units (92) when compared to image searchers' (266). Catalogers' descriptions were not broken down into smaller fragments, where references to image descriptions made in image searchers' correspondence was typically made up of individual sentences spread throughout the correspondence. Because of their rich descriptive nature, catalogers' descriptions were predicted to result in the highest number of relationship instances. The analysis in Table 11 shows, however, that image researchers' correspondence resulted in 898 relationship instances compared to 553 resulting from catalogers' descriptions. This noteworthy finding demonstrates the value of considering image description as a social practice that engages more than just catalogers working in institutions and organizations with an archival function. Image searchers' descriptive propositions accounted for 21% more of the total number of relationship instances than did statements made by catalogers. The curators' correspondence in this sample resulted in only 204 relationship instances, or 12% of the total, which supports the notion that their role as mediators in the discursive space of archives consists more of administrative tasks (e.g., explaining reproduction rights, pricing, availability, and so on) than image description tasks. For example, as demonstrated earlier in Figure 25, curators are concerned with customer service issues as expressed in the statements, “I’m sure we can find something suitable for you,” and “Let me have your fax number, when you get a chance. [We] will probably want to fax you images of what we have.”

Table 10: Frequency of semantic units based on recording unit type.

Recording unit type	Frequency count	Percent of sample	Cum. percent
Image searcher’s query	266	65.0	65.0
Curator’s correspondence	49	12.0	77.0
Cataloger’s description	92	23.0	100.0
Could not determine	0	0.0	0.0
TOTAL	407	100.0	100.0

4.3.5 Relationship type

In Question 8, the coder determines the relationship type expressed in the factual assertion. There are seven broad categories from which to choose: 1) Attribution, 2) Case, 3) Inclusion, 4) Meronymy, 5) Spatial, 6) Synonymy, and 7) Temporal. These seven relations and their occurrences in the data sample are examined in the sections that follow.

Table 11:Frequency of relationship instances by type.

Relationship type	Frequency count				Percent count	Cum. Percent
	Cataloger	Curator	Image Searcher	Total		
Attribution	264	48	331	643	39.0	39.0
Case	129	63	210	402	24.4	63.4
Inclusion	2	9	22	33	1.9	65.3
Meronymic	16	22	40	78	4.7	70.0
Spatial	101	41	161	303	18.3	88.3
Synonyms	5	4	13	22	1.3	89.6
Temporal	36	17	112	165	9.9	99.5
Can't determine	0	0	9	9	.5	100.0
TOTAL	553	204	898	1655	100.0	

4.3.5.1 Attribution

As indicated in Table 11, attribution, sometimes called the 'have/has relationship', represents the single largest class of relationship types to emerge in this study. A clarification of the term *attribution relationship* is in order for it appears that attributes, which are unquestionably monadic predicates, are once again being treated as if they are relationships. The apparent inconsistency was addressed earlier at the end of § 4.3.3.3. The short answer is that attribution is not a relationship, but in this study it is common practice to combine attributes with other propositions to create what this researcher calls a *relational proposition of attribution* or *attribution relationship* for short.[10]

A total of 643, or 39% of all instances, were identified as attribution relationships. Image searchers expressed attribution relations 331 times in the sample. This is interesting considering that image searchers in this study, for the most part, describe non-existent photographs. The next largest group was catalogers who expressed attribution relations 264 times in the sample. Curators expressed attribution relations in only 48 instances. Table 12 lists the nine attribute relationships that when considered individually represent more than 1% of the total number occurrences in the sample. There are 67 other attribute relationships that occur less than 1% of the time.

The attribute relationship represented most often is the *subject_of* relation as in, "'Industrial factories is the *subject_of* photograph P-8467." In instances where only a subject term exists in the semantic unit, for example, "Industries," and there is no named photograph, the semantic unit could be interpreted or parsed as a monadic predicate *R(a)* or *subject_of(industries)*. But in so doing, Hood's criteria of relation would not be met. Therefore,

---

[10] Formulating propositions where attributes are applied in dyadic predicates rather than used as *monadic* or *one-termed* predicates has important implications for how information professionals model information about image descriptions in knowledge-based systems. There are those who would argue that an attribute is a monadic relation, a relation that holds between a subject and a predicate. The argument goes something like this: If a photographic image has the subject "industries," then a relationship exists between subject and industries. That is, subjectness inheres in industries. The argument put forth in this dissertation is that humans talk about subjects in relation to images, so it is useful to formulate propositions that represent this relationship. The terminological solution offered here is to call this new entity a *relational proposition of attribution*—that is, a relational proposition that arises as an implicature between two propositions: one that names an image attribute and one that states objects depict image attributes. The object in question may be a known photograph, in which case it is named, or it may be an unknown photograph and designated by the term 'photograph.'

in cases where a human expresses a subject associated with an unnamed photograph, which is common in image researchers' requests, the content analyst applies the *subject_of(x,y)* construction thus making it a dyadic relation. Given the universe of discourse—a photograph archives and image descriptions—the subject is presumably predicated of a photograph, so arg1 is the subject and arg2 is the term 'photograph' as in *subject_of(industries,photograph)*. This interpretive process does not sacrifice the integrity of the sample data, which remains in tact in the semantic unit and the analytic process is transparent in the Corpus of Relationships. The question, then, arises as to how this formulation might behave in a knowledge-base system—a question that is outside the scope of this dissertation.

Table 12: Most common attribution relationship types exhibited in sample

	Frequency count				
Property-types	Cataloger	Curator	Image searcher	Total	Percent of Sample
accession_code_of	92	2	4	98	15.24%
date_of	23	0	5	28	4.35%
Format_of	2	0	11	13	2.02%
occupation_of	5	0	9	14	2.18%
photographer_of	8	0	1	9	1.40%
pix_of	0	0	17	17	2.64%
relative_of	10	0	0	10	1.56%
role_of	4	2	14	20	3.11%
subject_of	69	28	172	269	41.84%

In the current investigation, the nature of the subject_of relationship depends on how it is created during the discourse among catalogers, image searchers, and curators. How subjects are marked in descriptions and generalizing their nature is a difficult task. For example, the image researcher asking for a portrait of Andrew Carnegie conceives of the subject of a photograph as a conceptual integration of two different objects: the depicted referent Andrew Carnegie and the object showing depiction, in this case a photographic print, which is implicit in the request shown here:

[045SEMNL01] *"Do you hold the portrait of Carnegie at the top of the site?"*

There is evidence of this in the factual assertion *portrait_of(photograph,"Carnegie")* accompanying the above semantic unit. The searcher does not ask for Andrew Carnegie the pre-photographic referent. The request is for a portrait of him and as stated earlier, implicit in this request is that the object showing the depicted referent is a photographic print. This example also demonstrates integrating the abstract entity {portrait} with the subject Carnegie and the physical carrier of the image, a photographic print.

The preposition 'of' in the phrase "portrait of" can be interpreted as expressing composition. That is to say, the picture—a portrait—is made of this entity's visual information. The entity in this instance is the depicted referent Carnegie. The preposition "of" directly follows a noun 'portrait', indicating the material of which it is made. In this instance the material is a mental construct "portrait." Another interpretation could explain "of" as introducing the subject matter following the noun "portrait" to introduce the subject matter of the portrait.

One thing that this discussion highlights is that the process of explaining subject is an active social process that is the result of human beings describing image content—known images and unknown, but desired images. It is part of the description process that is carried out by agents on an image object, imagined or real, physical or mental, known or unknown. The result or outcome of the process is the naming of the attribute that the photograph is an image of.

It is a person's knowledge of the world that leads her/him to interpret some prehending entities (arg1 in subject_of constructions) as being conceptually integrated with the photographic print (e.g. a picture of Andrew Carnegie) and not others (e.g. emigration). The subject_of constructions can be further categorized by a narrower interpretation where arg1 is a spatial expression, for instance

[015FACTS0103] *subject_of('tank in snowstorm',"3087")*

This can be read as "The subject of the photograph with accession code 3087 is 'tank in the snowstorm'." The *photographer_of(x,y)* construction, on the other hand, strongly enforces the narrower interpretation that arg1 and arg2 are in an agent-object case relation. That is, a relationship exists between the photographer (agent) and the photograph (object) that takes part in the event.

The following factual assertion was parsed from the original semantic unit written by a cataloger, "Andrew Carnegie. Bedroom. Scotland. C1073."

[045FACTS90281] *location_of("Scotland","C1073")*

This can be read as "The photograph with accession code C1073 was taken in Scotland." The reader of the original correspondence brings a certain amount of background knowledge to this reading. The subject of the picture is Andrew Carnegie's bedroom in Scotland. In this example location_of tells us that Scotland is integrated into the photograph. The integration of arg1 Scotland and arg2 C1073 amounts to near identity. A photograph with location_of Scotland shows us a scene in Scotland. This is in contrast to ‘from’ where the description of a photograph from Scotland implies the photograph followed a path, beginning its journey in Scotland but may not contain subject content pertaining to Scotland.

How or if a listener integrates the subject and photograph depends on her knowledge of the world and the context or domain of interest in which the subject is stated. If a taxi driver utters “Woodrum sleeping car” it sounds incomplete and completely unnatural. If the same sentence fragment is included in a letter sent by an image researcher to the curator of a photograph archive it is understood to be a request for photographs expressing or showing this particular subject.

Photographs are supposed to have subjects, so it is natural for researchers to list subjects where ‘picture of’ or ‘photo of’ is implicit or stated only once and then it becomes background knowledge for the remainder of the reference event. Because of what humans know about subjects, the phrase *x is a picture of y* means that in the listener’s mental integration, the subject of the photograph and what the term ‘picture’ represents are virtually the same in scope. When curator’s knowledge of the world informs her that the subject is intrinsic to the depicted referent, conceptual integration is particularly thorough. For example,

[045FACTS3441] The subject of the photograph is Col. Anderson Library.

[045FACTS3445] The subject of the photograph is Homewood House, Pittsburgh.

[045FACTS4961] The subject of the photograph is Battle of Homestead.

Because of what the reader or listener knows about photographs and depictions of events and states of affairs, any word or phrase expressed as a subject reflects a natural integration of subject and the object depicting the image. Even when extrinsic relations are expressed, it is understood in the discursive space of reference incidents that the meaning is brought to the image by the individual who is narrating.

Before leaving this analysis of the family of attribution relations there is another problem to considered: The extent to which the functional predicate $f(x)$ is consistent with the usage of the image descriptions of catalogers, image seekers, and curators of the Pittsburgh Photographic Library. Perhaps a stronger case for attribution relationships is needed.[11]

The analysis in this section follows Aristotle's theory of predication as interpreted by Allan Back (2000) who describes the standard Aristotelian subject-predicate sentence of the form 'S is P', which is to be read as 'S is existent as a P'. So, for instance, the proposition 'Harold Corsisni is a photographer' [015SEMNL01] is to be read as 'Harold Corsini is existent as a photographer'. This reading, based on Back's interpretation, has compound truth conditions. The proposition 'Harold Corsini is a photographer' requires that both Harold Corsini be existent and that Harold Corsini be a photographer. This is to say 'photographer' signifies one of the attributes of Harold Corsini. Read more strictly, 'Photographer is predicated of Harold Corsini.' Therefore, 'S is P' is a conjunction of 'S exists' and 'P is predicated of S'. Back refers to his explanation of this theory as the *aspect theory* of predication, "as the predicate is supposed to stipulate a certain aspect of existence of the subject" (Back, 2000, p.3).

The argument is made here that when expressing attributes of images, S can be extended with f(S) to be read 'S is the function of' and when added to the original conjunction 'S is P' it

[11] The researcher views attribution relationships in two different senses: structural and functional. Attributes hold complex relations with the physical manifestations of photographs and the concept {photograph}. These relational structures are the subject of this dissertation. Attributes can be also functionally complex in information systems mapping values of attributes to instances of images—a viewpoint that is briefly introduced in this section.

becomes f(S,P) to be read 'S is the function of P'. Considering this example of a functional predicate as a "photographer formula,' a categorical sentence evolves of the form *photographer_of {photograph} = {person}* which can be read 'some person is the photographer_of some photograph.'

Another way to explain this is to describe a function called *photographer*. If one were to input the photograph named "Tank in 1950 Storm," what does the function photographer produce? In other words, what is "photographer of Tank in 1950 Storm"? Photographer_of "Tank in 1950 Storm" = "Harold Corsini". Viewing this proposition in prefix notation it would be written as

photographer_of("Harold Corsini","Tank in 1950 Storm")

Taking one more example that views attributes as functions consider the most common attribute *subject*, or what might be named here the "subject formula." If one were to make *subject* a function and input "Harold Corsini" into *subject*, then *subject* produces "Corsini at the Taj Mahal" (a 35mm slide of Harold Corsini taken circa 1941-1949). This can then be extended to subject_of(photographer_of("Tank in 1950 Storm")). First, photographer_of("Tank in 1950 Storm") was already determined to equal "Harold Corsini". Therefore, this new formula is the same as subject_of ("Harold Corsini"), which equals "Corsini at the Taj Mahal".

To summarize, the most the commonly occurring attribution relations outlined in Table 23 can be generalized to a **(?X{thing} *attribute_of* ?Y{photograph})** construction.[12] This type

---

[12] The term *construction* (borrowed from Construction Grammar), is used throughout the remaining chapters of this dissertation to describe structural patterns that have meaning much

of relationship is broken down further into different subtypes of constructions or *formulas*, each one linking a particular attribute with a photograph. The formulas are named after the attribute types and subtypes. For example, in Table 12 the *accession_code_of* formula links accession codes with photographs; the *subject_of* formula links subjects with photographs.

Following this explanation of applying functions to predicates, the idea is tested and applied to accession codes. A function *accession_code_of* produces an expression that denotes an object that functions grammatically as a new individual name. When predicates are applied to one or more names they produce new sentences. Functions, on the other hand, denote single new individuals.[13] For example, applying the phrase 'accession code of' to the term 'photograph' results in the phrase 'accession code of photograph', which designates the individual L1272 (assuming that is the accession code that the cataloger intended). The new phrase 'accession code of photograph' is called a *functional expression* and acts grammatically as a name (Nolt, Rohatyn, & Varzi, 1998). In this study the phrase is used in grammatical contexts in which a photograph's accession code is appropriate. In incident number 013, the image searcher makes it explicit that the functional expression 'alphanumeric code of' applies to an individual photograph, and for the purposes of this study, alphanumeric code is considered synonymous with accession code. This incident is used to demonstrate how the functional expression

---

like individual words have meaning. Rather than piecing together meaning from individual words, *Construction Grammar* proposes that the meaning of a sentence arises from both lexical units and from patterned *constructions* in which the lexical units are situated. The research constructions applied in this study focus less on grammatical properties and more on syntactic patterns and semantic categories. The notation used for representing constructions is borrowed from Fahlman's (2008) blog *Knowledge Nuggets*. A construction representation includes variables (e.g., ?X, ?Y), semantic categories (e.g., {person}, {object}), and the relation term, which can be specific or general (e.g., *subject_of* or attribution).

[13] To distinguish one accession code from another in the relationship corpus, each factual assertion has a corresponding factual assertion statement number.

‘accession code of’ can be applied in any context where an accession code is appropriate. The image searcher writes

The image in question has the alphanumeric code of "L1272," and is accompanied by the text "Pittsburgh in the Evenings, c 1940s, Lorant Collection".

In this incident the searcher names a photograph by listing a single accession code: L1272. The functional expression ‘accession code of’ can be attached to the name “L1272” to create a predicate resulting in the sentence, “The accession code of the photograph is L1272.” This sentence could be represented in the language of predicate logic as either a one-place or two-place predicate. This study adopts the view that a dyadic predicate produces a more desirable, finer-grained representation, which can accommodate either a named photograph (instantiation) or the type ‘photograph.’ Therefore, the accession_code_of function is expressed as an attribution relation in a two-place (dyadic) predicate where the function accession_code_of assigns to the class ‘photograph’ or instance of the class photograph a second term which is the accession code. In this sense a functional expression is used as a predicate to create new sentences. It could be stated in natural language a number of different ways:

1. Accession code of photograph is X.
2. X is accession code of photograph.
3. Photograph has accession code of X.

This study applies example two, ‘X is accession code of photograph.’ The object assigned by an accession_code_of function is called the *value* of the function for the ordered pair. In the current example under examination, the value of the accession_code_of function for the argument ‘photograph’ is “L1272.” In the relation Corpus this is expressed as:

accession_code_of(”B1272”,photograph)

This factual assertion can be read as “B1272 is the accession code of the photograph.” The two terms inside the brackets are ordered, meaning that they must occur in a particular order. In other words, it would not make sense to say, “Photograph is the accession code B1925.”

Once this proposition is made explicit in a knowledge base system, then the information can be associated with that named individual. In a subject_of(x,y) construction, x represents the subject and y, the second term in the ordered pair, represents the named individual designated by an accession code. Image searchers’ queries sometimes consist of lists of common or proper nouns. For instance, in incident 011 the searcher lists: “Duquesne Works,” “Lucy Furnaces, “Carrie Furnaces,” “open hearth process,” and “Beaver Falls Pennsylvania Mill.” These are re-presented in the relation Corpus as two-place subject_of or image_of attribution relations, not functional expressions. Here the subject_of relation is not treated as a functional expression because it does not denote unique individuals because the researcher does not ask for specific photographs in the collection. When subject_of is applied to any photograph that has more than one subject, it denotes not one subject term, but several.

There are no further refinements of the attribution relation formally examined here. There is nonetheless an interesting aspect of attribution that emerges from this study. It is in the variety of attributive relations humans express when describing visual information. The question examined in Chapter five is whether these different instances can be grouped into a broader class of relationship types.

To summarize, the analysis shows that humans most often describe photographic images in terms of subject, that subjects occur in ordinary language use during reference incidents, and that people commonly write image descriptions with subject predications of photographs. The phrase *subject predications of photographs* is used here to mean subjects are assigned to images using semantic constructions that follow the general pattern **(?X{thing} *is_predicated_of* ?Y{photograph})** or **(?X{thing} *is_said_of* ?Y{photograph}),** where *X* is a subject term and *Y* is a known or unknown photographic image. Furthermore, the expressions of subject terms such as “industries” are fragmentary and vague, but intelligible given the context. Contextual factors bring meaning to what is being expressed. In this particular case, the curator reading the cataloger’s description infers that “industries” refers to steel mills. Throughout this study, such fragmentary representations reflect ordinary cataloging practices and discourse among searchers and curators.

4.3.5.2 Case roles

Case roles help describe natural language expressions of everyday experiences. Sowa defines them as, “conceptual relations that link the concept of a verb to the concepts of the participants in the occurrent expressed by the verb” (Sowa, 2000, p.506). There is no universal agreement from one Case grammarian to the next on the type and number of case relations that exist (Somers, 1987), but as Cruse puts it, “linguists appear to feel that there is a finite number” (Cruse, 2000, p.282). This makes analysis in the current study difficult and particularly so when natural language expressions are sentence fragments and decisions have to be made whether to bring in known, but not explicitly stated facts. Consider the cataloger’s description reconstructed from incident 044 *Jacob J. Vandergrift built oil pipelines*. There are two main participants in this event: Jacob J. Vandergrift and oil pipelines. These participants, however, have different relationships to the act of building. Jacob is the doer, called *agent*, and supplies the force required to get the oil pipelines built; oil pipelines result from the action identified by the verb and fill the thematic role *patient*.

Information on case relationships in Story (1993) provides a useful synthesis of the work carried out by several linguists on this topic including Shank and Abelson (1977), Chaffin, et al. (1988), Chaffin and Herman (1984, 1987), and Landis, et al. (1987). Storey defines case relationships as events "structured according to scripts or plans" that are used to provide a knowledge structure for relationships in the everyday world (Storey, 1993, p. 470). Story describes six categories of case relationships and these are used to measure types of case relationships in the content analysis form. Three types involve agents, two types involve actions, and one type represents attributes, which is treated separately in this study.

Table 13: Frequency of case relations based on recording unit type.

Recording Unit Type	Type of Case Relationship Frequency Count							
	Agent-Action	Agent-Instrument	Agent-Object	Action-Recipient	Action-Instrument	TOTAL	Percent of sample	Cum. percent
Image searcher's query	42	1	0	6	0	49	48.0	48.0
Curator's correspondence	6	2	3	1	0	12	11.8	59.8
Cataloger's description	20	3	9	6	3	41	40.2	100.0
TOTAL	68	6	12	13	3	102	100.0	100.0
Percent of Total	66.7	5.9	11.7	12.7	3.0	100.0		

Table 13 shows there were 102 instances of the five kinds of *a priori* case relationship types measured in this study. There were 300 additional instances of 76 case relationship types not defined by Storey's categories. These are examined more closely in Chapter 5. Storey's five categories of case relations are defined as follows:

1) The *agent-action* case relationship is an event that takes place between an agent and a corresponding action, usually performed by the agent as in: *A photographer performs photography,* or more simply*, a photographer photographs.* In this example the agent is "photographer" and the action or verb is "photographs." These account for the largest number, 32 per cent, of the case relationship instances reported in the sample. In the data sample, semantic unit 011SEMNL13 includes a curator's description of a photograph of "workers acting as 'lookouts' for strikebreakers." The curator describes the image in terms of an event where workers are carrying out an activity of 'looking out for strikebreakers.' In a descriptive cataloging system where only image attributes are indexed, the semantics of this event would be lost, yet the event is the essence of the scene. The arguments to this relationship include an agent "worker" and a predicate "acting as lookout for strikebreakers." The preposition 'for' directs the action toward the strikebreakers.

The phrase 'lookouts for strikebreakers' can be explained as a *genitive of purpose* where the marker 'strikebreakers' identifies the intended recipient of the head noun 'lookout.' The genitive case in this instance describes a relationship between nouns that are nested within the overarching case relationship. Other genitive of purpose relationships occur throughout the sample. In 040SEMNL01, a curator writes in correspondence to an image searcher, "Enclosed are 'contact cards' for photographs that we have of the Carnegie Music Hall." The phrase 'contact cards for photographs' is a genitive of purpose relationship where the marker ‘photographs’ identifies the recipient of the noun phrase ‘contact cards.’

This interpretation of the *x for y* genitive case construction where both *x* and *y* are nouns, also explains “writer for Post Gazette” in 045FACTS93297 and “…in Pittsburgh for convention” in 045SEMNL55. While it is valuable to study the semantic roles of these smaller units, recognizing these features or elementary units that characterize relationships is complicated and not central to the goals of this investigation.

2) The *agent-instrument* case relationship exists between an agent of an event and the instrument the agent uses, for example: *A photographer uses a camera.* The arguments to this relationship should include an agent, the instrument, and how the agent prehends the instrument. The instrument is the recipient of the action. These account for only 12 per cent of the 75 case relationship instances. The phrase, ‘monkey is riding a bike’ explains the content of an image being described by a curator in semantic unit 027SEMNL01. The agent in this relationship is ‘monkey,’ the instrument is ‘bike,’ and the agent prehends the instrument by riding it.

3) The *agent-object* relationship exists between the agent and the object that takes part in the event or that the agent uses or makes as in: *A photographer* and a *photograph.* In this case relationship the object “photograph” is neither the recipient nor the instrument. In semantic unit 036SEMNL02, the curator writes, “I have located one photo from July 1952 of black and white children swimming together in a pool.” The curator is the agent and the photo is the object. The factual assertion statement of interest here is

located(curator,(of(photograph,'black and white children')))

where arg1 is ‘curator’ and arg2 is ‘photograph of black and white children’ and the relational term is ‘located.’ This is another example of a genitive case explaining a relationship within the overarching *located* relationship. The genitive in this instance is a *genitive of substance*, which is a feature of a significant number of relationship instances in this study. One easily recognizable category of relationship instances that can be explained as genitives of substance are *subject_of* relationships. Its consideration clarifies the features of the *subject_of* relationship and is the topic of Ch. 4 § 4.2.6.1 and § 4.2.6.2.

In the current assertion, ‘black and white children’ are analyzed as *part of* the picture. The preposition "of" is interpreted as expressing composition or substance. In other words, the claim is being made that the picture is made up of the visual information ‘black and white children.’ "Of" directly follows the noun 'photograph' indicating the material out of which the picture is made—understood in this context to be a photographic print.

4) The *action-recipient* relationship exists between actions and the thing that receives the action, for example: *take* and *photograph*. These occur commonly among image searchers’ descriptions and in total account for 22.7 per cent of all case relationships. An image researcher in 011SEMNL24 interested in the Homestead Strike of 1892 in Homestead, Pennsylvania, asks for photographs expressing, “effects of the strike on the workers and the town.” Once again, this illustrates how humans describe

images in terms of events and states of affairs. The action in this event is the Homestead strike and the recipients of the action are the workers and town.

5) Lastly, there is the *action-instrument* relationship that relates the action in an event to the instrument used in the action as demonstrated in: *Picture taking uses cameras.* A sample incident is 045SEMNL72, in which a cataloger describes an image as, “Silver trowel used in laying the Memorial Stone of the Carnegie Free Public Library…” The action is ‘laying the Memorial Stone’ and the instrument is the ‘silver trowel.’ This portion of the semantic unit is formalized as a *used_in* relationship in the corpus:

used_in(trowel,(of('laying the Memorial Stone',"Carnegie Free Public Library")))

4.3.5.3 Inclusion

Inclusion relationship, here restricted to class inclusion, is the standard *is-a* or *subtype/supertype* relationship. Winston states: “Class inclusion or hyponymy is usually expressed in the frames, ‘Xs are a type of Y,’ ‘Xs are Ys,’ ‘X is a kind of Y,’ and ‘X is a Y’” (Winston, Herrmann, Chaffin, 1987, p. 427). There were only five occurrences of inclusion coded in the sample. For example, the assertions describing relationships in semantic unit 011SEMNL34,

Growth of coke industry--workers pulling coke from beehive ovens/1890. Especially interested in coke production under Frick's ownership.

included the statements *kind_of(‘coke industry’,industry)* and *kind_of(‘beehive oven’,oven)*. There were examples of assertions using *is_a* relationship, but these were coded as ‘relationship type cannot be determined.’ An example is the assertion *is_a("Duquesne Works",'production facility')*

One possible explanation for this inability to identify and code the relationship is that the *a priori* relationship types did not account for instances of classes. An *instance* refers to what is particular in the world. Instances are also referred to as tokens and individuals. In this case, for example, “Duquesne Works” is an individual instance that stands in space and time. ‘Production facility’ is a specific class of facilities of which ‘Duquesne Works’ is an instance. Instance-level relationships are important for understanding class-level relationships because they stand in a type of instance-level relationship to some corresponding instance of a class. Put simply, it is hard to make sense of the assertion *part_of(photograph,’photograph collection’)* without understanding that an instance of the class *photograph* stands in an instance-level *part_of* relationship to another corresponding instance of the class *collection*.

To summarize, this section has shown that to say X *is_a* Y means that X is a subtype of Y. The relationship *is_a* implies *is a subtype of* not *is an instance of*. An instance is an individual. Instance level relationships are important for understanding class-level relationships. Therefore, the ontology discussed in chapter 5 extends the original relationship types to include a relationship type *instance_of*.

4.3.5.4 Meronymic

Meronymic relationships occur between something and its parts and are labeled *part_of* in the Corpus of Relationships (Winston et al., 1987). For example, factual assertion 013FACTS0413 *part_of("L1272","ACCD's smoke-control campaign")* states, “a photograph L1272 is part of the ACCD’s smoke-control campaign.” Meronymic relationships are central to the relationship types used for describing photographs’ part-whole relations with collections, but occurred in only 4.7

per cent of the total number of relationship instances in this study (see Table 19). One factor contributing to this relatively low count could be that this study focused on image descriptions and not on descriptions of how photographs are organized and stored in collections. There are six types of meronymic or "part-whole" relationships considered in this study and defined by Winston, Chaffin, and Herrmann (1987).

1) *Component/Integral Object* describes "components and the objects to which they belong" (Winston, Chaffin, & Herrmann, 1987, p. 421). Examples include:

a shutter is part of a camera
a roof is part of a house

2) *Member/Collection* involves membership of an entity in a collection, for example:

photographs are parts of collections
Harold Corsini's photographs are part of the Pittsburgh Photographic Library collection.

3) *Portion-Mass* involves homeomerous entities, that is, "Having parts which are similar to each other and to the wholes which they comprise" (Ibid., p. 423). For example:

this slice is part of a pie

4) *Stuff-Object* relation typically employs the "is partly" and "is made of" frames as in:

the bike is partly steel

5) *Feature-Activity* uses "part" to indicate features or phases of activities and processes. For instance:

paying is part of shopping

6) Finally, *Place-Area* is a type of meronymy that describes the relation between areas and places and locations within them. The place or location cannot be separated from the area of which it is a part. For example:

the Everglades are part of Florida

Based on the findings as shown in Table 13, there were instances of the first two definitions found in the sample and no instances of definitions 3-6. There were, however, other statements in the Corpus that were coded as meronymic but did not share the features of any of the above classes. Examples of these, which were coded as 'can't determine,' included:

045FACTS84234 *part_of((part_of("Cast House Crew","Eliz Blast Furnace Department)),"Jones & Laughlin")*

This instance describes people as part of a department and the department as part of the larger organization Jones & Laughlin.

013FACTS0413 *part_of("L1272","ACCD's smoke-control campaign")*

This *part_of* relationship holds between an individual photograph L1272 and its use in a specific smoke-control campaign. In the following statement made by a cataloger found in semantic unit 041SEMNL29, the content analyst interprets the statement *under(view,"Views')* as an instance of meronymy, but codes the type as 'cannot determine.'

View of Pittsburgh. Engraved 1871. Same view under 'Views' P-5468

This interpretation is based on the belief that in this context of archives, when a cataloger says ‘under Views’ s/he is referring to a subject heading under which another similar photograph with the same view is filed. This assertion could be formalized as a spatial expression, but instead the view shown in photograph P-6568 was interpreted as being part of a set of similar views, another instance of which can be found under the subject heading ‘Views.’

Finally, 027FACTS0102 *image_of(photograph,"Kiddieland")* is coded as a ‘cannot determine’ type of meronymy. This interpretation treats Kiddieland as a part of the picture. The preposition ’of’ is interpreted as expressing composition or substance. The picture is made of Kiddieland’s visual information. ‘Of’ follows a noun 'image', indicating the material (a photographic print) of which it is made.

Table 14: Frequency of meronymic relations based on recording unit type.

Recording unit type	Type of Meronymic Relationship						Percent of sample	Cum. percent
	Component/ Integral Object	Member/Collection	Stuff/Object	Feature/ Activity	Can't Determine	TOTAL		
Image searcher's query	16	16	3	1	4	40	50	50
Curator's correspondence	7	16	0	0	0	23	31	81
Cataloger's description	7	6	0	0	2	15	19	100
TOTAL	30	38	3	1	6	78	100	100
Percent of total	38.5	48.7	3.8	1.3	7.7	100		

4.3.5.5 Spatial

Spatial relationships describe where something is situated in space in relation to some reference entity or it provides motional information. Spatial expressions account for the second largest category of relationships in the study. As shown in Table 20, there are 303 instances of spatial expressions representing 18.3% of the total sample. The table also shows that the content analysis measured spatial location and motion under eight different subcategories, although there were no occurrences in two of these subcategories: 1) spatial prepositions not followed by arguments, and 2) cant' determine.

As the research unfolded it became evident that Cooper's set of necessary and sufficient conditions predicting spatial expressions did not account for the use of prepositions in describing other semantic fields. For instance, these are samples of occurrences in the sample where locative expressions (prepositions) were used in semantic fields other than spatial:

- *The old symphony, founded in 1895.*
- *Milton Lomask in first chair. [That is, Lomask holds the first chair position in the violin section of the orchestra.]*
- *Photo by [Clyde] Hare.*
- *Horse car in operation.*

Useful approaches to this problem can be found in Gruber's Thematic Relations Hypothesis as explained by Ray Jackendoff (1983). Gruber (1965) recognizes there are similar grammatical patterns that extend across unrelated semantic fields and that the semantics of the locative expression "in," for example, suggests a one-dimensional pseudospace in the temporal expression [4.8]. In other words, the mind adapts machinery it already has in place for spatial cognition to non-spatial fields (Jackendoff, 1983, pp. 188-189).

The variable in Table 20 named 'Spatial expression applied in some other semantic field' was intended to account for these instances of locative expressions in semantic fields other than spatial. As it turns out, the relationship 'The old symphony, founded in 1895,' and others like it were classified as a temporal relationships, not as spatial expressions from "other semantic fields.' This variable was retained, however, to see what other interesting semantic fields might utilize spatial prepositions other than temporal. Two example from the twenty instances that surfaced include these:

(1) 016FACTS0315 under('view_of building',construction)

(2) 041FACTS2191 in('key figure',foundation)

In example (1), the preposition *under* is used to describe an event—a building in the process of construction—and in example (2), the preposition *in* is used to express a key figure being part of a foundation or a member of a foundation.

Table 15: Frequency of spatial relations based on recording unit type.

Recording unit type	Type of spatial relationship	Categories of Spatial Relationships								Percent of sample	Cum. percent
	Static locative	Directional locative	Both static and directional	Preposition followed by another preposition	Ambiguous	Preposition not followed by an argument	Spatial expression applied in other semantic field	Can't Determine	TOTAL		
Image searcher's query	125	15	4	2	1	0	14	0	161	53.0	53.0
Curator's correspondence	31	3	0	0	3	0	4	0	41	13.5	66.5
Cataloger's description	83	13	2	1	0	0	2	0	101	33.5	100.0
TOTAL	239	31	6	3	4	0	20	0	303	100.0	
Percent of total	78.9	10.2	2.0	1.0	1.3	0.0	6.6	0.0	100.0		

Following Herskovits's terminology, the subject of the preposition in spatial expressions refers to the *located entity* and the object refers to the *reference entity* (Herskovits, 1986). The sentence fragment "mass rally at Pitt Stadium" represents a simple spatial expression. In this example, "at" is the locative preposition, "mass rally" is the subject of the preposition and "Pitt Stadium" is the object. The subject refers to the *located entity* and the object refers to the *reference entity*.

The data sample shows that catalogers describe in detail the relative position of key individuals within the frame of an image where image searchers only specify who should be present. The following constructions represent catalogers' descriptions placing individuals relative to each other where *X* is the *located entity* and *Y* is the *reference entity*.

(?X{person} ***poses_on_left_with*** **?Y{person} and ?Z{person})**
(?X{person} ***standing_left_of*** **?Y{person})**
(?X{person} ***seated_right_of*** **?Y{person})**

In the second and third examples the *left_of* and *right_of* spatial relations are signaled by the genitive case with the force of "of" placing emphasis on the named individual that follows. For example, in 041FACTS24110 *right_of("John Phipps","Andrew Carnegie"),* expressing direction with "right of" followed by the name Andrew Carnegie indicates possession: John Phipps is on Andrew Carnegie's right. A *postion_of* relation is implicit in statements like the following:

Mr. and Mrs. Carnegie and some Pittsburgh Friends. Seated, L to R, Miss Lauder, Mrs. Andrew Carnegie, Mr. Carnegie, Mrs. Samuel Harden Church. Standing, L to R George Lauder, Mrs. William J. Holland, Mrs. William N. Frew.

Several spatial assertions result from this semantic unit, the last two expressing the *position_of* relations:

- *direction_of(seated,'left to right')*
- *direction_of(left to right',identification)*
- *seated_left_of("Miss Lauder","Mrs.Andrew Carnegie")*
- *seated_right_of("Mrs. Andrew Carnegie","Miss Lauder")*

- *seated_left_of("Mrs.Andrew Carnegie","Mr. Carnegie")*
- *seated_left_of("Mr. Carnegie","Mrs. Samuel Harden Church")*
- *direction_of(standing,'left to right')*
- *standing_left_of("George Lauder","Mrs. William J. Holland")*
- *standing_left_of("Mrs. William J. Holland","Mrs. William N. Frew")*
- *position_of(Mr. and Mrs. Carnegie and some Pittsburgh friends row 1",seated)*
- *position_of("Mr. and Mrs. Carnegie and some Pittsburgh friends row 2",standing)*

In incident 45, a cataloger describes the following scene: "The first Board of Trustees of the Carnegie Corporation. Margaret Carnegie Miller, shown seated between her father and mother, is still an honorary Trustee." The second sentence introduces a triadic predicate spatial construction:

X seated between Y and Z

Another construction common to catalogers who describe images they are viewing locates objects in the geometric plane of the photographic print, as in 045SEMNL97 "Andrew Carnegie's first office extreme right," where the phrase 'extreme right' is referring to the right side of the photographic print.

In instances where image searchers requested images of people, the searchers did not specify relative positions of two or more people as was common in catalogers' descriptions. It was more common to see simple assertions using the preposition *with* to locate a person with another person or with an object as in these assertions derived from incident 33:

- *with("Margaret Carnegie","Andrew")*
- *with("Henry Bessemer",furnace)*
- *with("McCluskie",workers)*

Except for one lone instance, there was little evidence of image searchers specifying the location of picture elements within the frame of a print using phrases such as “upper left,” “far right,” or “center.” The exception came in incident 28 when an image searcher requested an image seen earlier. He described the photograph saying, “The bottom of the picture had Carnegie [sic] Library on them.” This was parsed into a factual assertion 028FACTS0101 *bottom_of(“Carnegie Library”,picture).*

In addition to these locative expressions, there were instances where catalogers and searchers expressed spatial relationships in which the subject of the preposition was a clause as in:

045SEMNL46 *Exterior of their first house in New York after they were married…*

In this instance “Exterior of their first house” is treated as the subject of the preposition “in.”

Question 17 in the coding instructions asks for a distinction between locative and directional spatial expressions. "Locative" denotes sets of places (locations) and defines a preposition that is associated with place structures and gives a locative reading, that is, it denotes a place*: in, on, under,* or *behind* (Gehrke, 2006). "Directional" indicate motion and are associated with path structures; denote sets of paths made up of places (Jackendoff, 1983). Directional prepositions define a location of an object relative to another object by means of a direction. Directions are expressed by projective prepositions (Cooper, 1968) and include *to, from, about,* and *around.* Direction and motion are indicated in incident 11 with the use of the prepositions ‘across’ and ‘to.’

Delivery of iron across the "Hot Metal Bridge": Carrie Furnaces to Homestead

There are ambiguous instances where locative prepositional phrases act as part of a directed motion event. The "from" and "to" prepositions represent motion information and may require a different interpretation than that offered in Cooper's model.

4.3.5.6 Synonym

During content analysis, synonymy was defined simply as “a relation among words with similar meaning” (Murphy, 2003, p. 133). No further distinctions were made. This section reexamines the relationships that were classified as synonymous in Table 21 and compares them with specializations that are discussed in computer science, anthropology, and linguistics.

In computer science, Raphael (1968) describes synonymy as an equivalence relation where X and Y name the same object and possess the properties of symmetry and transitivity. To be symmetric means that X is synonymous with Y implies that Y is synonymous with X. To be transitive means that if X is synonymous with Y and Y is synonymous with Z, then X is synonymous with Z.

Table 16 shows that only a few synonyms, twenty-two in all, were expressed during reference transactions and most of these represent statements made by image searchers. The content analyst classified all of them as synonym-like relationships, but not all of the factual assertions used the relationship term *synonym_of*. In incident 040, for example, a curator writes to an image searcher stating, “Enclosed are ‘contact cards’ for photographs that we have of the Carnegie Music Hall.’ The content analyst formulated a *used_for* relationship between the phrase “contact card” and “picture card catalog card.” (Figure 23 in chapter three shows an example of what is meant by contact card. Contact cards are catalog cards that have photographic contact prints glued to their surface.) The *used_for* relation does not appear to be a semantic relationship when viewed in light of Raphael’s definition. It presents itself more as terminology or natural language processing problem. It is beneficial, therefore, to look more closely at the synonymy relations by determining how the relational terms are similar to each other and then group them according to their similarities.

Table 16: Frequency of synonym relations based on recording unit type.

Recording unit type	Frequency count	Percent of sample	Cum. percent
Image searcher's query	13	59.0	59.0
Curator's correspondence	4	18.0	77.0
Cataloger's description	5	23.0	100.0
Could not determine	0	0.0	0.0
TOTAL	22	100.0	

In another example, incident 045, a cataloger describes a photograph as showing, "Dr. Arthur A. Hammerschlang president of Carnegie Institute of Technology (now Carnegie Mellon University)." In this instance the relational term *now* is used to associate the name *Carnegie Institute of Technology* with *Carnegie Mellon University* and the factual assertion is classified as a synonym type relationship. Keeping the original expression in this instance sheds new light on a type of synonymy that depends on the indexical "now." There was a particular time in history when these two names began referring to the same entity.

It appears there is no simple statement that serves as a common base for the formulation of synonymy. It is usually implied in the context of the description. For example, an image searcher in incident 011 begins by asking for *subjects/images* that *tell* a story and later they request images that *depict* events. The equivalence of *subject, tell,* and *depict* is only approximate, but their meanings in the context of a researcher's request for images is regarded as being close.

In other cases, one term is defined as being an abbreviation of another term. For example, in incident 14, a searcher refers to *B/W image scans* and in this instance the content analyst uses the formulation *abbreviation_of("B/W image","black and white image")* classifying it as a type of synonym relationship. Reconsidering it in light of Raphael's definition and the work of Casagrande and Hale, which follows, it too appears to be more of a natural language processing problem than a semantic relation expressing synonymy.

A closer look at relational models in anthropology and linguistics helps refine the analysis of synonyms in this study. As part of an American Indian language study in 1961, Casagrande and Hale (1967), pioneers in lexical semantic relations in linguistic anthropology, categorized 800 Papago folk-definitions according to relationship types. They followed a procedure that analyzed complex sentences into simple sentences and then those were reduced to binary relationships. Following this procedure, Casagrande and Hale arrived at thirteen general categories of lexical-semantic relations and included in this list is a type of synonymy relationship named *comparison*. They define comparison as, "X is defined in terms of its similarity and/or contrast with Y. Usually relates items on the same level of a taxonomy" (p. 178). For example, *wolf* is defined as "they are rather like coyotes, but they are big" and *bat* "which looks like a mouse, but it has wings" (p. 179).

Casagrande and Hale treat the *comparison* relationship as a 'comparative assertion' and represent its construction as: *X is like, or looks like Y.* A similar representation model can be applied to the synonyms in the current study. The variables X and Y are used to represent the two items being related and the relationship is described using natural language. There are some repeating patterns even in this small sample. For example, *same_as* is explicit in three of the propositions and is a specialized type of synonymy where the names being associated with one another share the same referent.

040FACTS1037	*same_as("A-939","P-7374")*
041FACTS28129	*same_as((of(image,"P-5205")),"P-5901")*
043FACTS0338	*same_as("Bellefield Bridge","Carnegie Library Bridge")*

Turning briefly to linguistics, Stasio, Herrmann, and Chaffin (1985) determined that there is one *family element* common to all similarity relations called ‘intersection of meaning’ and that the degree of intersection varies from complete overlap to total inclusion. They defined four types of *similars* as shown in Table 17:

Table 17: Types of synonym relations defined by Stasio, Herrmann, and Chaffin

Relation name	Examples
Action subordinate	Cook-Fry
Dimension similarity	Furniture-Chair
Synonymity	Car-Auto
Necessary attribution	Bachelor-Unmarried

To summarize, this section presented a more precise criterion for synonymy than was applied during content analysis. As a result, certain instances of relations that were originally interpreted as synonyms, such as ‘X is an abbreviation of Y’ and ‘X is a nickname for Y,’ were reconsidered and more appropriately labeled natural language processing problems rather than semantic relationships.

4.3.5.7 Temporal

Temporal relations are the final category of relations examined in question 8 of the coding form. The temporal relationship type accounts for about 10% of the total sample of 1,655 relationship instances. Temporality in traditional descriptive cataloging practices is concerned with birth and death dates of creators and with dates as a property of photographing and printing events. That is, catalogers say a photographic image has a capture date and a print has a printing date. In the domain of image descriptions there are other temporal entities that need explanation.

In the content analysis coding, a distinction is drawn between two categories of temporal relations as defined by Allen (1984) and described in IDEF5. They are *time-points* and *time-intervals* (Perakath et al., 1994). *Time-points* describe points in time. The date a photograph is presumably taken—known as capture date—is expressed as a point in time or *time-point*. In the data sample, the capture date is rarely expressed in the form of a proposition as in “Photograph X was taken on date Y.” Catalogers in the sample typically record a photograph’s capture date at the end of the description as an isolated date as shown in Figure 23.

Image searchers commonly express time-points when describing events. For instance, in incident 10 a searcher asks for, “1877 Railroad Riot and burning of Union Depot.” The term adopted in this study for expressing the association between an event (including photographing events) and a point in time is *date_of*. For example, the binary predicate expressing the *date_of* relation associated with the Railroad Riot of 1877 is:

date_of(photograph,”1877”)

In instances like this where an unknown photograph is expressed in the relationship, the year is associated with the common noun “photograph.” In this same reference incident, the image searcher describes another temporal relationship between the year and a named event using the preposition “in.”

in(“Railroad Riot”,”1877”).

There was only one instances where an image searcher used the *in* relationship without a reference to a year or century: “I am trying to find a photo of the stands, perhaps with fans, of Three Rivers Stadium in its early days.” This semantic unit was parsed into the following factual assertion:

029FACTS0103 *in("Three Rivers Stadium",’early days’).*

The curator’s response to this query included an explicit reference to a date, “I have three photos of the Stadium (c.1972),” which was parsed into the following factual assertion:

029FACTS0206 *date_of('photo of Stadium',"c.1972")*

The data in Table 18 indicates that 62% of the temporal relations sampled are time-point temporal relations. The second category of time-interval relationships accounts for 38% of the total temporal relationships. IDEF5 describes *time-interval* as a binary relation having three attributes: “*beginning*, which returns the beginning point of an interval; *end*, which returns the end point of an interval; and *duration*, which returns the length of an interval (Perakath et al., 1994, p. 134). The following request submitted by an image searcher, semantic unit 010SEMNL06, contains two references to time intervals and is used as an example to help explain how time-interval relations are instantiated in the corpus.

Lafayette Hall and early Republican conventions of mid-1850s; anything on Republican gatherings of that 1850-1860 era.

In this example, the phrases “mid-1850s” and “1850-1869 era” are represented as arguments in two propositions:

during(“Republican conventions”,”mid-1850s”)
during(“Republican gatherings”, "1850-1860 era")

The relationship represented in these assertions is *during* where arg1 in both instances represents events and arg2 represents time intervals. These two relational assertions are considered instances of the *during* temporal relation, which is a type of *time-interval* relation. IDEF5’s definitions of “during” are not strictly applied here because IDEF5 specifies that a “during” temporal relation is a binary relation between two time-intervals. The relationship literature consulted during this investigation (Table 6, Ch. 3, §3.3) explicitly excludes temporal information and IDEF5 methodology limits what can be represented about temporal relations that hold in the domain of photographs, as well as objects involved in the internal structure of image descriptions.

Table 18: Frequency of temporal relations based on recording unit type

Recording unit type	Categories of Temporal Relationships			Percent of sample	Cum. percent
	Time-point	Time-interval	TOTAL		
Image searcher's query	60	52	112	67.9	67.9
Curator's correspondence	13	4	17	10.3	78.2
Cataloger's description	30	6	36	21.8	100.0
TOTAL	103	62	165	100.00	
Percent of total	62.4	36.6	100.0		

One final comment should be made regarding time-point and time-interval expressions in image descriptions. The sample exhibited fourteen propositions where the semantic structure of English sentences contained references to temporal relationships without explicitly stating a date or range of dates. The task in each of these instances is for the content analyst to make clear what is being suggested by the relationship term. This has been difficult to do. Previous proposals in LIS have left unanalyzed many phenomena including important temporal references like the following prepositional-like structures:

- 045FACTS4657 *after('exterior of house',marriage)* – the construction *X after Y* is used to mean *X follows Y in time* or *X is later than Y.*
- 045FACTS0203 *at("Andrew Carnegie",'all ages')* the preposition construction *X at Y* is used to mean *X and Y occur simultaneously.* The assertion *taken_at(photographs,'other times')* is a simple statement, but is semantically complicated when the temporal reference 'other times' is interpreted to mean a simultaneous occurrence.
- 026FACTS0102 *before('Red Light District',"Cultural District Transformation")* – the preposition construction *X before* Y is used to mean *X earlier than Y in time*
- 015FACTS0106 *from(images,'same day')* – the construction *X from Y* is interpreted as meaning *X with origin or stating point Y.*

- 011FACTS0208 *in(event,"American history")* The construction *X in Y*, occurs often in temporal relations where arg2 is a date or time interval. This is an instance where there is no date, but the assertion claims that an overall event or some aspect of an event is located in American history. Deeper analysis shows that the semantics of this proposition are related to an earlier proposition in the correspondence where the image searcher explains, “I am working on picture research for a History Channel documentary about the Homestead Strike of 1892 in Homestead, Pennsylvania.” This makes explicit that ‘event’ refers to an interval of time during which the Homestead Strike took place during the year 1892. This makes the structuring of interpropositional relations important, but this is not the topic of the current investigation.
- 011FACTS3382 *through(images,'turn of century')* – Temporal use of the preposition *through* normally indicates the end point included in range with *from* being implicit as in “Images *from* 1850 *through* the turn of the century.” But in this example the beginning of the range is not state explicitly.
- 010FACTS0114 *time_span_of(pictures,'300 years')* – The *of*-construction is applied when there is an interchangeable genitive counterpart.[14] The genitive counterpart of “time span of 330 years” is “300 years’ time span” so the *of*-construction is applied. The only other example of *time_span_of* in the Corpus is parsed from incident 10 where an image searcher asks for pictures of “Black Pittsburghers of any era voting or in political situations.” This is parsed into various assertions, including *time_span_of("black Pittsburghers",'any era')*. The *of*-construction is interchangeable with the genitive counterpart “any era’s time span,” so the *time_span_of* relationship is applied. This factual assertion would work equally well with the *during* relationship. This study treats both assertions as instances of binary *time-interval* relationship.

Before leaving the analysis of samples occurring in *a priori* relationship categories, a final section addresses the class of relations labeled “Can’t Determine.”

[14] Determining applicability of *of*-constructions to temporal relationships was influenced by the work of Rolf Kreyer (2003) and the distinctions he makes between genitive and *of*-constructions.

4.3.5.8 Can’t Determines

During the first pass of analysis, not all relationships were neatly captured under *a priori* headings. This resulted in set of 150 “Can’t Determine” relationships. The following examples indicate the flavor of some of these unclassified relationship instances.

1. 045FACTS89274 *at("Margaret Carnegie","age 22")*
2. 045FACTS89272 *child_of("Margaret Carnegie","Andrew Carnegie")*
3. 041FACTS0313 *depicting(pictures,"Andrew Carnegie")*
4. 037FACTS0101 *down(looking,"Ohio River")*
5. 041FACTS29133 *engraved("P-5468","1871")*
6. 038FACTS0411 *cropped_at(photo,angle)*
7. 041FACTS0314 *depiction_of(pictures,'steel empire')*
8. 016FACTS0523 *from(taken,distance)*
9. 045FACTS69164 *imbued("George E. Lauder",(with("Andrew Carnegie"),'love of poetry'))*
10. 023FACTS0620 *instance_of("Hand Bridge",'covered wooden bridge')*
11. 043FACTS0525 *view_of(exterior,"Forbes Field")*
12. 045FACTS93298 *writer_for("Donald Miller","Post Gazette")*

Two points can be made about these unclassified relationship instances. The first is that there are obviously borderline cases where the analysis could be carried out different ways. For instance, example (1) could be interpreted at least three different ways. This is the original semantic unit from which the factual assertion was reconstructed:

Margaret Carnegie. Andrew Carnegie's only child. Married at age 22 to Roswell Miller, Naval Ensign. P-5238

In one reading, the factual assertion in (1) could be classified as a kind of temporal relation between a person and her age where the preposition *at* refers to a particular place located on a timeline. In another interpretation it could be interpreted as an *age_of* relational proposition of attribution. That is, a person is said to have an age, as in *age_of(”22”, “Margaret Carnegie”)*. And finally, the factual assertion could be specified as a *marriage_age_of* relationship holding between Margaret and 22, as in *marriage_age_of(“”22”, “Margaret Carnegie”).* It is reasonable to say that in situations like this clear criteria are needed for making choices. Again, there are two possibilities: a relationship is classified based on what can be recognized as grammatical distinctions in the original semantic unit; or an attributive relation can be justified if an *of*-construction is available in any context in the domain of interest. The *date_of* attribute is used in 50 other factual assertions and the *marriage_date_of* could be considered a specialization of *date_of*. There already exists in the Corpus *death_date_of* of and *birth_date_of* relations.

The second point is that after the analysis of the sample is complete and viewed as a large collection of instances, new, domain-related families of relationships become apparent that are not easily recognized when parsing one factual assertion at a time. For example, among the twelve Can’t Determine instances listed at the beginning of this section, three emerge as *viewing* or *point-of-view* kinds of relations, including *down(looking,"Ohio River"), from(taken,distance),* and *view_of(exterior,"Forbes Field").* This suggests an important question concerning the content analysis method applied to capturing relationships. How effectively can an analyst predict relationship types in a domain of interest by proceeding in a strict linear fashion—one incident, one semantic unit, and one factual assertion at a time? It is shown during *post analysis* that the possibility of making significant generalizations about relationship instances seems to be a strong motivation for reconsidering the class of Can’t Determines after the Corpus is complete.

Two general examples serve to make this point: first, repeating patterns across many incidents sometimes point to a new relationship category unique to a particular domain, so it is not surprising to see a number of *viewing* type relations emerging in the domain of image descriptions—relations not specifically defined by the *a priori* categories; second, when a wide range of natural language statements are examined, it appears that exceptions to rules crop up again and again and there is no clearly defined category in which to place instances of a particular kind of relationship. This was especially true with case roles where many accounts of roles were difficult to distinguish applying Story's five categories discussed in § 4.3.4.2.

First, the point of new emerging relationship types is considered with what the data suggests is a *viewing* class of relationships. A solution for their explanation is sought in Gruber's Thematic Relations Hypothesis.[15] This approach seems simple and direct. The viewing category is especially interesting because it is unique to the domain of visual information and includes the relations *point_of_view*, *view_from*, and *view_of* relationships. Representative instances are taken from a cross section of incidents and presented in Table 19.

[15] Gruber's Thematic Relations hypothesis states that in any semantic field of states and events, the principle state, event, path, and place functions are a subset of those used for analyzing spatial location and motion. Fields vary from one another in only three possible ways: 1) The kinds of things that emerge as located entities, 2) the kinds of things that emerge as reference entities, and 3) the type of relationship type that fills the role played by location in the semantic field of spatial relations.

Table 19: The point_of_view, view_point, and view_of relationships

Column 7	Column 8	Column 9	Column 11	Column 12	Column 14
FACTS#	FACTS	RECTP	ARGU1	ARGU2	RELAT
041FACTS0949	point_of_view(aerial,photograph)	1	aerial	photograph	point_of_view
040FACTS1141	point_of_view(looking_out,"Dorothy Height")	1	looking out	Dorothy Height	point_of_view
040FACTS1140	point_of_view((on('Music Hall interior',stage)),camera)	1	Music Hall interior on stage	camera	point_of_view
033FACTS1520	view_of(present,"original location_of Col. Anderson Library")	1	present	original location_of Col. Anderson Library	view_of
032FACTS1127	point_of_view(camera,distant)	2	camera	distant	point_of_view
033FACTS2738	point_of_view('exterior shot',"Braddock Library")	2	exterior shot	Braddock Library	point_of_view
016FACTS0626	point_of_view(aerial,"Civic Arena")	2	aerial	Civic Arena	point_of_view
040FACTS0306	view_from('stage area',stage)	2	stage area	stage	view_from
040FACTS0307	view_from(audience,stage)	2	audience	stage	view_from
016FACTS1031	view_of(closeup,'Civic Arena exterior'\)	2	closeup	Civic Arena exterior	view_of
016FACTS1033	view_of(photograph,'Civic Arena exterior')	2	photograph	Civic Arena exterior	view_of
045FACTS82220	point_of_view(interior,'steel mill')	3	interior	steel mill	point_of_view
043FACTS0525	view_of(exterior,"Forbes Field")	3	exterior	Forbes Field	view_of
041FACTS29132	view_of("P-5468","Pittsburgh")	3	P-5468	Pittsburgh	view_of
045FACTS73189	view_of("P-5975","West Side")	3	P-5975	West Side	view_of

Within this category of relations, the nouns "interior" and "exterior" designate indoor space and exterior surfaces of structures from the perspective of the camera or photographer taking the picture, which presents a particular viewpoint in the picture.

- *camera views Music Hall interior*
- *camera views exterior of Braddock Library*
- *camera views exterior of Civic Arena*
- *camera views of exterior of Forbes Field*
- *photograph of exterior of Civic Arena*

On the one hand, if the informational emphasis is placed on the subject being photographed, it is reasonable to interpret these factual assertions as describing spatial expressions and thus they should be categorized as instances of spatial relationships. On the other hand, if emphasis is placed on the action that is taking place in the photographing event, the assertions could be interpreted as case relations that relate the action of viewing to the camera. Point of view, therefore, is said to be present in the image and can be explained as an extrinsic relationship between the camera and the subject. Thus, depending on which perspective is taken, *view-type* relations can be considered members of spatial or case relations.

In Table 25 there are two instances of *view_from* relations.

view_from('stage area',stage)
view_from(audience,stage)

Gruber states that in applying the Thematic Relations Hypothesis, semantic fields vary in only three possible ways: the types of entities that play the role of located entity, the types of entities that play the role of reference entity, and the type of relationship that assumes the role played by location in locative expressions. In the two instances stated above, the located entities are “stage area” and “audience.” These are the starting points or initial references from which viewing originates. The reference entity is “stage” in both instances and represents the endpoint or destination. The relationship *view_from* expresses traversal of a path from the located entity to the reference entity.

Thus far it has been determined that the relations *view_from*, *point_of_view*, *view_point*, and *view_of* share certain properties in common: all of them describe a viewing event involving a viewing agent and points of origin and endpoints in space that are viewed. The eight remaining relationships illustrated in Table 19 are now considered in light of these properties. The assertions 3 and 5-7—the remaining *view_of* relationship instances—all share the same properties as the *view_of* relations analyzed earlier. The term “present” in assertion 3 is unique in that it marks a time-point as the point of origin rather than a point in space.

1. *point_of_view(aerial,photograph)*
2. *point_of_view(looking_out,"Dorothy Height")*
3. *view_of(present,"original location_of Col. Anderson Library")*
4. *point_of_view(aerial,"Civic Arena")*
5. *view_of(photograph,'Civic Arena exterior')*
6. *view_of("P-5468","Pittsburgh")*
7. *view_of("P-5975","West Side")*
8. *point_of_view(camera,distant)*

The four remaining *point_of_view* relations can also be defined using the same general properties described thus far with some refinements. In assertion 2, the agent is a speech competitor named Dorothy Height, which means that the viewing agent may be an individual. This same assertion is unique in another way. The term in arg1 describes a property of Dorothy Height, which suggests a more accurate formulation for this assertion would be *view_of(“Dorothy Height”,looking_out)*. This formulation clarifies that the viewing event’s point of origin is Dorothy Height. The phrase in arg2 “looking out,” however, is not typical of the nouns representing endpoints in the other assertions. The phrase “looking out” suggests a path or line of sight with no particular endpoint in mind. Another look at the original semantic unit from which this factual assertion was parsed suggests that the endpoint is the audience. That is to say, she is standing on stage looking out, presumably at the audience. The only other anomaly occurs in assertion 3, arg1, where the viewing point of origin refers to a point in time rather than a spatial location.

To summarize, all of the *viewing-type* relationships discussed in this section describe a viewing event along three dimensions: 1) there is viewing agent, which can be an individual, an apparatus (e.g., camera), or object depicting an image (e.g., photographic print), 2) a viewing point of origin, which is either a place, path, or time-point, and 3) a viewing endpoint fixed in an object, place, or path.

Next, the distinction between case roles is considered. This is an example of one of the more perplexing problems in this study because Story's (1993) five categories do not explain explicitly how to define case relations by semantic predicates. For instance, Story describes the first case relation—*agent-action*—as an event that takes place between an agent and a corresponding action, usually performed by the agent. A classic example of the standard theory of *agent-action* is *John opened the door* (Fillmore, 1968). There are two problems using this definition and example in the current study. First, such sentences as *John opened the door* are not likely to occur in image descriptions made by catalogers, curators, or image searchers, so there is the problem of recognizing unconventional instances of *agent-action* in the domain of image descriptions based on conventional linguistic examples. Second, there is the problem of representing this relationship and the other case relations as factual assertions. To say in Story's terminology that "*John opened the door* is the relation of agent-action," or simply, "John is the one who is doing the action," does not meet Hood's criteria of relation—that is, an expression containing two arguments and a relational term. Therefore, the convention in this study during content analysis has been to represent sentences like this as *open("John",door)*. While this expression meets Hood's criteria, it does not make clear which participant is the agent and which participant is the recipient of the action.

This reexamination of case relations now turns to the data sample for a real-world example of agent-action case relations. In incident 011 an image researcher writes to the curator of the Pittsburgh Photographic Library with a list of subject requests concerning the Homestead Strike of 1892 in Homestead, Pennsylvania. One of the subject requests (1) and its corresponding factual assertion (2) are presented here:

(1) *workers storming the barges*
(2) *storming(workers,barges)*

Factual assertion (2) is categorized in the Corpus under Column 15 as case relation type 1, which means the agent-action type case relationship is applied. This explains the case expressed at the surface, but there are two participants in this relation—workers and barges—and only one is being accounted for as an agent. The solution proposed here, inspired by Kuntz (1987), is to adopt a predicate, or *semantic case label* that explains arg2 in the relation and this opens the possibility for explaining in a semi-formal manner subject requests like *workers storming the barges*:

(3) *has_agent(storming,workers)*
(4) *has_recipient(storming,barges)*

Factual assertions (3) and (4) describe the agent and patient of storming. Factual assertion (3) is an example of Story's agent-action relation and it makes clear "worker" is the agent. Factual assertion (4) brings to light a different nuance. Story (1993) describes an action-recipient relationship that exists between an action and the thing that receives the action and (4) describes the recipient of storming. Story's definition might then translate into a dyadic relation of the form *recipient_of(barges,storming)*. Both constructions are possible alternatives, that is, where *has_recipient(x,y)* can be substituted by the *of*-construction *recipient_of(y,x)*.

Turning now to the Can't Determine category, incident 014 includes an assertion that had classification problems centering on the actions of *opening* and *naming*. The curator is writing to an image researcher and states:

The first federal hospital was opened in 1778 and named after General Hand.

Post analysis and the new approach for representing case relations bring new light to the problem. The original analysis failed to recognize the case relation in (1), labeling it as a temporal relation, and the original analysis failed to recognize any relationship type in (2).

(1) *open('first federal hospital',"1778")*
(2) *named_after('first federal hospital',"General Hand")*

While the semantic case label helps solve the dyadic predicate problem, a new problem surfaces in (1), namely, the date or point-in-time participant—an object not explained by Story's five categories. This suggests a need to return to the literature to examine other families of case relations. When faced with the problem of how to formally define case relations, Kuntze (1987) observes, somewhat tongue in cheek, "there are more proposals and systems than authors" (p. 302). Sowa (2000) does an adequate job of synthesizing the work of many of the authors including Moravcsik (1973), Somers (1987), Dick (1991), and Pustejovsky (1995), creating a matrix that extends Story's categories to nineteen. While adding more categories makes classification decisions more difficult, Sowa sheds light on the case relation of current interest as well as providing semi-formal definitions of each case relation in his proposed matrix. With assistance from Sowa's matrix, factual assertions (1) and (2) can be formulated as case relations (3) – (6).

(3) *has_PointInTime(opening,"1778")*
(4) *has_theme(opening,'first federal hospital')*
(5) *has_beneficiary(naming,"General Hand")*
(6) *has_theme(naming,'first federal hosptial')*

Even with the force of a completed Corpus there remains considerable tension between the ambiguity of the Can't Determine expressions and specializing those expressions until they can be classified into families of relations. For example, there were six Can't Determine assertions where image searchers, and in one case a cataloger, expressed explicit *showing* relationships.

041FACTS1661	*showing(pictures,"Andrew Carnegie")*
041FACTS0521	*showing('fiery,dramatic scene',"Bessamer furnace")*
036FACTS0309	*showing(pools,'blacks and whites swimming together')*
041FACTS0836	*showing(scenes,at(business,work))*
040FACTS0615	*showing("P-5593",'organ and organ case')*
041FACTS1767	*showing(view,'steel plants')*

In the first example in which arg1 is "pictures," *showing* seems interchangeable with or equivalent to *subject_of* and could be reformulated as a *subject_of* relationship without changing the semantics of the relation. In the next two assertions, however, *subject_of* does not make sense. A practical solution for these instances was to interpret the activity 'show' as a semantic frame and reformulate certain factual assertions as case relations. This resulted in generating a series of semantic case relation *frames* describing interesting events and activities that had occurred throughout the Corpus of relations. For instance, Table 20 illustrates an example of the *take event* and its case roles, including agent, point in time, motivation, subject, and location. The context is a cataloger describing photograph explaining when the picture was *taken*. Sometimes there are events within this event, for example, the event of a baseball game being captured by the event of taking a photograph.

Table 20: The 'take' semantic frame

043FACTS0329	has_agent(take,photographer)
043FACTS0330	has_PointInTime(take,"July 5, 1909")
043FACTS0331	has_motivation(take,"Independence Day"(event_of(photograph)))
043FACTS0332	has_subject(take,'traditional fireworks'(feature_of("Independence Day")
043FACTS0333	has_subject(take,'balloon ascension'(feature_of("Independence Day")
043FACTS0334	has_location(take,"Schenley Park")
043FACTS0335	has_subject(take, 'baseball game' (event_of("Independence Day)))
043FACTS0336	has_subject(take,has_feature('baseball game',doubleheader))
043FACTS0337	has_subject(take,has_location('baseball game',"Forbes Field"))

This completes the analysis of *a priori* categories of relationship types listed in the coding form under question 8 with their subtypes listed under questions 13-17.

4.4 CONCLUSIONS

The analysis in this chapter was carried out for two reasons: First, it was hypothesized that relationship instances could be captured from natural language expressions, recorded in a Corpus of Relationships, and their types predicted using ontological tools and methods; and second, no description of relationships in this domain, based on empirically-grounded evidence, existed.

The main body of analysis was divided into eight sections according to broad relationship types: 1) Attribution, 2) Case roles, 3) Inclusion, 4) Meronymic, 5) Spatial, 6) Synonym, 7) Temporal, and 8) Can't Determines.

Evidence shows that Attribution is the largest category of relationship instances in the study, accounting for 39% of all relations, and an important factor in characterizing how catalogers and image searchers express meaning when describing images. A breakdown of all 643 instances of attribute-type relations shows the predictable result that the *subject_of* attribute accounts for the majority, or 42% of the total. One surprising finding in this category is that there were only six occurrences of *name_of* and *caption_of* (or the inverse *has_caption)* attribution and no occurrences of *title_of*. It is common knowledge that most institutions who adopt description standards insist on assigning titles to photographs in order to identify the photograph (Klijn, 2003). This signifies that photograph titles may be cued by LIS description standards and would otherwise not be a part of the discourse surrounding images.

The category of case roles was particularly problematic because there was no explicitly stated method in the literature for formulating this type of relation as a dyadic predicate. Moreover, recognizing specific case relationship categories was made difficult by Story's (1993) informal definitions. In spite of these difficulties, progress was made in this category during analysis of the Can't Determine category where there were many unidentified sub-categories of case relations. There a system was developed for expressing case relations as dyadic predicates and Story's (1993) set of five case roles was extended adopting Sowa's (2000) nineteen thematic roles. A problem remains, however, in how best to capture and represent as factual assertions the events and actions found in image descriptions and whether case roles are the most effective tools for doing this. There is reason to believe that solutions will emerge by carrying out deep analysis in domains where descriptions of photograph processes and events are more prevalent, for example, in art history, ethnography, and the fine arts.

A surprising finding in the category of Inclusion, which is the backbone of LIS thesauri and taxonomies, was that there were only thirty-three occurrences of *kind_of*, *is_a*, and *type_of* relationships. A possible explanation for this is that while taxonomies are useful for structuring controlled vocabularies, inclusion relations are not a naturally occurring phenomenon in the discourse surrounding image descriptions. In fact, the *is_a* relationships recorded in the Corpus were categorized as Can't Determine because they were not true class/superclass relations, but rather class *instances*. This brought to light that the *instance* relationship was not included in the coding instructions and form.

Three important findings resulted from analysis of Meronymic relationships, which represented only 4.7% of total instances: First, only two out of the six subcategories of part-whole relationships being tracked in the study occurred with any frequency in descriptions. This is surprising considering that part-whole relations are critical for describing the unique nature of photographs belonging to named collections housed within departments that are parts of institutions. Then again, this is only one viewpoint of how images are organized and stored and evidence in this study shows that in only one instance did an image researcher describe an image in terms of its location within a collection; second, in certain cases relationships were determined to be instances of more than one class of relationship type, so it is imperative that computational systems representing relationships support multiple hierarchical inheritance; and third, there were a substantial number of cases in the Corpus where the preposition "of" and certain locative expressions were interpreted as having partitive effects, but did not neatly fit within one of the *a priori* meronymic categories. The interpretation applied in the Corpus went something like this: in constructions like **(?X{photograph} *image_of* ?Y{thing})** and **(?X{photograph} *picture_of* ?Y{thing})***)*, where ?X represents an unknown photograph, ?Y describes a subject, and when *of* follows "image" or "picture," the *of*-construction is interpreted as expressing composition or substance. That is, the picture or material (the print) is interpreted as being made of visual information ?Y. This could be formulated as *part_of(subject,photograph)*, which offers an interesting alternative view of image subject.

A total of 303 spatial instances represented the third largest category of relationship types in the study just behind case relations, which represented 24.4% of the total. One particular finding that stood out was the absence of spatial expressions locating picture elements by describing their position on the geometric surface of the print, for example, upper-left corner or bottom. In Benson (2010), fine art photographers who were members of a Pittsburgh artist's consortium were observed during meetings discussing individual prints and series of prints. In that setting it was common for photographers to describe images by making reference to spatial locations on the surfaces of prints. This suggests that while photographers describe images differently than image searchers, curators, and catalogers, the ontological relationship categories developed for organizing relationship instances should be general enough to accommodate these differences.

At the outset, synonymy was treated as a primitive with no subcategories. Evidence shows, however, that while synonymy is a kind of 'similar' relationship type, it did not account for every kind of similar in the data sample. This suggests that the *equivalence* relation used in library thesauri (that is, the *use* and *use for* relations) may be semantically overloaded.

Finally, a critical discovery, although not surprising, is that unlike catalogers' and curators', image searchers in only rare instances describe known photographs. This carries with it significant implications for knowledge sharing and reuse and the practical use of relationship corpora that include image searchers' descriptions. The argument can be made, and justly so, that a content analyst cannot build factual assertions from image searchers' descriptions of unknown, possibly non-existent photographs because these assertions are neither true nor false. In other words, what an image searcher states about photographs cannot be considered for inclusion in an information system until the arguments of the relationship are assigned referents. The image searcher who said in incident 036, "*I am looking for photos of Pittsburgh pools showing blacks and whites swimming together*," and the analyst's reconstruction *showing(photos,'blacks and whites swimming together')* could describe an infinite number of photographs. That is, photos showing *blacks and whites swimming together*. The common noun *photo*, for instance, does not in itself communicate anything, and the assertion that there is a "photo showing blacks and whites swimming together," on its own, is not an item of knowledge. It is, as D. Alan Cruse puts it, a naked proposition.

The content analytic method introduced in this study addresses the argument straight on in the following manner. In the context of the relationship corpus, the factual assertions derive their energy from the force of the incident. That is, the analyst claims the assertion is true of an image searcher's request, which is assigned a unique alphanumeric code. It would be cumbersome to notate every assertion using codes, but if it were applied, the earlier stated assertion might be expressed as

showing(photos$_{036FACTS0309}$*,'blacks and whites swimming together')*

As described in the beginning of this chapter, the content analysis form adopts an alphanumeric coding system that links the original correspondence and catalogers' descriptions to semantic units in the Corpus and from semantic units to factual assertions. The usefulness of image searchers' assertions in future knowledge base systems, however, is questionable. Their purpose in this study is to contribute to the Corpus of how humans express relationships and that in turn informs the families of relationship types that emerge in the ontology of relationships introduced in chapter five.

Little is known about what constitutes an ontology of relationship types in the domain of image descriptions. Structuring and representing an ontology of relationships is the task carried out in the next chapter.

5.0 ONTOLOGY OF RELATIONS

The main portion of this study used content analysis techniques to ascertain how humans express relationships in English when searching for images in the Pittsburgh Photographic Library. The research examined correspondence between curators and image researchers, which at times also involved catalogers' descriptions. Within this discursive space, relationships associated with image descriptions were identified, analyzed, and recorded in a Corpus of Relationships. The Corpus shows transparently how the original natural language expressions are reconstructed as semiformal assertions, each one representing one instance of a relationship.

The view taken in the Corpus was based on the types of concepts being related. In other words, assertions were broken down into relational terms and arguments and these components were studied in order to classify them into categories of relationship types. The view taken in this chapter is different. Emphasis is placed on specifying and classifying relationship types into broader categories of relationships and on the conceptualization and specification of the image description itself. The goal of the current chapter, therefore, is to determine the properties of image descriptions, how image description relationships are connected with one another, and provide an empirically based ontology of the relationship types that emerge from the Corpus of Relationships. This work is important for three reasons: 1) It emphasizes organizing the image description concept and its properties based on semantic relations; 2) It establishes a hierarchy for classes of relationship types; and 3) It represents at the symbol level how relationships are used for inference.

The chapter opens with a clear statement of how to measure success followed by a definition of ontology and an explanation of how an informal specification of the ontology will be made. The balance of the chapter is divided between two main sections: 1) The first section illustrates how an ontology of image description relationships (OIDR) looks when represented as a semantic network consisting of nodes linked together with relations. It explains diagrammatically (Figure 30) how an instance of a relationship within an image description links conceptually to a relationship type in an ontology of relationships. 2) The second section specifies the classes or *families* of relationships that emerge from the Corpus of Relationships. These link conceptually to the {relation} concept in Figure 30.

5.1 CIRCUMSCRIBING ONTOLOGICAL BOUNDARIES

In one respect, the task of modeling an ontology of relationships is similar to the task faced in predicting relationships in the Corpus of Relationships. While building the Corpus, the logical syntax of relations—that a relationship instance had to specify certain features—helped define the goals of the analysis and it provided a means for measuring success. Now the task is to circumscribe the ontological boundaries of each relationship type. To accomplish this, each particular relationship type is defined. That is, conditions are made explicit for being a type of temporal relationship, for example, as opposed to being another sort of relationship.

Characterizing relationship types according to semantic primitives extends the definition of a relationship type. Semantic primitives are elementals, fundamental properties that cannot be explained by using other primitives (Huhns and Stephene, 1989). They are general enough to apply to relationship types, but not instances. The Relation Element Theory of Chaffin and Herrmann (1987) exploits the idea of semantic primitives, differentiating relations by relation elements. For example, the proposition "Photograph L1273 is part of the ACCD's smoke-control campaign," was analyzed in the Corpus as being a type of part-whole relationship. In this chapter it is important to clarify the nature of the relationship types, identifying which types of entities they relate. Using this same proposition as an example, the relationship can be explained as an **(?X{physical object}** ***part_of*** **?Y{event})** construction where **?X** represents an individual photograph (a physical object) and **?Y** represents a smoke control campaign (an event).[16]

This study also proposes adopting two primitives, symmetry and transitivity, based on a literature survey of Huhns and Stephene (1989), Perakath C. B. et al. (1994), and Winston, M., Chaffin, R., and Herrmann, D. (1987). These primitives are chosen as being useful for the purpose of describing properties of relations in the domain of image descriptions. Symmetry and transitivity are defined as follows:

> **Symmetrical:** A property that holds of a given relation R if, for any objects x and y (of the appropriate kinds), if x bears the relation R to y, then y bears R to x. For example, in the relation *brother_of*, if x is the *brother_of* y, then y is the *brother_of* x. Other kinship relationships are *asymmetrical*, for example *mother_of.*
>
> **Transitive:** A property that holds of a given relation R if, for any objects x, y, and z (of the appropriate kinds), if x bears the relation R to y, and y bears R to z, then x bears R to z. For example, if Dorothea Lange's photograph "Migrant Mother" is *part_of* the Farm Security Administration-Office of War Information Collection and the Farm Security Administration-Office of War Information Collection is part of the Prints and Photographs Division of the Library of Congress, then the photograph "Migrant Mother" is part of the Prints and Photographs Division of the Library of congress.

---

[16] The semantic construction notion adopted here is taken from Fahlman's (2008b) blog *Knowledge Nuggets*. The pattern **(?X{physical object}** ***part_of*** **?Y{event})** is a generalization of a more specific part-whole relation. The term "photograph" is generalized to the semantic category {physical object} and the term "campaign" is generalized to the semantic category {event}.

5.2 DEFINING ONTOLOGY

This study adopts Sowa's (2000) definition of ontology as "a catalog of the types of things that are assumed to exist in a domain of knowledge $\mathcal{D}$ from the perspective of a person who uses a language $\mathcal{L}$ for the purpose of talking about $\mathcal{D}$" (Sowa, 2000, p. 492). The domain of interest examined in this investigation is the image description. The observable events in this domain are the entities that the researcher perceives as being relationships. The 'types' in this ontology are concepts and relationship types. The relationship instances recorded in the Corpus of Relationships, which can be validated against the original semantic expressions made by catalogers, curators, and image searchers, are the researcher's ontological commitments.

This study presents a *semi-formal* ontology specifying concepts and relationship types in natural language. This is not to say the present ontology is less rigorous than a formal ontology. However, there are ambiguities in natural language that are removed when expressing these same concepts and relationship types in higher levels of explicitness.[17] Representing Corpus relationships in machine-understandable language is outside the scope of this dissertation.

To summarize, there are four levels of informal specification that can now be identified and adopted in this chapter. The first offers a semi-formal definition of the relationship types in natural language. The second expresses relationship types in a semantic network labeled nodes and edges. The third explains relationships in terms of their properties. And the fourth represents a relationship as a Semantic Construction generalized to the types of semantic categories joined together by the relationship.

---

[17] Formal information systems ontologies are expressed in formal languages such as SUO-KIF (http://ontolog.cim3.net/file/resource/reference/SIGMA-kee/suo-kif.pdf) or Scone's representation system (http://www.cs.cmu.edu/~sef/scone/Scone-User.htm).

A major challenge for the Library and Information Science (LIS) community is how to make the complexity and practical knowledge of ontology approachable and usable by everyone who has knowledge organization and representation needs. For the task of modeling an ontology of relationships with future applications in LIS, the argument is made that relation specifications must follow a simple and concisely stated natural language format that is accessible to indexers, catalogers, and other information professionals in fields as diverse as biomedical information and art history.

5.3 METALANGUAGE FOR ANALYZING IMAGE DESCRIPTIONS

Discussing an ontology of relationships in the domain of photographs and photographic images requires drawing distinctions between real-world events, abstract structures humans impose on images, concepts, and textual representations of these entities. In Ch. 1, §1.4, a preliminary framework was presented for understanding meaning as it pertains to image descriptions, defining five categories of meaning: referent, depicted referent, object showing depiction, text, and the conceptual layer. It is useful to introduce a metalanguage that accounts for these realms as this research moves from natural language semantic units and factual assertions to more explicit specifications in computational environments.

To help distinguish between these layers of meaning, this chapter adopts the following conventions: Words that signify concrete substances—that is, referents and the objects depicting referents—will be left unaltered (e.g., photographic prints, steel mills, human beings); text symbolizing depicted referents will be placed in double quotes (e.g., the name “Harold Corsini” refers to an individual human being and the name “Homestead Steel Mill” refers to an individual steel mill); concepts will be placed in curly brackets (e.g., {Harold Corisin}, {steel mill}, {photograph}).

5.4 ONTOLOGY OF IMAGE DESCRIPTION RELATIONSHIPS

In this first part, the ontology of image description relationships (shortened to OIDR from this point forward) is presented as a graph-theoretical structure consisting of terms that form the nodes of each corresponding graph, which are linked together by means of edges called *relations*. The edges (relations) are applied in a manner consistent with the Relationship Corpus. In the section immediately following this section, the OIDR is presented as a taxonomy of relationship categories or *families*. A standard syntax is employed throughout to help ensure future integration with a knowledge base. These criteria help ensure reliable curation of the ontology and formal structures and building a solid base for future ontology development in information sciences.

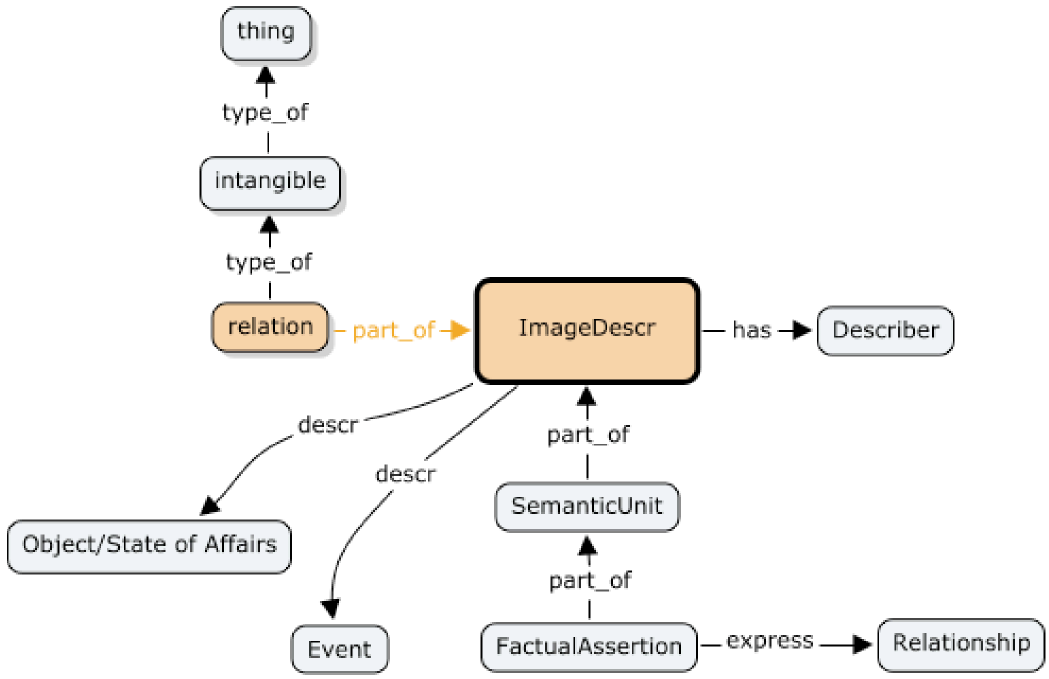


Figure 28: Concept *relation* joined with the concept *ImageDescr*

Knowledge in the OIDR consists of instances, classes, and slots or properties. In the introductory Figure 28, the goal is to represent how the {ImageDescr} (image description) concept is situated in the landscape of Scone's core knowledge base.[18] Scone, an existing knowledge base and upper-level ontology, is used in this study to help illustrate how the concept of an image description might connect with other, more fundamental units of existence. Following Scone's upper ontology, the top node shown in Figure 28 is the root concept {thing}

[18] Scone is an open-source knowledge-base system currently under development in the Language Technologies Institute of Carnegie Mellon University (Fahlman n.d.) Scone supports representation, searching and limited forms of common sense reasoning, all features that are useful for expressing and using knowledge relating to image descriptions. Knowledge is stored in Scone as a set of files written in a computer language called Common Lisp (Fahlman 1982). Its application here is to provide a framework for linking the concept of an image description and its properties to an existing upper-level ontology.

with subclasses {intangible} and {relation}. The concept {ImageDescr} and its association with the concept {relation} are established at this level in the ontology. The concept *relation* is *part_of* the concept {ImageDescr}.

The concept {ImageDescr} has certain properties. It has a {Describer}, which can have properties, subclasses, and instances. In the current investigation, for example, {Curator} is a type of {Describer} and "Gillbert Pietrzak," the current curator for the Pittsburgh Photographic Library, is an instance of the class {Curator}. The entity {ImageDescr} has two roles: 1) It describes (descr) objects and states of affairs, and 2) It describes events. Furthermore, an image description has subparts: {FactualAssertion} is *part_of* a {SemanticUnit} and {SemanticUnit} is *part_of* {ImageDescr}. These classes can have properties, subclasses, and instances. For example, a {SemanticUnit} has a {SemanticUnitNumber}, which coincides with the coding system applied in the Corpus of Relationships. A {FactualAssertion} has an {arg1}, {arg2}, and in some instances an {arg3}, as well as a relational term {Predicate}. Many of these classes, subclasses, properties, and instances are illustrated in Figure 29.

Given this level of detail, it is possible to show where an instance of a relationship within an image description links to the concept {relation} in an ontology of relationships. It occurs between an instance of {Predicate} and a subtype of {relation} as illustrated in Figure 29.

The ontological view illustrated in Figure 29 can be explained in terms of regions or domains: On the left-hand side of the diagram is the domain of relationship types that emanate as subclasses of the concept {relation}. On the right-hand side of the diagram is the domain of image descriptions where instances of relationships (predicates) reside. In this domain, instances emanate from relationships expressed in factual assertions.

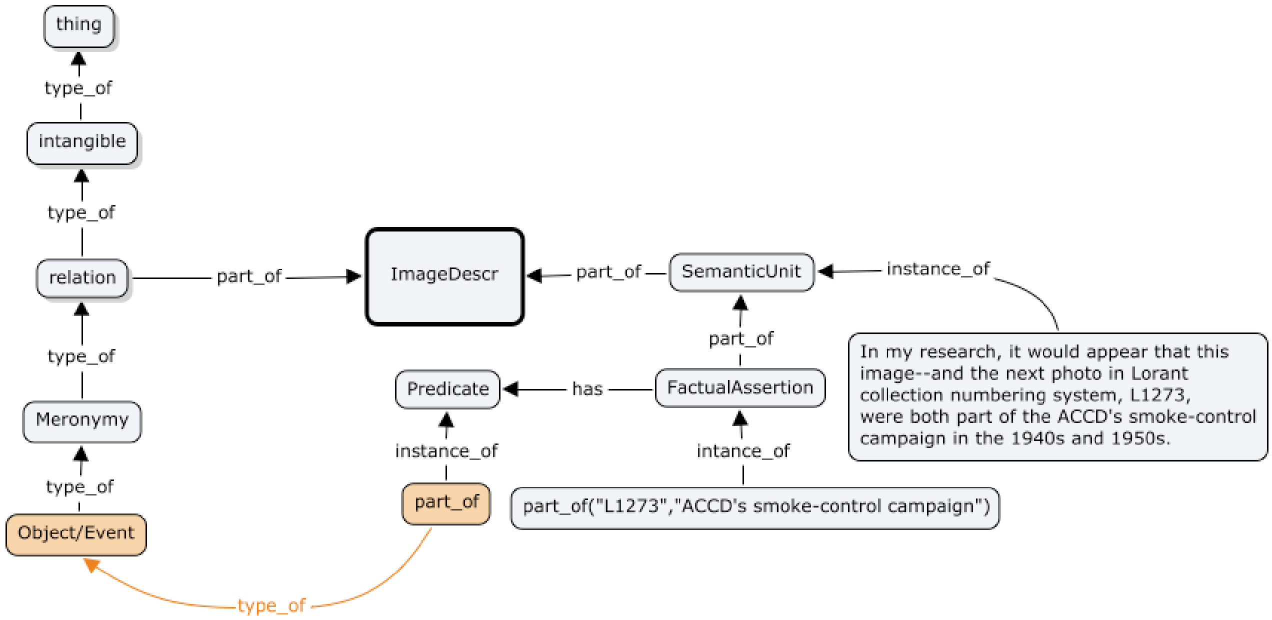


Figure 29: Link between relation instance and relation type

The analysis now moves to a final, more detailed view of the OIDR, which is illustrated in Figure 30. The illustration in Figure 30 is a small part of a larger situation or process called {photography}. In the larger situation {photography}, there are an infinite number of events and states of affairs. The events and states described most often in the literature begin with the {photograph} concept and the notion of a scene being photographed, followed by a photographer who captures the scene to film. The latent image that lies on the film's surface is chemically processed and then a print of the resulting negative is made on photographic paper. The photographic print is a perceptible object and a viewer can bring meaning to the image by describing it. It is the outcome of this last activity—describing—that is partially explained in Figure 29. It is a partial explanation because it attempts to represent only the types of description entities captured in the Corpus.

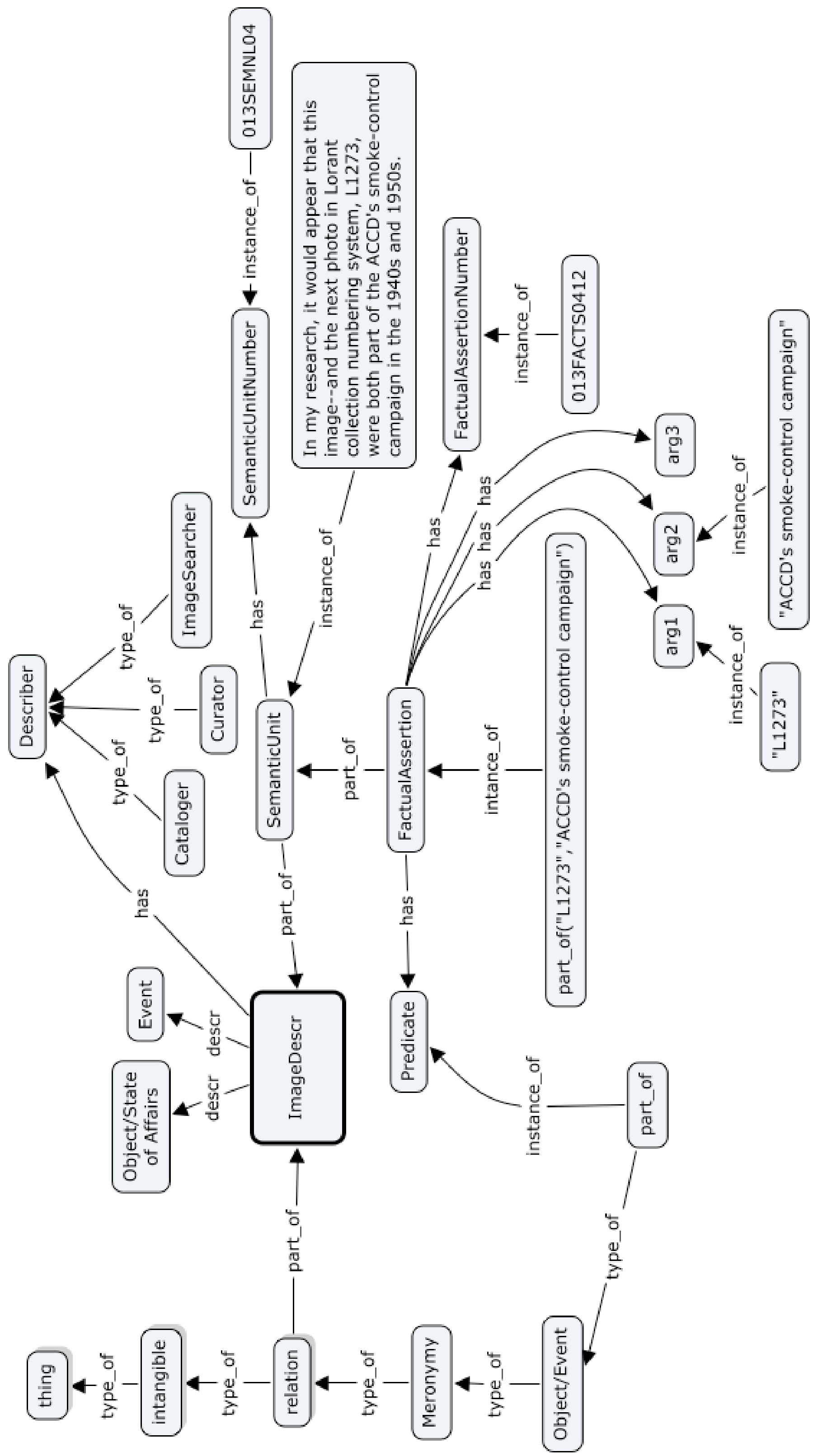


Figure 30: Diagrammatic view of an instance of image description

A better understanding of what is meant by *relational inference* can be achieved by considering the factors shaping the representation shown in Figure 30. This representation makes explicit a basic set of facts that when implemented in a knowledge base such as Scone, the inference mechanism of the knowledge base can infer additional facts. For example, if {FactualAssertion} is part of {SemanticUnit} and {SemanticUnit} is part of {ImageDescr}, then {FactualAssertion} is part of {ImageDescr}. This demonstrates property of relations called *transitivity*. Transitivity also applies to the classification relation *type_of* that links the relation *part_of* to the concept {Object/Event}. In other words, if *part_of* is a type of {Object-Event} relationship and {Object-Event} is a type of {Meronymy}, then *part_of* is a type of meronymic relationship. The actual instance of the *part_of* relationship resides in an image description in the Corpus labeled 013FACTS0412. It is this last detail—linking to an instance in the Corpus—that, in part, establishes empirical grounding for the claims made in this study.

To summarize, the goal in this section has been to engage ontology's role in structuring and representing information relating to image description relationships. A series of three views were presented at varying levels of detail. First, in Figure 28, a general overview showed how the concept {relation} in Scone's upper-level ontology links through a *part_of* relationship to the concept {ImageDescr} (image description). In a more detailed view, Figure 29, an instance of a relationship taken from the Corpus is linked to the ontology of relationships, which enables the instance to inherit all of the properties of its superclasses (including {relation}, {intangible}, and {thing}). In the broadest overview, Figure 30, a framework of classes, subclasses, properties, and instances are interlinked through a network of relationships. This larger picture demonstrates how ontology at the symbol level helps structure and represent an instance of an image description occurring in the Corpus.

Finally, the suggestion was made that the ontology of image description relationships (OIDR) be viewed as two domains or regions. Looking at the illustration in Figure 28, the region on the right represents an ontological view of the concept {ImageDescr}. The region on the left represents a taxonomy of relationship types. It is the region on the left where the *families* of relationships described in the next section are organized as subclasses of the concept {relation}.

5.5 FAMILIES OF RELATIONS

Categorizing relationships is not an easy task, and as it was pointed out in the literature review, Library and Information Science researchers disagree on the number and types of relationships that exist. Nevertheless, the approach taken in this investigation has been to rely on real-world samples of relationships expressed during image searching incidents in the Pittsburgh Photographic Library and that this would provide the starting point for creating an empirically grounded ontology of image description relationships. The relationship types and subtypes used for organizing the 1,655 relationship instances recorded in the Corpus of Relationships serve as the primary basis for the top-levels of the ontology of relationships. The next and final step is to integrate into a coherent whole the relationship types and subtypes that emerged in the Corpus of Relationships and identify the properties of family members.

In the Corpus of Relationships, relationship instances were classified as being a particular type of relationship, but many phenomena were left unaccounted for, and although instances were analyzed and determined to conform with certain ontological rules, it remains unclear whether each instance in a class corresponds to a single more general one, or that instances are more complex than previously thought and necessitate forming subclasses. Nevertheless, the research efforts carried out in the Corpus of Relationships provide a starting point for creating an ontology of relationship types. The relationship categories and their subcategories adopted in the Corpus and the subcategories that emerge during the analysis carried out in Chapter four provide the primary basis for the domain-specific taxonomy of relationships that describe visual information. The ontological analysis begins with the attribution relationship.

5.5.1 Attribution relationship

Attribution relationships represent the largest family of relationships to emerge from the data sample. The analysis carried out in Chapter four showed that there were 643 occurrences of attribution relationships expressed in the data sample. The data in Table 22 shows that these 643 instances can be grouped into 100 different types of relationships. It is interesting to note that 80% of these 100 categories occurred in the data sample less than four times each. In other words, only 20 relationship categories account for 43% of all the attribution relationship instances. If frequency of occurrence were the only criteria for determining relevance in this study, the top 20 attribution relationships listed in Table 21 would be significant.

Table 21:The 20 most frequently occurring attribution relationship types

Rank	Property-types	Frequency count			Total Occurrences
		Cataloger	Curator	Image searcher	
1	subject_of	69	28	172	269
2	accession_code_of	92	2	4	98
3	capture date_of	23	0	0	23
4	role_of	4	2	14	20
5	pix_of	0	0	17	17
6	occupation_of	5	0	9	14
7	format_of	2	0	11	13
8	relative_of	10	0	0	10
9	feature_of	0	4	5	9
10	photographer_of	8	0	1	9
11	depiction_of	0	0	7	7
12	age_of	5	0	1	6
13	genre_of	3	1	1	5
14	mother_of	1	0	4	5
15	view_of	2	0	3	5
16	birth_date_of	0	0	4	4
17	color_of	2	0	2	4
18	death_date_of	0	0	4	4
19	photograph_of	1	1	2	4
20	theme_of	0	1	3	4

A closer look at these 20 categories reveals that some of the categories share certain properties in common. For example, numbers 3, 16, and 18 express {time points}; numbers 9, 11, 19, and 20, *feature_of, depiction_of, photograph_of,* and *theme_of* respectively, strongly suggest a link to the highest ranked category *subject_of*; and *relative_of* and *mother_of* types of kinship relationships. This evidence suggests the possibility of grouping the 100 attribution categories into broader families of relations. The goal in this section, therefore, is to test this hypothesis and examine the data in Table 22 to determine what important elements the relationship types share in common.

Table 22: Classes of attribution relationships

Count	Property-types	Frequency count Cataloger	Curator	Image searcher	Total	Percent of Sample
1	about	0	1	1	2	0.31%
2	accession_code_of	92	2	4	98	15.24%
3	address_of	1	0	0	1	0.16%
4	age_of	5	0	1	6	0.93%
5	alphanumeric_code_of	0	0	1	1	0.16%
6	arrangement_of	0	0	1	1	0.16%
7	aspects_of	0	0	1	1	0.16%
8	bedroom_of	1	0	0	1	0.16%
9	birth_date_of	0	0	4	4	0.62%
10	brother_of	2	0	1	3	0.47%
11	burning_of	0	0	1	1	0.16%
12	called	0	0	1	1	0.16%
13	caption_of	1	0	1	2	0.31%
14	capture_date_of	23	0	0	23	3.58%
15	child_of	1	0	0	1	0.16%
16	photographic_process_of	2	0	2	4	0.62%
17	connected_to	0	1	0	1	0.16%
18	content_of	1	0	0	1	0.16%
19	convention_of	1	0	0	1	0.16%
20	date_of	0	0	5	5	0.78%
21	daughter_of	0	1	1	2	0.31%
22	death_date_of	0	0	4	4	0.62%
23	depiction_of	0	0	7	7	1.09%
24	elements_of	0	0	1	1	0.16%
25	father_of	0	1	1	2	0.31%
26	feature_of	0	4	5	9	1.40%
27	first_cousin_of	1	0	0	1	0.16%
28	format_of	2	0	11	13	2.02%
29	founder_of	0	0	1	1	0.16%
30	genre_of	3	1	1	5	0.78%
31	giant_of	0	0	1	1	0.16%
32	growth_of	0	0	2	2	0.31%
33	has_caption	1	0	0	1	0.16%
34	has_estate	1	0	0	1	0.16%
35	has_first_cousin	1	0	0	1	0.16%

Count	Property-types	Frequency count			Total	Percent of Sample
		Cataloger	Curator	Image searcher		
36	has_IntervalInTime	0	0	1	1	0.16%
37	has_occupation	1	0	0	1	0.16%
38	has_origin	0	1	0	1	0.16%
39	has_referent	0	0	2	2	0.31%
40	has_role	1	0	1	2	0.31%
41	has_skill	0	0	1	1	0.16%
42	has_superintendent	0	0	1	1	0.16%
43	has_top	0	0	1	1	0.16%
44	has_view	1	0	0	1	0.16%
45	heading_of	1	0	0	1	0.16%
46	height_of	0	0	1	1	0.16%
47	home_of	0	0	1	1	0.16%
48	imbued_with	1	0	0	1	0.16%
49	indexical_of	0	0	1	1	0.16%
50	interest_of	0	0	1	1	0.16%
51	iron_forge_of	1	0	0	1	0.16%
52	life_and_times_of	0	0	1	1	0.16%
53	life_of	0	0	1	1	0.16%
54	location_of	0	0	3	3	0.47%
55	married_to	1	0	0	1	0.16%
56	Memorial_Stone_of	1	0	0	1	0.16%
57	mother_of	1	0	4	5	0.78%
58	mural_of	0	0	1	1	0.16%
59	name_of	1	0	2	3	0.47%
60	nearest_relative_of	1	0	0	1	0.16%
61	occupation_of	5	0	9	14	2.18%
62	of	1	1	0	2	0.31%
63	on	0	0	3	3	0.47%
64	only_child_of	1	0	0	1	0.16%
65	opening_of	0	0	1	1	0.16%
66	owner_of	1	0	0	1	0.16%
67	paintings_of	0	0	1	1	0.16%
68	photo_of	0	0	1	1	0.16%

Count	Property-types	Frequency count			Total	Percent of Sample
		Cataloger	Curator	Image searcher		
69	photograph_of	1	1	2	4	0.62%
70	photographer_of	8	0	1	9	1.40%
71	pix_of	0	0	17	17	2.64%
72	point_of_view	1	0	1	2	0.31%
73	portrait_of	1	1	1	3	0.47%
74	position_of	0	0	2	2	0.31%
75	president_of	1	0	0	1	0.16%
76	publisher_of	2	0	0	2	0.31%
77	purchasing_agent_of	2	0	0	2	0.31%
78	quality_of	0	0	1	1	0.16%
79	quantity_of	0	0	3	3	0.47%
80	relating_to	0	2	0	2	0.31%
81	relative_of	10	0	0	10	1.56%
82	replacement_of	0	0	1	1	0.16%
83	representation_of	0	0	2	2	0.31%
84	role_of	4	2	14	20	3.11%
85	scene_of	0	0	2	2	0.31%
86	scenes_of	0	0	1	1	0.16%
87	scope_of	0	0	1	1	0.16%
88	shows	0	0	1	1	0.16%
89	side_number_of	0	0	1	1	0.16%
90	sisters_of	0	0	1	1	0.16%
91	size_of	0	0	1	1	0.16%
92	son_of	1	0	0	1	0.16%
93	subject_of	69	28	172	269	41.84%
94	taken_by	0	0	1	1	0.16%
95	taken_on_board	2	0	0	2	0.31%
96	theme_of	0	1	3	4	0.62%
97	uncle_of	2	0	0	2	0.31%
98	view_of	2	0	3	5	0.78%
99	wedding_of	0	0	1	1	0.16%
100	wife_of	0	0	1	1	0.16%
TOTAL		264	48	331	643	100.00%

The features of the relationship types shown in Table 22 were examined to determine different ontological categories. This resulted in the taxonomy of attribution relations presented in Table 23.[19]

Table 23: Preliminary taxonomy of nine families of attribution relations

1. **accession_code_of**
2. **age_of**
3. **photographic_process_of**
4. **date_of**
   a. birth_date_of
   b. capture_date_of
   c. death_date_of
4. **format_of**
5. **genre_of**
   a. portrait_of
6. **occupation_of**
7. **role_of**
   a. photographer_of
8. **subject_of**
   a. about
   b. depiction_of
   c. feature_of
   d. on
   e. photo_of
   f. photograph_of
   g. pix_of
   h. relating_to
   i. representation_of
   j. scene_of
   k. theme_of
   l. view_of

---

[19] Table 23 shows the category *date_of* subdivided by time point distinctions. The *birth_date_of* relationship (or read in the opposite direction *has_birth_date*) is an attribute and within this relationship there is hidden an existential quantification: "Every person has exactly one birth date." Therefore, the construction **(?X{animal} has_birth_date ?Y{date})** is an attribute and also a temporal relation since the restriction on ?Y is a kind of time point or time interval. This explains why the *date_of* relationships in Table 23 are also included in the temporal ontology illustrated in Figure 36.

The quantity and variety of attribution relations suggests that no taxonomy of attribution relations can be exhaustive in the domain of image descriptions and that a thorough analysis of the attribution relations in Table 22, or for that matter, any of the other categories of relationship types in the Corpus, would extend this dissertation beyond the scope of its original intent. To assist in reigning in what might be an unlimited number of categories of relationship subtypes, only the relations that occur with the most frequency—in the case of attribution relations, only those listed in Table 21—will be used to cluster and identify families of relationship types. The analysis now proceeds with defining and describing the nine families of attribution relations listed in Table 23. The *kinship* relationship is tentatively removed from its classification as a subtype of attribution and is treated separately in its own class below.

An *attribute* is a property or characteristic of something. At the most abstract level of meaning, an attribution relationship links one entity to another entity and has the form **(?X{thing} has an attribute ?Y{thing})**, where {thing} includes physical objects as well as mental objects. For example: *The pencil is sharp.* However, in this investigation the first and second concepts {thing} are more specialized expressing specific types of attributes. For example, the subject of a photograph is an expression of the form **(?X{thing} subject_of ?Y{photograph})**, the name of a photographer who takes a picture is an expression of the form **(?X{person} photographer_of ?Y{photograph})**, and the occupation of someone depicted in an image is an expression of the form **(?X{thing} occupation_of ?Y{person})**.

Attribution relations are characterized as being asymmetric. That is, for all ?x and all ?y, if ?x is an attribute of ?y, then it is not the case that ?y is an attribute of ?x.

The first attribution relationship examined is *accession_code_of*.

5.5.1.1 Accession code

An *accession code* is an alphanumeric code—a sequential number preceded by a letter—assigned to each unique photographic negative and print as it is entered into the collection. The accession code attribute links accession codes with photographs. All 98 instances of *accession_code_of* relationships follow the form *accession_code_of("L1273",photograph)*, where arg1 is a named entity—a unique alphanumeric code—and arg2 is the concept {photograph}. This can be represented as a **(?X{string}** ***accession_code_of*** **?Y{photograph})** construction.

5.5.1.2 Age

Age refers to how long something has existed. The age attribute links a named individual to a number representing the individual's age. The age_of relationship follows the form *age_of("Andrew Carnegie","16")*, where arg1 is a named entity and arg2 is a numeric quantity. This can be generalized as a **(?X{person}** ***age_of*** **?Y{quantity})** construction. There is one instance that does not fit this formula. In reference incident 12, an image searcher writes

> *I am told that the hotel was located adjacent to the Mt. Washington Streetcar Tunnel and near the old P & L E Railroad Depot.*

This is parsed into several factual assertions, including *age_of("P & L E Railroad Depot",old)*, where arg1 is a named entity and arg2 is a quality. The quality {old} applied to a building carries a slightly different interpretation of age—that the building is not new, that it has been around a long time—where 'old' applied to a person means having lived a long time. This suggests there may be a useful ontological distinction drawn between age_of relationships integrating animate and inanimate entities.

5.5.1.3 Photographic Process

A photographic process is a type of process (noun). A photographic process relationship is a type of attribution relationship. For example, a photographic print may be described as having been made using a gelatin silver process, or that a photograph has the properties or characteristics of a gelatin silver print. While the process itself is a complex event which lasts some amount of time with a beginning and ending and different stages of time where the process is going on, here the only the result of the event--the photograph embodiment--is of concern. There are several possible values, but in the Corpus all but one of the references to process were limited to color negatives, color photographic prints, and black and white photographs. The *photographic_process_*of relationship follows the form *photographic_process_of('black and white', photograph)* or *photographic_process_of(color,photograph)* where arg1 is the concept {photograph} and arg2 designates the photographic process: {color} or {black and white}. This can be represented as a **(?X{photograph}** ***photographic_process_of*** **?Y{photographic process})** construction.

There is one exception in reference incident 037 where a cataloger describes the color of a pre-photographic referent—an object that existed in the scene when the photograph was taken:

Ohio River tow. Whitish material in three of the barges is sulphur.

This suggests that an ontological distinction should be made between the concept {color} describing depicted referents, pre-photographic referents, picture elements, and photographic processes. The layers of meaning outlined in Ch. 1, § 1.4 would accommodate all of these distinctions with the possible exception of {color process} if it is described as an event taking place over time.

5.5.1.4 Date

In the Corpus of Relationships, *date* is a specified day, month, year, decade, century, or span of time. This ambiguity must be removed and the relationship type refined when moving to a formal ontology representation in a computational environment. In the majority of instances, the date formula links dates to referents—that is, objects that existed at the time the photograph was taken. The date_of relationship follows the form *date_of("1936","Liberty Avenue Flood")* where arg1 is a date and arg2 is an event. This can be represented as a **(?X{date} *date_of* ?Y{event})** construction.

There are three subtypes of the date_of relationship that emerge from the Corpus: *birth_date_of*, *death_date_of,* which refer to referents—people depicted in photographs—and *capture_date_of*. The *capture_date_of* relationship occurs most often and follows the form *capture_date_of("1862","L-201")* where arg1 is a date and arg2 is a named photograph (designated by an accession code). This can be represented as a **(?X{date} *capture_date_of* ?Y{photograph})** construction.

The *birth_date_of* and *death_date_of* relationships follow the form *birth_date_of("March 1, 1807","James Laughlin")* and *death_date_of_("1825","Joseph McClurg")* where arg1 is a date and arg2 is a named individual. These can be represented as **(?X{date} *birth_date_of* ?Y{person})** and **(?X{date} *death_date_of* ?Y{person})** constructions.

5.5.1.5 Format

The term *format* suffers from semantic overload both in the literature and in this investigation. In the literature, format describes film size and the associated cameras and equipment for shooting a particular film size. For instance, *large format film* is generally 10x13 centimeters or larger. Format refers to image file formats in digital photography. For example, a digital image is described as being JPEG format, PNG format, or TIFF format. In the data sample analyzed in this study, the relationship format_of was used generally to describe a property of the physical manifestation of a photograph, as in *format_of('still image',photograph)* where {still image} is a single, static photographic image as distinguished from a moving picture as in *format_of('moving picture,'visual image').* In one instance, the *format_of* relationship was used in the Corpus to described a photograph that was taken of a drawing as in *format_of(drawing,photograph).*

This research is targeted at providing a semi-formal ontology of conceptually driven relationship types. Providing this depends on clarifying the nature and the formal properties of the relationship. The data sample and subsequent content analysis does not provide the level of clarity necessary for defining or characterizing what kind of entity a *format_of* relationship is or the kinds of concepts it relates. Deeper analysis of the ontological status of format is needed.

5.5.1.6 Genre

Genre is an expressive style in photography; a kind of endeavor or work with a characteristic form and style of depiction. For example, in the landscape genre, photographs depict natural scenery. The genre attribute links genres with photographs. The *genre_of* relationship follows the form *genre_of(landscape,photograph),* where arg1 is a type of genre and arg2 is the concept {photograph} or a named entity. For example, the photograph with accession code A-159 as in *genre_of(documentary,"A-159").* This can be represented as a **(?X{genre}** ***genre_of*** **?Y{photograph})** construction.

The genre "portrait" occurred frequently in the data sample and is examined closely for its inclusion as a sub-type of the genre_of relationship. In reference incident 045, *portrait_of(photograph,"Carnegie")* provides a striking case of conceptual integration, making as one the depicted referent and the object showing depiction, in this instance *a portrait of Andrew Carnegie*. About this factual assertion the following can be said:

1. 'of' serves as a preposition expressing composition.
2. The collocation portrait_of indicates a function
3. The term in position arg1 refers to the object showing depiction of a visual image, in this case a photograph image.
4. The {photograph} concept is implicit in the image searcher's query, "Do you hold the portrait of Carnegie…?" It is essential in arg1 to make explicit the nature of the object showing depiction, either as photograph in general or a named photograph specifically.
5. The term 'Carnegie' in position arg2 refers to another object, the subject of the portrait, the prehended entity of the proposition.
6. The function/predicate portrait_of indicates the material of which the photograph is composed. In this instance the photograph is composed of a mental construct 'portrait' with value Carnegie. They are integrated one within the other.

7. The relational term determines the kind of entity that can fill the arg2 slot. A portrait_of relationship presents the prehended entity (arg2) as a person, which is an abstract entity as a depicted referent.

5.5.1.7 Occupation

An *occupation* is an activity, line of work, or job carried out by a referent depicted in a photographic image. The occupation attribute links types of occupations to persons depicted in photographs. The occupation_of relationship follows two forms: 1) *occupation_of("John Potter",(of('general superintendent',"Homestead Mill")))* where arg1 is a named person and arg2 is a sub-proposition describing the person's occupation in the arg1 position and a named organization in the arg2 position, joined together by the preposition 'of,' which expresses the individual's belonging to the organization through their occupation title or position, and 2) *occupation_of("Michelle Madoff",'city councilwoman')*, where arg1 is a named person and arg2 is a type of occupation. the concept {photograph}. These can be represented as **(?X{person} *occupation_of* (?Y{occupation} *of* ?Z{organization}))** and **(?X{person} *occupation_of* ?Y{occupation})** constructions respectively.

5.5.1.8 Role

A *role* is an action or activity assigned to or expected of a animate or inanimate object. The role attribute links a role to an entity depicted in a photograph, the photograph itself, or to an entity external to the image or its physical carrier. This ambiguity must be resolved before the relationship type can be specified in a precise, formal representation language. One role_of relationships expressed in the Corpus follows the form *role_of('tell story','visual record')* where arg1 is an activity and arg2 is a physical object (photograph).

A more complex variant of the role_of relationship involves a sub-proposition describing an entity's role in an activity: *role_of('main player',(in("Amalgamated Association of Iron and Steel Workers",strike)))*. The role is expressed in arg1and arg2 is a sub-proposition naming an organization in the arg1 position and an activity in the arg2 position, joined together by the preposition 'in,' which denotes a state—that is, the organization is in a state of being on strike. This can be represented as a **(?X{role}** ***role_of*** **(?Y{organization}** ***in*** **?Z{activity})).**

5.5.1.9 Subject

Subject is the last attribution type relationship discussed in this investigation. *Subject* is the topic a viewer attaches to the meaning of a photographic image. In terms of frequency of occurrence, subject is the chief currency in talking about the image description domain. While curators, catalogers, and image searchers bring vast amounts of background knowledge to their describing activities, they make explicit the subjects that are salient to their activities.

The subject_of relationship links subjects to photographic images and follows the form *subject_of("Johnstown Flood",photograph)* or *subject_of("George Westinghouse","A417")*, where arg1 is anything that a viewer interprets to be the topic of the photograph. In these examples arg1 expresses instances of events and persons—named entities—and arg2 is the concept {photograph}, which can be a named individual or the generic concept {photograph}. The subject attribution relationships can be represented as a **(?X{thing}** ***subject_of*** **?Y{photograph})** construction.

The idea that a photograph has a subject seems very basic. It is part of the natural process of description and the most prevalent attributive relation with the most subtypes reported in this study. The subtypes expressed in the Corpus are listed in Table 24.

Table 24: Types of attribution relationships

About
Depiction_of
feature_of
On
photo_of
photograph_of
pix_of
relating_to
representation_of
scene_of
theme_of
view_of

Subjects are expressed in a variety of ways in the relationship corpus. Image searchers express it as (terms expressing subject are italicized):

- [011SEMNL03] "*Images depicting* the growth of the steel mill—both Bessemer and open hearth process."
- [011SEMNL04] "*Documentation of* houses in town; workers; family life—any images to give us a sense of life in the mill town."
- [011SEMNL05] "*Depictions of* both skilled and unskilled laborers at their jobs in the mill."

Curators express subject as:

- [018SEMNL02] "Now, I do have a *photo of* the Grant Building which is located in Downtown Pittsburgh and there was also a Bell Telephone building, but they were not located near each other."
- [039SEMNL010 "Here is the *photograph of* George Westinghouse from thePittsburgh Photographic Library Collection."

And finally, catalogers express subject like this:

- [041SEMNL21] "*In this* 1862 *photograph*, the twenty-seven year old Andrew Carnegie (on the left) poses with George E. Lauder and Thomas N. Miller. Lauder was his uncle; it was he who imbued the young man with the love of Poetry. Miller, the purchasing agent of the Ohio and Pennsylvania Railroad and a partner in the Pittsburgh iron forge of the Kloman brothers, was a key figure in the foundation of Carnegie's fortune. L-201"
- [042SEMNL02] "*Heading:* Pittsburgh, Districts. South Side. #L356."

It turns out that from the variety of expressions representing or marking a subject in a description, the notion of subject requires explanation and when explained it can become quite complex. Most importantly, the types of attribution listed in Table 24 are not treated as types of conceptual relationships in this study. They are treated in the same manner as synonyms and near-synonyms, lexical information that may be useful in natural language processing.

This concludes the ontological analysis of attribution relationship types. A summary of categories, semantic category constructions, and sample instances are presented in Table 25.

Table 25: Summary of attribution relationship types and subtypes.

Types	Constructions	Instances
accession_code_of	**(?X{string}** ***accession_code_of*** **?Y{photograph})**	*accession_code_of("L1273",photograph)*
age_of	**(?X{person}** ***age_of*** **?Y{number})**	*age_of("Andrew Carnegie","16")*
photographic_process_of	**(?X{photograph}** ***photographic_process_of*** **?Y{photographic process})**	*photographic_process_of(color,photograph)*
date_of	**(?X{date}** ***date_of*** **?Y{event})**	*date_of("1936","Liberty Avenue Flood")*
birth_date_of	**(?X{date}** ***birth_date_of*** **?Y{person})**	*birth_date_of("March 1, 1807","James Laughlin")*
capture_date_of	**(?X{date}** ***capture_date_of*** **?Y{photograph})**	*capture_date_of("1862","L-201")*
Death_date_of	**(?X{date}** ***death_date_of*** **?Y{person})**	*death_date_of_("1825","Joseph McClurg")*
format_of	**(?X{thing}** ***format_of*** **?Y{thing}).**	*format_of('still image',photograph)* *format_of('moving picture,'visual image')* *format_of(drawing,photograph)*
genre_of	**(?X{genre}** ***genre_of*** **?Y{photograph})**	*genre_of(landscape,photograph)* *genre_of(documentary,"A-159")*
portrait_of	**(?X{photograph}** ***portrait_of*** **?Y{person})**	*portrait_of(photograph,"Carnegie")*
occupation_of	**(?X{person}** ***occupation_of*** **(?Y{occupation}** ***of*** **?Z{organization}))** **(?X{person}** ***occupation_of*** **?Y{occupation})**	*occupation_of("John Potter",(of('general superintendent',"Homestead Mill")))* *occupation_of("Michelle Madoff",'city councilwoman')*
role_of	**X{role}** ***role_of*** **?Y{physical object})** **(?X{role}** ***role_of*** **(?Y{organization}** ***in*** **?Z{activity}))**	*role_of('tell story','visual record')* *role_of('center of activity',"Music Hall")* *role_of('main player',(in("Amalgamated Association of Iron and Steel Workers",strike)))*
photographer_of	**(?X{person}** ***photographer_of*** **?Y{photograph})** **(?X{photograph}** ***taken_by*** **?Y{person})**	*photographer_of("Harold Corsini","3087")* *taken_by(image,"Harold Corsini")*
Subject_of	**(?X{thing}** ***subject_of*** **?Y{photograph})**	*subject_of("Johnstown Flood",photograph)* *subject_of("George Westinghouse","A417"),*

Figure 31 presents a diagrammatic view of the ontology of attribution relationships. To avoid visual clutter, the links between nodes are not labeled in the remainder of the chapter. It can be assumed that all of the links represent class inclusion relationships moving from specific (the lowest nodes in the diagram) upward to more general concepts.

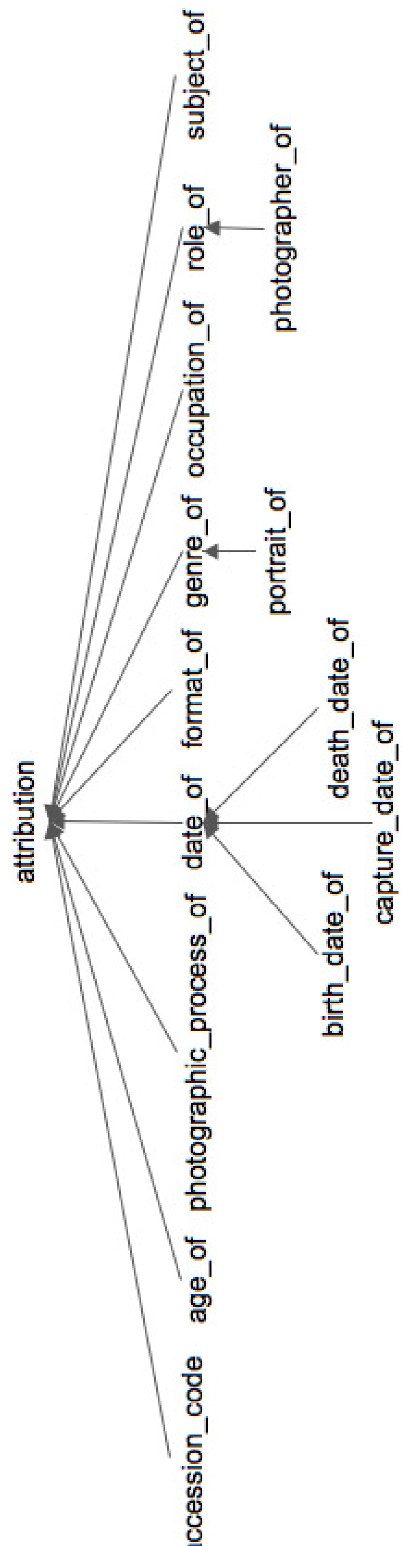


Figure 31: Ontology of attribution relationships

5.5.2 Case relations

Case relations describe parts of real-world situations and the participants taking part in those situations. Typically, the participants include agents that initiate actions, instruments or objects, and results or recipients of the action. These can be extended to include different types of initiators, objects, and recipients of actions. For example, in the event “Dorothea Lange took a picture of Florence Owens Thompson using a medium format press camera,” the *take* event has an agent Dorothea Lange, the instrument is a medium format press camera, and the recipient of the action is Florence Owens Thompson. This could be extended to include, for example, a PointInTime-Action case relation that makes explicit the *date* the photograph was *captured*. Indeed, the analysis in chapter four showed that several other types of case relations emerged in the domain of image descriptions that did not fall under one of Story’s categories. These are listed by type (76 in all) with frequency of occurrence in Table 26.

Table 26: Frequency of case relationship types not defined by Story (1993)

	Case-type	Frequency count				
Count		Cataloger	Curator	Image searcher	Total	Percent of Sample
1	action_between-action	0	0	1	1	0.3%
2	action-agent	3	1	9	13	4.4%
3	action-effector	0	0	1	1	0.3%
4	action-referent	0	0	1	1	0.3%
5	action-location	1	0	0	1	0.3%
6	action-medium	0	0	7	7	2.4%
7	action-object	1	5	2	8	2.7%
8	action-patient	1	0	0	1	0.3%
9	action-PointInTime	1	0	0	1	0.3%
10	action-theme	2	0	3	5	1.7%
11	action-viewpoint	0	0	1	1	0.3%
12	agent-event	0	1	0	1	0.3%
13	beneficiary-action	1	1	1	3	1.0%
14	cause-action	0	0	1	1	0.3%
15	coparticipant-action	1	0	0	1	0.3%
16	creator-action	2	0	0	2	0.7%
17	description-action	0	3	2	5	1.7%
18	direction-viewpoint	1	0	0	1	0.3%
19	down-action	1	0	0	1	0.3%
20	effector-action	3	3	7	13	4.4%
21	era-action	0	0	2	2	0.7%
22	era-viewpoint	0	0	2	2	0.7%
23	experiencer-action	1	0	0	1	0.3%
24	feature_of-action	0	1	0	1	0.3%
25	genre-action	1	0	1	2	0.7%
26	goal-action	1	0	5	6	2.0%
27	instrument-action	2	2	0	4	1.3%
28	instrument-viewpoint	0	1	0	1	0.3%
29	IntervalInTime-action	2	1	6	9	3.0%
30	IntervalInTime-viewpoint	0	0	1	1	0.3%
31	issue-action	1	0	0	1	0.3%
32	killer-action	0	0	1	1	0.3%
33	located_entity-action	1	1	1	3	1.0%
34	located_entity-viewpoint	0	0	1	1	0.3%

	Case-type	Frequency count				
Count		Cataloger	Curator	Image searcher	Total	Percent of Sample
35	location-action	4	4	18	26	8.8%
36	location-viewpoint	0	1	3	4	1.3%
37	manner-action	5	2	1	8	2.4%
38	medium-action	0	0	7	7	2.4%
39	medium-statement	1	0	0	1	0.3%
40	message-statement	1	0	0	1	0.3%
41	motivation-action	1	0	0	1	0.3%
42	occasion-action	1	0	1	2	0.7%
43	occasion-statement	2	0	0	2	0.7%
44	path-action	2	0	4	6	2.0%
45	patient-action	1	1	12	14	4.7%
46	place-action	1	0	0	1	0.3%
47	pointInTime-action	5	5	14	24	7.7%
48	pointInTime-agent	0	0	1	1	0.3%
49	pointInTime-statement	1	0	1	2	0.7%
50	pointInTime-viewpoint	0	0	1	1	0.3%
51	position-action	3	0	0	3	1.0%
52	purpose-event	0	2	0	2	0.7%
53	recipient-action	1	2	2	5	1.7%
54	reference_entity-action	0	0	1	1	0.3%
55	reference_entity-viewpoint	0	0	1	1	0.3%
56	referent-action	1	0	1	2	0.3%
57	referent-viewpoint	2	1	1	4	1.3%
58	representation-action	6	0	0	6	1.7%
59	represented-action	1	0	0	1	0.3%
60	result-action	3	1	5	9	3.0%
61	scene-action	0	1	5	6	2.0%
62	side1-action	1	0	0	1	0.3%
63	side2-action	1	0	0	1	0.3%
64	speaker-statement	2	0	0	2	0.7%
65	subject-action	6	2	4	12	4.0%
66	theme-action	5	5	14	24	8.1%
67	thing_depicted-action	0	0	1	1	0.3%
68	topic-action	0	1	1	2	0.7%

	Case-type	Frequency count				
Count		Cataloger	Curator	Image searcher	Total	Percent of Sample
69	topic-statement	1	0	0	1	0.3%
70	towards-action	1	0	0	1	0.3%
71	victim-action	1	0	1	2	0.7%
72	view_of-action	0	1	0	1	0.3%
73	viewOf-viewpoint	0	0	1	1	0.3%
74	viewpoint-action	1	1	4	6	2.0%
75	viewpoint-look	0	1	0	1	0.3%
76	viewpoint-shot	0	1	0	1	0.3%
	TOTAL	88	51	161	300	100.0%
	Percent of total	29%	16.80%	54.20%	100.00%	

Case relations are characterized as being asymmetric. For example, in the relation between agent and action, it can be said that for all ?x and all ?y, if ?x ?y is an agent-action, then it is not the case that ?y ?x is an agent-action; and for all ?x and all ?y, if ?x ?y is an agent-instrument, then it is not the case that ?y ?x is an agent-instrument; and if ?x ?y is an action-recipient, then it is not the case that ?y ?x is an action-recipient, and so on.

The representation used in this section features two systems of expression, both already used earlier: the factual assertion prefix notation applied in the Corpus, as *has_representation(photograph-v,"A-159")* and the construction expressing semantic categories as **(?X{photograph-v} has_representation ?Y{photograph}).** A new approach, however, is needed for understanding description within a situation or circumstantial context. In the current study, the description viewpoint is a photographing event or situation and schemas provide the core structure for representing case relations in this conceptual context. First, case relations and their properties are explained and then the schema concept for explaining a photographing event is described. For a more detailed explanation of Storey's case roles, see Ch. 4, § 4.3.5.2.

The schema concept (schemata or schemas are its plural forms) is not new to Library and Information Science.[20] Schemas are familiar to librarians as headings lists, the most common being the Library of Congress Subject Heading (LCSH) schema. Librarians also use a variety of metadata schemas, for instance, the Dublin Core metadata schema. Schemas, however, are not limited to these entities and can also help explain case relations.

Photography may be conceived as a schematic system constructed as a multi-part process or *situation-template* whose elements include a set of participants (physical and mental, for example, photographers and scenes), relationships integrating concepts into a coherent wholes, an underlying representation language, and conditions that specify how variables are to be filled to obtain instances of relationships.[21] The set of semantically related entities emerging from the relationships in Table 26 and linked to the situation-template introduced here are *take*, *view*, *look*, *depict*, *show*, and *photograph-v*.[22] These account for the majority of case relations in the data sample and are the focus of ontological analysis in this section.

Each entity evokes different sub-frames or schema. The *take* action (as in a photographer takes a picture) places emphasis on the agent, instrument, and agent's motivation, backgrounding representation, medium, and printing; in a *describe* event, the focus is on the representation (the image description) and its interpretation, backgrounding the instrument (the camera). Knowing

---

[20] Schemas are information *structures* used in this study as frameworks around which events in image descriptions are built.
[21] Conceptions similar to situation-templates have developed in other fields. Most notably in AI, Marvin Minsky's *frame* described as a "data structure representing a stereotyped situation," (Minsky, 1975, p. 212) and Fahlman's multiple inheritance hierarchies and default reasoning with exceptions (Fahlman, 1979, 2008a); in psychology, the schemata idea of Bartlett (Bartlett, 1932), and linguist Charles J. Fillmore's (1976) *frame semantics* where words activate frames of semantic knowledge for the speaker/listener.
[22] The term *photograph-v* is used to distinguish the verb form of "photograph," which means to take a photograph of something, from the noun form of "photograph," which refers to a picture produced by photography.

what takes place in an image description, therefore, requires knowing what takes place in a situation. The situation-template presented here is only a partial representation of a situation, or rather of reality. Accordingly, the photograph-v template, for example, may be regarded as giving partial or fragmentary descriptions of the concept of photography or rather of the relational concept photograph-v.

Schematically, an individual situation-template looks like this:

(Template-name
(Schema$_1$-name   Schema$_1$-instance)
(Construction$_1$)
(Schema$_2$-name   Schema$_2$-instance)
(Construction$_2$)
…))

Subsequent sections deal with descriptions of situation-templates and their schemas.

5.5.2.1 Take-template

Take means to make a photograph. For example, the cataloger in incident 043 describes a photograph as follows: "On Monday, July 5, 1909, when this picture was taken, Independence Day was being celebrated with the traditional fireworks, plus ballon[sic] ascensions in Schenley Park and a doubleheader at Forbes Field." The take-template is explained in Table 27.

Table 27: Take-template

Schema Name	Schema Instance/Construction
agent-action	has_agent(take,photographer)
	(?X{take} has_agent ?Y{person}), where the first variable is to be filled by the VP “take” and the second variable is to be filled by a proper name of the individual taking the photograph.
PointInTime-action	has_PointInTime(take,"July 5, 1909")
	(?X{take} has_PointInTime ?Y{time point}), where the first variable is to be filled by the VP “take” and the second variable is to be filled by a {time point} representing day, month, year.
motivation-action	has_motivation(take,"Independence Day$_1$"(event_of(photograph$_1$)))
	(?X{take} has_motivation ?Y{thing}), where the first variable is to be filled by the VP “take” and the second variable is to be filled by a named event or object being photographed. Object can be a physical or mental thing.
subject-action$_1$	has_subject(take,'traditional fireworks'(feature_of("Independence Day"))) has_subject(take,'balloon ascension'(feature_of("Independence Day"))) has_subject(take,"baseball game"(feature_of("Independence Day)))
	(?X{take} has_subject (?Y{feature} ?Z{event})), where the first variable is to be filled by the VP “take,” the second variable is to be filled by a feature of an event depicted in the image, and the third variable is a named event.
subject-action$_2$	has_subject(take,has_feature('baseball game',doubleheader))
	(?X{take} has_subject (?Y{event} has_feature ?Z{thing})), where the first variable is to be filled by the VP “take,” the second variable is to be filled by an event, and the third variable is to be filled by a feature of the event.
location-action	has_location(take,"Schenley Park")
	(?X{take} has_location (?Y{take} has_location ?Y{place})), where the first variable is to be filled by the VP “take” and the second variable is to be filled by a named place.

5.5.2.2 View-template

View in the domain of image descriptions is a concept concerned with a position from which objects are viewed. Context plays a role in determining position, who or what is in a given position, what is viewed, angle of view, and so on. Not all of these aspects are demonstrated in the restricted domain of the data sample. Image descriptions given in the context of fine art photography practices, for instance, are sure to give rise to a rich schematic structure underlying the view-template.

Among photographers, the relationship between the camera and the object being photographed is described in terms of angle of view and distance. There was no evidence in the data set of terms such as bird's-eye view, or high, low, and oblique angles of view. An image searcher in incident 041 instead describes a photograph as follows:

"View of Pittsburgh in the late 19th century, showing steel plants. The idea here is that many cities were identified with specific industries-I.e. Pittsburgh and the steel industry."

This offers a typical example of a description marked with viewing terms, an explicit topic being viewed, and the position from which viewing takes place is implicitly understood to be the camera or photographer. The point here is that the viewer—in this case an image searcher describing an imagined image—uses the word "view" to intentionally direct the curator's attention to a place "Pittsburgh" in order to communicate a desired photographic subject. For this reason, the researcher formalizes the case relation as *has_location(view,"Pittsburgh")* and calls this a location-action schema. The object steel plant is linked to the viewing act with an object-action schema in *has_object(view,'steel plant').*

The link between the concept {steel industry} and the concept {Pittsburgh} is a causally dependent relationship resulting from the image searcher's point of view in the image description and the evidence that lies in an image depicting steel mills in Pittsburgh. The researcher introduces a causal dependence relation to account for the relation between {steel industry} and {Pittsburgh} in the relationship *depends_on_causally('steel industry',"Pittsburgh").*[23] This core

[23] {Steel industry} is causally dependent on {Pittsburgh} if {steel industry} is the result of Pittsburgh's existence. The destruction of Pittsburgh, however, does not cause {steel industry} to cease existing.

structure is represented as a binary relation **D** where **D** is the viewpoint of an image searcher in a an agent-viewpoint schema *has_agent(viewpoint*,D).

The depict-template has an agent-action schema generalized as *has_agent(depict,person)* as well as a medium-action schema *has_medium(depict,photograph).* The concept {view} is linked to the image searcher through an agent-action schema in the generic assertion *has_agent(view,person).*

In the same incident, viewpoint is again marked by a viewing term, but this time the searcher introduces an interval-in-time schema asking, "Are there any views (dating from the late 19th century) of railroad cars pulled up to one of Carnegie's steel plants (loading and unloading) with the name Carnegie written on the railroad cars or visible on the plant?"

It is trivial for the researcher to understand this expression describing a viewpoint (camera position) where a viewer of a photograph (another viewpoint) may see the name "Carnegie" written on railroad cars or a steel plant. To translate this knowledge into a higher order formalism that is machine-understandable, however, is problematic and outside the immediate concerns of this dissertation. The constructions offered throughout this chapter do offer limited insight to the natural language processing problem by labeling the types of semantic categories relationships join together.

Aspects of the view-template as it relates to incident 041SEMNL17 are described in Table 28.

Table 28: View-template for semantic unit 041SEMNL17

Schema Name	Schema Instance/Construction
location-action	has_location(view,"Pittsburgh")
	(?X{view} has_location ?Y{named place}), where the first variable is to be filled by the VP “view” and the second variable is to be filled by a named place.
PointInTime-action	Has_PointInTime(view,'late 19th century')
	(?X{view} has_PointInTime ?Y{time point}), where the first variable is to be filled by the VP “view” and the second variable is to be filled by a time interval.
object-action	has_object(view,’steel plant’)
	(?X{view} has_location ?Y{physical object}), where the first variable is to be filled by the VP “view” and the second variable is to be filled by a physical object.
agent-action	has_agent(view,person$_1$)
	(?X{view} has_agent ?Y{person}), where the first variable is to be filled by the VP “view” and the second variable is to be filled by a named person.
agent-viewpoint	*has_agent(viewpoint,(depends_on_causally(‘steel industry’,”Pittsburgh”)))*
	(?X{viewpoint} has_agent (?Y{steel industry} depends_on ?Z{thing})), where the first variable is to be filled by the VP “view” and the second variable is to be filled by an embedded proposition of the form *depends_on(x,y)*.

Some viewpoints situate objects in the scene, for example, “people behind rope.” Other viewpoints describe a general point in space, for example, the interior or inside of a building. The view-template is examined further in 041SEMNL06 described in Table 29.

Table 29: View-template applied to semantic unit 041SEMNL06.

Schema Name	Schema Instance/Construction
scene-action	has_scene(view,of(interior,'steel plant'))
	(?X{view} has_scene ?Y{building}interior), where the first variable is to be filled by the VP “view” and the second variable is to be filled by a generic or named building and the ADJ interior.
scene-action	has_scene(view,of(exterior,'steel plant'))
	(?X{view} has_scene ?Y{building}exterior), where the first variable

	is to be filled by the VP “view” and the second variable is to be filled by a generic or named building and the ADJ exterior.
IntervalInTime-viewpoint	has_IntervalInTime(viewpoint,"1870's-1880's")
	(?X{view} has_IntervalInTime ?Y{time interval}), where the first variable is to be filled by the VP “view” and the second variable is to be filled by a time interval.

5.5.2.3 Look-template

Look in the current domain of interest means a person depicted in a photograph directing his or her eyes toward some entity or phenomenon. An image searcher in incident 040 describes a photograph as follows: “Dorothy Height speech competition-gave speech at Music Hall. On stage looking out. Late 20's early 30's.” The look-template is explained in Table 30.

Table 30: Look-template

Schema Name	Schema Instance/Construction
entity_perceiver-action	has_entity_perceiver(look,”Dorothy Height”)
	(?X{look} has_entity_perceiver ?Y{person}), where the first variable is to be filled by the VP “look” and the second variable is to be filled by a named person.
reference_object-action	has_reference_object(look,audience)
	(?X{look} has_reference_object ?Y{physical object}), where the first variable is to be filled by the VP “look” and the second variable is to be filled by a physical object.
agent-action	has_agent(look,"Dorothy Height")
	(?X{look} has_agent ?Y{person}), where the first variable is to be filled by the VP “look” and the second variable is to be filled by a named person.
IntervalInTime-action	has_IntervalInTime(look,"late 20's early 30's)
	(?X{look} has_IntervalInTime ?Y{time interval}), where the first variable is to be filled by the VP “look” and the second variable is to be filled by a time interval.
path-action	has_path(look,’looking out’)

	(?X{look} has_path ?Y{spatial path}), where the first variable is to be filled by the VP "look" and the second variable is to be filled by a spatial expression describing a path.

5.5.2.4 Depict-template

Persons describing images communicate a message stating an image characterizes or *depicts* an entity. In incident 041, an image searcher asks, "Are there any beautiful colored engravings of paintings of any of Carnegie's steel plants? As you know, many industrial facilities of the period were depicted in aerial-types of views done in colored engravings."

Table 31: Depict-template

Schema Name	Schema Instance/Construction
medium-action	
	(?X{depict} has_medium ?Y{physical object}), where the first variable is to be filled by the VP "depict" and the second variable is to be filled by a physical carrier of the depiction.
depicted_entity-action	has_depicted_entity(depict,'steel plant')
	(?X{depict} has_depicted_entity ?Y{thing}), where the first variable is to be filled by the VP "depict" and the second variable is to be filled by a depicted referent.
view-action	has_view(depict,aerial)
	(?X{depict} has_view ?Y{viewpoint}), where the first variable is to be filled by the VP "depict" and the second variable is to be filled by a point of view.

5.5.2.5 Show-template

Persons use the term *show* to express how the photograph medium causes an entity or phenomenon to be perceived by a viewer. In the context of semantic unit 041SEMNL16, the image searcher describes a desired image to the curator, a perceiver, by requesting, “Pictures showing Andrew Carnegie in his office or in some other context at one of his plants.” In the Corpus, the relationship expressed between Andrew Carnegie and his office is defined as a spatial expression. Here, however, it is described as an instance of the medium-action schema where the medium depicting a referent *shows* an entity or phenomenon. This illustrates the *common argument criterion* (introduced in § 5.5.6), that the same scene or subject can be explained using different relational expressions.

In incident 020, an image searcher is seeking out a known photograph describing it as:

The photograph shows Amelia Earhart's Avro Avian, side number G-EBUG nose down in a ditch after a forced landing near Pittsburgh in 1928.

The phrase “photograph shows” marks the description as an image description and not a description of the scene at the time the photograph was taken. Table 32 provides examples of schema describing *show*.

Table 32: Show-template

Schema Name	Schema Instance/Construction
medium-action	has_medium(show,photograph)
	(?X{show} has_medium ?Y{physical object}), where the first variable is to be filled by the VP “show” and the second variable is to be filled by a physical carrier of the depiction.
depicted_entity-action	has_depicted_entity(show,”Andrew Carnegie”)
	(?X{show} has_depicted_entity ?Y{person}), where the first variable is to be filled by the VP “show” and the second variable is to be filled by a named person.
location-action	has_location(show,office)
	(?X{show} has_location ?Y{place}), where the first variable is to be filled by the VP “show” and the second variable is to be filled by type of place or

	named place.
Scene-action	has_scene(show,'fiery,dramatic scene')
	(?X{show} has_scene ?Y{scene}), where the first variable is to be filled by the VP "show" and the second variable is to be filled by the description of a scene.
object-action	has_object(show,"Bessamer furnace")
	(?X{show} has_object ?Y{physical object}), where the first variable is to be filled by the VP "show" and the second variable is to be filled by a physical object.

5.5.2.6 Photograph-v-template

Photograph-v describes a photographer who produces or makes a photograph. The photograph serves as a representation of an entity or entities (physical and mental) including states and events, which is the result of the *photograph-v* activity.

Table 33: Photograph-v-template.

Schema Name	Schema Instance/Construction
agent-action	has_creator(photograph-v,"Lewis W. Hine")
	(?X{photograph-v} has_agent ?Y{person}), where the first variable is to be filled by the VP "photograph-v" and the second variable is to be filled by a named person.
genre-action	has_genre(photograph-v,documentary)
	(?X{photograph-v} has_genre ?Y{genre}), where the first variable is to be filled by the VP "photograph-v" and the second variable is to be filled by a type of photographic genre.
represented-action	has_representation(photograph-v,"A-159")
	(?X{photograph-v} has_agent ?Y{photograph}), where the first variable is to be filled by the VP "photograph-v" and the second variable is to be filled by a named photograph.
instrument-action	has_instrument(photograph-v,camera)
	(?X{photograph-v} has_instrument ?Y{camera}), where the first variable is to be filled by the VP "photograph-v" and the second variable is to be filled by a type of photographing device.

manner-action	has_manner(photograph-v,sensitive)
	(?X{photograph-v} has_manner ?Y{person}), where the first variable is to be filled by the VP "photograph-v" and the second variable is to be filled by character of agent's expression.
depicted_object-action	has_depicted_object(photograph-v,"Slavic millworker")
	(?X{photograph-v} has_depicted_object ?Y{physical object}), where the first variable is to be filled by the VP "photograph-v" and the second variable is to be filled by a type of physical object or named entity.

To summarize, the goal was to establish both the empirical grounding of case relations by making reference to semantic units in the Corpus, and describe how case relationships link together in an ontology of relationship types. Situation-templates provide a framework for describing classes of case-based relationships. Relationships between events and participants are partially defined by schematic structures. In some cases, terms expressing an event and participants were explicit in the text and a schema could explain aspects of relationship instances easily. Where the event or participants were implicit, pragmatic inference is needed to schematize the case relationships.[24]

While additional situation-templates might be motivated by exploring all of the case relations in the Corpus, it is reasonable to conclude that a complete set would have to include something like the schema represented in Tables 27-33 to cover case relations in image descriptions adequately.

The ontology resulting from the forgoing analysis is illustrated in Figure 32.

[24] While it is not the goal of this dissertation to address natural language processing concerns, it can be generally stated that the normal link marked by case relations is between a verb and noun group. In one instance, however, the case relation was between an event and an embedded proposition (the agent-viewpoint schema, Table 28).

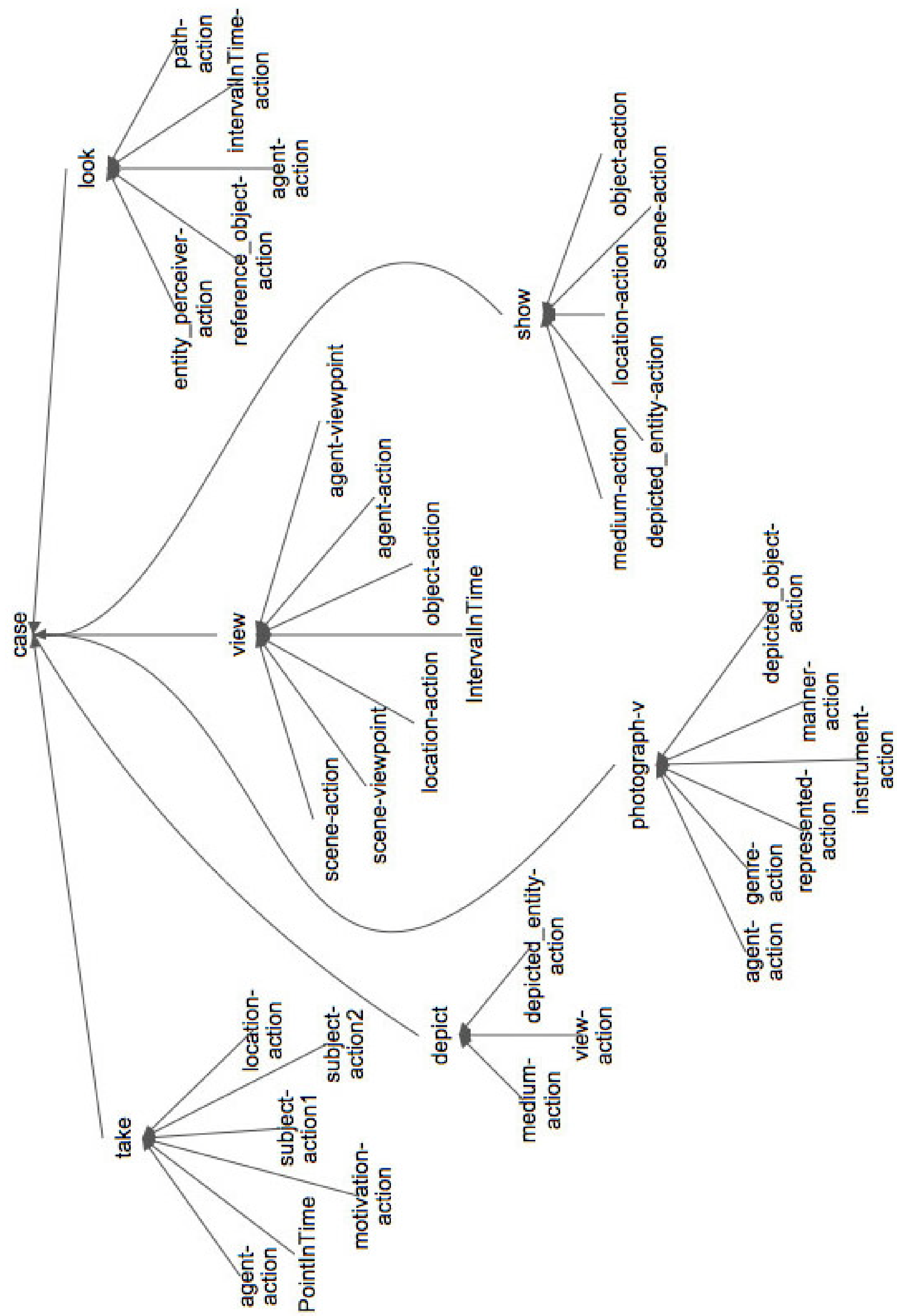


Figure 32: Ontology of case relationships

5.5.3 Inclusion

Storey (1993) defines *inclusion* relationships as “situations where one entity type comprises or contains other entity types” (p. 460). The type of inclusion examined here is class inclusion. Meronymy and spatial relations—other types of inclusion—are treated separately. Class inclusion is the most important relationship in the area of relationship descriptions. It establishes a hierarchical structure of relationships connected to certain inheritance mechanisms where more specialized relationship types inherit information from their superclasses. This topic, and relational inference, was explained earlier in § 5.3.

The data analysis in chapter four demonstrated that class inclusion relationships are not explicitly stated phenomena in image descriptions. The position taken in this study is that the role of class inclusion in an ontology of image description relationship types is to explain standard subtype and supertype relationships that occur between relationship types, or when modeling concepts associated with {ImageDescr} in the first half of this chapter.

Class inclusion relationships have the properties of asymmetry. That is, for all ?x ?y, if ?x is a subclass of ?y, it is not the case that ?y is a subclass of ?x. Another way to express this is that if ?x is a subclass of ?y and ?y is a subclass of ?x, then ?x=?y. A class inclusion relationship is transitive, that is, for all ?x ?y ?z, if ?x is a subclass of ?y, and ?y is a subclass of ?z, then ?x is a subclass of ?z.

5.5.4 Instance_of relationship

An *instance_of* relationship is a primitive relation between a component instance and a class which it instantiates. For relationships, it is a primitive relation between a relationship instance and a class of relationship type which it instantiates. Not originally included in the content analysis form, the instance_of relationship is necessary in this chapter for expressing instance-level relationships between factual assertions in the Corpus and their corresponding relationship types.

5.5.5 Kinship relationships

Kinship relationships are used to represent a system of familial relationships describing image content. Kinship is a relationship between any entities that share a genealogical origin, through biological, cultural, or historical descent.

In the original analysis carried out in the Corpus, some relationship instances drew distinctions between family member relationships. These were classified as instances of attribution relationships and treated as image attributes much like describing the image of a workman in terms of his occupation 'steelworker.' During the data analysis portion of this investigation it became apparent that kinship-type relationships occurred often in the domain of image descriptions among all three categories of image describers.

Kinship relations can be characterized with properties that go beyond symmetry and transitivity. One type of kinship relationship may carry with it certain inferences that are not present in another type of kinship relationship. For example, if A is the mother of B, it can be inferred that A is a female parent and B is the child of A. In the case of father_of, however, if A is the father of B, the inference is that A is a male parent, but it still holds true that B is the child of A.

In the Scone core ontology with which his study aligns itself, there are concepts formalized that are relevant to kinship relationships, including {boy}, {girl}, {man}, {woman}, {son}, {daughter}, {parent}, and {offspring}. The kinship relationships that emerged in the Corpus, along with instances, are presented in Table 34. In the ontology of relationship types, the relationships in Table 34 are considered subtypes of the {kinship} relationship type.

Table 34: Kinship relationships

Kinship Relationship Type	Instance	Construction/Properties
brother_of	Brother_of("Andrew","Tom")	(?X{male} brother_of ?Y{male}) Brother_of is symmetric if the fact that ?X is the brother of ?Y implies that ?Y is the brother of ?X and both ?X and ?Y are male. If ?X is the brother of ?Y, then ?X is male. If ?Y is the brother of ?X, then ?Y is male.
child_of	child_of("Margaret Carnegie","Andrew Cargenie")	(?X{daughter} child_of ?Y {parent}) Child_of is asymmetric if the fact that ?X is the child of ?Y implies that ?Y does not stand in that relation with ?X. If ?X is the child of ?Y, then ?Y is the parent of ?X. If ?X is female, then ?X is the daughter of ?Y.
daughter_of	Daugter_of("Margaret","Carnegie")	(?X{female} daughter_of ?Y{parent}) Daughter_of is asymmetric if the fact that ?X is the daughter of ?Y implies that ?Y does not stand in that relation with ?X. If ?X is the daughter of ?Y, then ?X is female. If ?X is the daughter of ?Y, then ?Y is the parent of ?X.
father_of	father_of("Andrew Carnegie","Margaret Carnegie")	(?X{parent} father_of ?Y{daughter}) Father_of is asymmetric if the fact that ?X is the father of ?Y implies that ?Y does not stand in that relation with ?X. If ?X is the father of ?Y, then ?X is male. If ?X is the father of ?Y, then ?Y is the child of ?X. If ?Y is female and ?X is the father of ?Y, then ?Y is the daughter of ?X.

Kinship Relationship Type	Instance	Construction/Properties
first_cousin_of	first_cousin_of("Mrs. McKenna","Andrew Carnegie")	(?X{person} first_cousin_of ?Y{person}) First_cousin_of is symmetric if the fact that ?X is the first cousin of ?Y implies that ?Y is the first cousin of ?X.
mother_of	mother_of("Margaret Morrison Carnegie","Andrew Carnegie")	(?X{parent} mother_of ?Y{son}) Mother_of is asymmetric if the fact that ?X is the mother of ?Y implies that ?Y does not stand in that relation with ?X. If ?X is the mother of ?Y, then ?X is female. If ?X is the mother of ?Y, then ?Y is the child of ?X. If ?Y is male and ?X is the mother of ?Y, then ?Y is the son of ?X.
only_child_of	only_child_of("Margaret Carnegie","Andrew Carnegie")	(?X{daughter} only_child_of ?Y{parent}) Father_of is asymmetric if the fact that ?X is the father of ?Y implies that ?Y does not stand in that relation with ?X. If ?X is the father of ?Y, then ?X is male. If ?X is the father of ?Y, then ?Y is the child of ?X. If ?Y is female and ?X is the father of ?Y, then ?Y is the daughter of ?X.
relative_of	relative_of("John Patterson","Andrew Carnegie")	(?X{person} relative_of ?Y{person}) Relative_of is symmetric if the fact that ?X is the relative of ?Y implies that ?Y is the relative of ?X.

Kinship Relationship Type	Instance	Construction/Properties
sister_of	sister_of(image_of,"Margaret Carnegie")	(?X{image} sister_of ?Y{female}), where the first variable is to be filled by the NP "image" and the second variable is to be filled by a named person. In this instance, an image searcher asks for an image of Margaret Carnegie's sister. The sister_of relation infers that the image will be a picture of a female.
son_of	son_of("Andrew","William")	(?X{male} son_of ?Y{parent}) Son_of is asymmetric if the fact that ?X is the son of ?Y implies that ?Y does not stand in that relation with ?X. If ?X is the son of ?Y, then ?X is male. If ?X is the son of ?Y, then ?Y is the parent of ?X. ?Y is male, therefore ?Y is the father of ?X.
wife_of	wife_of("Margaret Ann Beatty Ferris",George Ferris")	(?X{female} wife_of ?Y{male}) Wife_of is asymmetric if the fact that ?X is the wife of ?Y implies that ?Y does not stand in that relation with ?X. If ?X is the wife of ?Y, then ?Y is the husband of ?X. If ?X is the wife of ?Y, then ?X is female and ?Y is male.

The *unnamed person* is a new idea of meaning that surfaced in searchers' descriptions of images using kinship relations. The format that accounts for this *unnamed person* description takes the form *x_of(image_of,y)* where *x_of* is a type of kinship relationship, *y* is a named person, and the image sought is a picture of an unnamed person. An ontology of kinship relationships is illustrated diagrammatically in Figure 33.

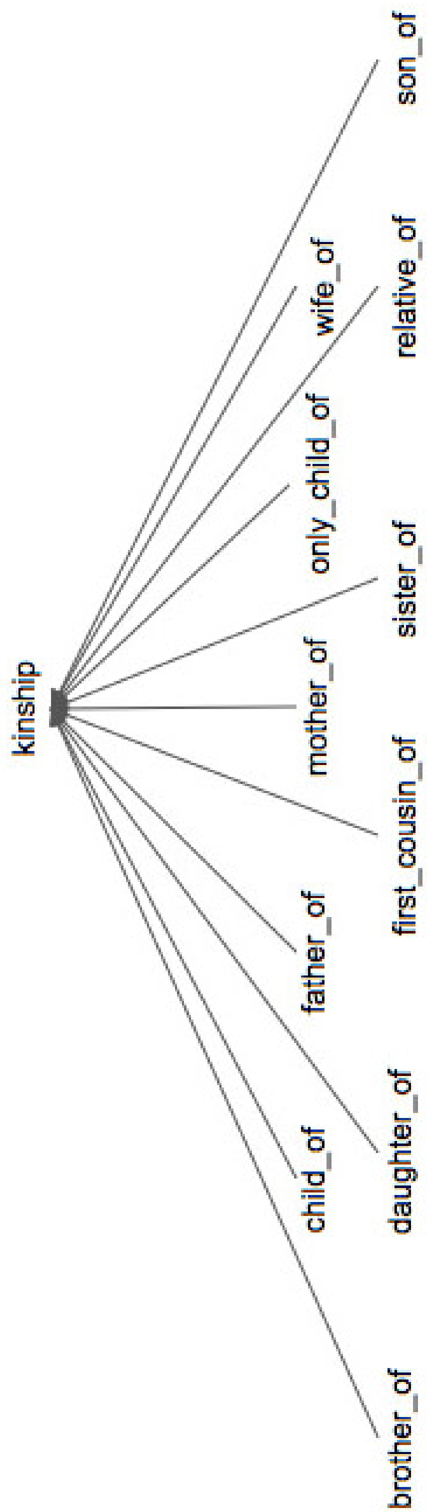


Figure 33: Partial ontology of kinship relationships

5.5.6 Meronymy

Meronymic or 'part-whole' relationships can be described as a part_of(x,y) construction. The task in this section is to determine what constitutes meronymic relationships in the domain of image descriptions, describe their properties, and model an ontology of the types/subtypes that emerge from the Corpus. The member-collection relationship, the most prevalent meronymic relationship that emerged in the Corpus, is examined here in some detail and the remainder of meronymic relationship types is described in Table 35. Consider the following description presented by an image researcher:

[013SEMNL03]*"The image in question has the alphanumeric code of 'L1272', and is accompanied by the text "Pittsburgh in the Evening, c1940s, Lorant Collection."*

The code "L1272" is the first entity in an ordered pair of arguments and it represents an individual photograph or member of a collection. The phrase "Lorant Collection" is the second argument in the proposition and is the entity that represents a collection. *Part_of* is the relational (meronymic) process implicit in the proposition. The elements of this relationship based on Stasio, Herrmann, Chaffin (1985) extended to image descriptions are: 1) Inclusion, that is, photograph L1272 is semantically included in the collection. Here {collection} is an abstract entity and inclusion at this level is marked by the "L" in the accession code, which stands for "Lorant"; 2) Partitive, that is, the physical manifestation of the photograph is part of a collection. A collection is a mental object and members of the collection—e.g., photographic prints and negatives—do not perform any particular function in a collection; and 3) There is an implicit part_of relationship between the collection and a superordinate entity, either another collection or institutional entity functioning as an archives or library. In the above part_of relation, arg1 is part of arg2 by social agreement. In other words, the archivist chooses to group these items together. Furthermore, the objects are related by an event: A book *Pittsburgh: The Story of an American City*, from which all of the images in the collection were taken.

The interesting point about this example is that the word "collection" marks a part-whole relationship type and the noun preceding the word "collection" names the collection. Knowing that the context is a cataloger's description adds force to the assumption that the words "image" and "accession code" signal that the image being described is part of the Lorant Collection. The relationship between the code and photograph is realized in a separate proposition that associates the identification code to an instance of photograph through the accession_code_of relation. This brings to light the necessity for a larger knowledge-based solution to semantic meaning of relations between not only words within sentences, but relations between sentences.

The following definitions are taken from Winston, Chaffin, and Herrmann (1987, p.421-426) and Story (1993, p. 463-466).

<u>Meronymic relationships between solid, physical objects</u>

- *Component-integral object* – "components and the objects to which they belong."
- *Portion-Mass* – "portions of masses, extensive objects, and physical dimensions…having parts which are similar to each other and to the wholes which they comprise."
- *Stuff-object* - "constituent of an object; the stuff cannot be physically separated from the object without altering its identity."

Meronymic relationships between abstract entities

- *Member-collection* – "membership in a collection; not requiring members to perform a particular function or possess a particular structural arrangement."
- *Feature-event* – "event may have parts (features) that occur at different moments in time."
- *Place-area* – "takes place between an area and special places or locations within it."

Transitivity properties hold between physical part_of relationships. That is, for all ?X ?Y ?Z, if ?X is a physical part_of ?Y, and ?Y is a physical part_of ?Z, then it can be inferred that ?X is a physical part_of ?Z. Winston, Chaffin, and Herrmann (1987) also characterized meronymic relations in terms of three relational elements: 1) Functional/nonfunctional, 2)homeomerous/nonhomeomerous, and 3) separable/inseparable. (See section Ch. 4, § 4.2 for a more detailed explanation of these relation elements.)

Winston, Chaffin, and Herrmann (1987) point out that it is difficult to distinguish between class inclusion (X is a type of Y, X is a kind of Y, X is a Y) and meronymy in situations describing activities. For example, in the photograph-v-template discussed under case relationships, it is ambiguous whether the activity of shooting landscape is a type of photography or a part of photography. Confusion also arises between class inclusion and member-collection because both concern membership in larger sets. Winston, Chaffin, and Herrmann (1987) argue that the distinction is made by determining if the relationship is based on an entity's similarities with members of a class, in which case it is class inclusion, or on the basis of spatial proximity or by social connection, which signals a member-collection part-whole relation.

Table 35: Types of meronymic relationships

Type of Meronymic Relationship	Instance	Construction
stuff-object	part_of(image,'same roll of film')	**(?X{image} part_of ?Y{film})**, where ?X is photographic image and ?Y is photographic film.
component-integral	of(gate,"Homestead Mill") of('board of directors',"Carnegie Steel”)	**(?X{component} of ?Y{thing}),** where "of" follows a noun indicating a given part. ?X is a component belonging to object ?Y.
member-IntervalInTime	part_of("Kennywood images",'pre-1940 time period')	**(?X{photograph} part_of ?Y{time interval})**, where ?X represents photographs and ?Y represents an interval in time.
member-collection	part_of(“Pittsburgh in the Evening”, “Lorant Collection”)	**(?X{photograph} part_of ?Y{collection})**, where ?X is a named photograph and ?Y is a named collection.
member-collection	from((of(photograph,"George Westinghouse")),"Pittsburgh Photographic Library Collection")	**(photograph_of ?X{person} from ?Y{collection})**, where ?X is a named person and ?Y is a named collection.
feature-event	part_of(fans,'overflow crowd')	**(?X{thing} part_of {event})**, where ?X is a thing (mental or physical object) and ?Y is part of an activity or event.
place-area	part_of(“Gold Room”,”City/County Building”)	**(?X{place} part_of ?Y{area})**, where ?X is a named room and ?Y is a named building.

This study extends Winston, Chaffin, and Herrmann’s (1987) set of six part-whole relations with the member-IntervalInTime relation, which is found in incident 031 when a curator describes a collection to an image searcher saying, “There are several Kennywood images from the pre-1940's time period.” The Kennywood photographs are members of a set defined by the time interval that presumably begins with the start of Kennywood amusement park and ends before the 1940s. An ontology of meronymic relationships is presented in Figure 34.

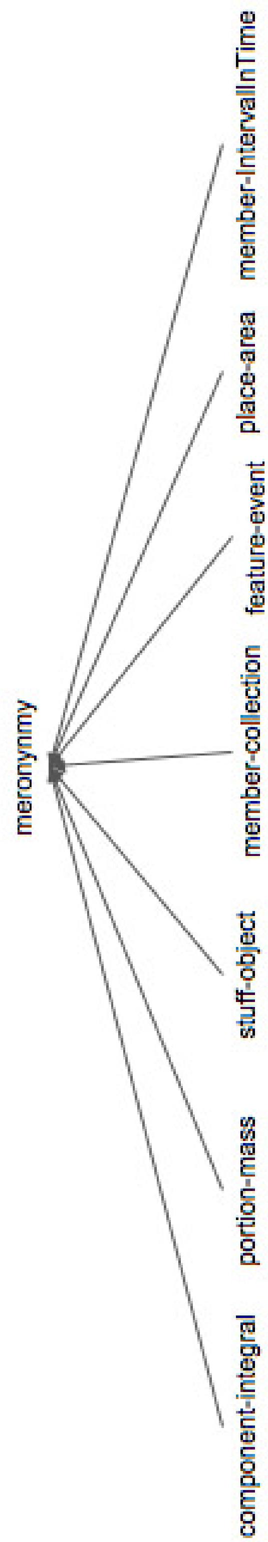


Figure 34: Meronymic relationships

5.5.7 Spatial

Spatial expressions specify that an object is at a place or that a situation holds at a location. The content analysis and data analysis carried out in Ch. 4, § 4.3.5.5, used Herskovits's (1986) *located entity/reference entity* model to explain simple locative expressions in the Corpus. This approach was insufficient for predicting all of the relationships involving prepositional phrases together with whatever the phrase modified. This study presented Gruber's (1965) Thematic Relations Hypothesis as explained by Ray Jackendoff (1983) (see Ch. 4, § 4.3.5.5) to make sense of the exceptions. While this proposal for analysis helped predict spatial expressions in semantic fields other than spatial, the predication of places and paths in image descriptions is shown to be much more complex than previously thought.

In the domain of image descriptions, if the situation being described is an activity or event, it is taking place at a location. If an object is specified, it is located at a place on the two-dimensional surface of a photographic print, or in the three-dimensional space where pre-photographic referents reside. Similar distinctions in meaning apply to events: an image viewer may describe an event depicted in an image or the photographic print participating in an event, such as a gallery showing. These distinctions fall roughly under the same categories of meaning described in Ch. 1, § 1.4, where the referent refers to a physical object in three-dimensional space and the depicted referent is part of the photographic image. A distinction that is less clear is the notion of depicted referents existing in different locations by virtue of the photograph's placement in different physical locations, for example, archives and gallery spaces.

The spatial phenomena considered during content analysis included seven different types of spatial relationships expressed in image descriptions:

1. Static locative
2. Directional locative
3. Both static and directional
4. Preposition followed by another preposition
5. Ambiguous
6. Preposition not followed by an argument
7. Spatial expression applied in other semantic field

Arguably, there are cases of spatial expressions being overlooked. For example, structures allowing for movement relationships that express paths of motion. No attempt is made to extend these descriptions beyond what is expressed in the Corpus of Relationships. Spatial relationships are binary relations except the *between* relation, which is ternary. The *near, far, across*, *faces*, and *beside* relations are symmetric, while the relations *right-of, left-of, above, behind, inside* are antisymmetric. Relations *left-of, right-of, above, behind,* and *inside* as well as their inverse relations are transitive.[25]

The table of spatial relationships (Table 36) and temporal relationships (Table 37) take a slightly different approach representing the kinds of semantic categories expressed by arg1, arg2, and arg3. The notation is simplified and uses the form *<relationship type name>, <semantic category1-semantic category2>*. This is located in column one and the more specialized subtype is in column two. An example of an instance taken from the Corpus is included in the third column. The types listed in column one are partially influenced by the original categories adopted during content analysis. Perhaps of greatest influence for a comprehensive set of spatial

[25] Moving these spatial expressions into a computational environment will necessitate finer-grained distinctions that take into account different frames of reference (from what viewpoint can it be said that X is standing to the left of Y) and different dimensions (e.g., a spatial reference in 3-dimensional space places one person in front of another even though the surface of a photograph is flat).

conceptual entities is Herskovits's (1986) set of elementary concepts divided at the top by a set of five sorts of entities: topological, geometrical, physical, projective (point of viewing is from perspective of person viewing photograph), and metric.

Table 36: Ontology of spatial relationships

Relationship Type, conceptual entities	Subtype	Instance
topological, event-physical object	across	across('iron delivery',"HotMetal Bridge")
topological, physical object-space	adjacent_to	adjacent_to(hotel,"Mt. Washington Streetcar Tunnel")
topological, physical object-space	arriving_at	arriving_at(trolley,park)
geometrical, geographical area-physical object	around	around(highway,building)
topological, institution-place	at	at("Edgar Thompson's Steel Works","Braddock, Pennsylvania")
topological, person-space	at	dining_at("Pittsburgh millionaires","Schenley Hotel")
topological, event-space	at	at('mass rally',"Pitt Stadium")
topological, person-space	at	at(laborer,(in(job,mill))
topological, person-place	at	at("Andrew Carnegie","Homestead")
topological, event-space	at	at(performance,"Pittsburgh Civic Arena")
topological, photograph of person-place	at	at((of(picture,"Carnegie")),"Skibo")
topological, person-event	at	at('bobbin boy',work)
topological, physical object-place	at	at("Ammon Swimming Pool","Bedford and Kirkpatrick")
topological, person-space	at	at("Slavic millworkers","Homestead boardinghouse")
topological, [person]-place	at	at(died,"Shadowbrook")
topological, person-physical object	at	at("Ethelbert Nevin",piano)
topological, event-space	at	at(all_around('shook hands'),"Carnegie Institute")
geometrical, person-physical object	behind	behind(fans,rope)
geometrical, physical object-space-geographical area	between	between('complex of building','second Avenue',"Monongahela River")
topological, physical object-photograph	bottom_of	bottom_of("Carnegie Library",picture)

Relationship Type, conceptual entities	Subtype	Instance
geometrical, photograph-space	hangs_at	hangs_at(image,"Eat n Park")
topological, physical object-environment	in	in(tank,snowstorm)
topological, person-space	In	in("Carnegie","Scottish landscape")
topological, space-place	In	in('telegraph office',"Allegheny City")
topological, [person]-place	In	in(lived,"Pittsburgh")
topological, event-place	In	in(married,"Pittsburgh")
topological, person-photograph	In	in("Carnegie and his mother",'same picture')
topological, physical object-space	In	in(blocks,background)
topological, photograph-book	In	appeared_in("A-159","Pittsburgh Survey")
topological, photogtraph-collection	In	in(photograph,"Carnegie Library's photo collection")
topological, geographical area-place	near	near("Monongahela River","Homestead")
topological, physical object-space	on	on('small hotel',"Carson Street")
topological, photograph-physical object	on	on((of(picture,"Frederick Law Olmsted Jr.)),Internet)
topological, person-space	on	on("Dorothy Height",stage)
topological, physical object-geographical area	on	on("Winnie Mae","Monongahela River")
projective, physical object-direction	on	on('tall building',right)
topological, physical object-space	on	on((with(building,""Christopher's Restaurant")),'top floors')
projective, person-direction	on	on("Andrew Carnegie",left)
topological, person-space	on	on("Andrew Carnegie",'golf course')
geometrical, physical object-geographical rea	over	over("St. Louis Bridge","Mississippi")
topological, space-activity	place_of	place_of("Monongahela Wharf",activity)
Projective, person-person	right_of	right_of("John Phipps","Andrew Carnegie")
geometrical, space-physical object	surrounding	surrounding(area,"Civic Arena")
geometrical, physical object-physical object	under	under(bridge,"Schenley Memorial Fountain")

Relationship Type, conceptual entities	Subtype	Instance
Projective, direction-identify	direction_of	direction_of(left to right',identification)
Projective, manner-direction	direction_of	direction_of(standing,'left to right')
Projective, manner-direction	direction_of	direction_of(seated,'left to right')
Projective, person-person	left_of	left_of("Andrew Carnegie","John Phipps")
Projective, person-direction	on	on("Andrew Carnegie",left)
Projective, person-direction	poses	poses("Andrew Carnegie",'on the left')
topographic, space-person	position_of	position_of('front row',relatives)
Projective, space-person	position_of	position_of('in the back',"Miss Lauder")
Projective, person-person	right_of	right_of("John Phipps","Andrew Carnegie")
topological, person-person-person	seated_between	seated_between("Margaret Carnegie Miller",father, mother)
Projective, person-person	seated_left_of	seated_left_of("Miss Lauder","Mrs.Andrew Carnegie")
Projective, person-person	seated_right_of	seated_right_of("Mrs. Andrew Carnegie","Miss Lauder")
Projective, person-person	standing_left_of	standing_left_of("George Lauder","Mrs. William J. Holland")

A diagrammatic view of the ontology of spatial relationships is presented in Figure 35.

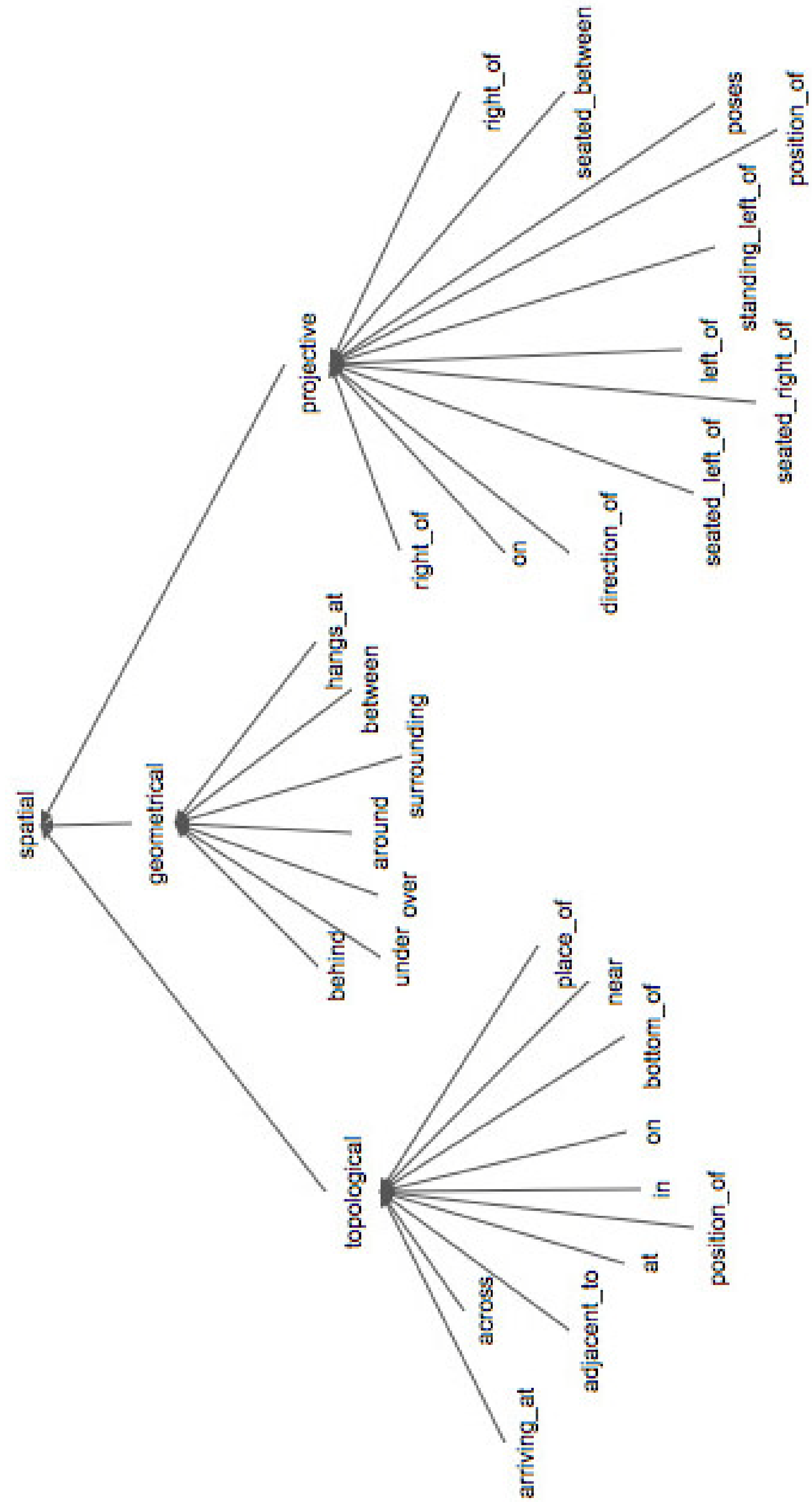


Figure 35: Ontology of spatial relationships.

5.5.8 Synonymy

Synonymy is not treated as a type of conceptual relationship in this study. Synonyms and near-synonyms are important relations for natural language processing and semantic lexicons, but there is no place for them in ontologies where concepts should be unambiguous and where, for example, all topological spatial concepts between a physical object and a space are expressed using the same concept. Therefore, from the viewpoint of this dissertation, the issues of what do with the instances of synonymy are external to the ontological modeling addressed in this chapter.

5.5.9 Temporal

Temporal relations are described by the interplay of two more specialized temporal relationship types: *time-point* and *time-interval*. The expression *time-interval* specifies the time duration or the period of time when a state persists. In connection with photographic images, it is used to specify the duration of depicted events and states, the period during which a situation exists or in what period of time it exits. The expression *time-point* corresponds to the intuitive notion of a point in time. In contrast, *time-interval* shows an explicit beginning and ending of a photography event, for example, the period of time during which a series is created, a collection is built, or an exhibition takes place as demonstrated in the following description:

> *"Luke Swank: Modernist Photographer," retrospective photography exhibition at Carnegie Museum of Art, November 5, 2005 – February 5, 2006.*

The ontology of temporal relations is presented in Table 37.

Table 37: Ontology of temporal relationships

Relationship Type, Conceptual Entities	Subtype	instance
time-point, space-event	after	after("Smokey City","Pittsburgh Renaissance")
time-point, physical object-event	after	after("Avro Avian",'forced landing')
time-point, space-event	after	after('exterior of house',marriage)
time-point, physical object-date	around	around("Civic Arena",1970)
time-point	before	before("Smokey City","Pittsburgh Renaissance")
time-point, space-event	before	before('Red Light District',"Cultural District Transformation")
time-point, person-event	before	before("Margaret Carnegie",death)
time-point, date-person	birth_date_of	birth_date_of("1827","Jacob Jay Vandergrift, Sr.")
time-point, physical object-date	built_in	built_in("Hand Bridge","1839")
time-point, space-century	date_of	date_of("Pittsburgh skyline","21st century")
time-point, photograph-date	date_of	date_of(photograph, "1852")
time-point,event-date	date_of	date_of("Johnstown flood","1889")
time-point, named photograph-date	date_of	date_of("L1272","c1940s")
time-point, person-date	death_date_of	death_date_of("1899","Jacob Jay Vandergrift, Sr.")
time-point, physical object-date	dedicated	dedicated(bridge,"September 2, 1918")
time-interval, event-era	during	during("Republican gatherings", "1850-1860 era")
time-point, space-event	during	during("Forbes Field and surrounding area",'"night game')

Relationship Type, Conceptual Entities	Subtype	instance
time-point, abstract object-decade	flourishing_in	fourishing_in('industrial economy','last three decades')
time-interval, photograph-time span	from	from('air pollution images',"1960s-1970s")
time-interval, event-time span	from	from(image,"1970s")
time-point, photograph-date	from	from(photo,"1961)
time-point, event-date	has_PointInTime	has_PointInTime('assassination attempt',"July 23, 1892")
time-interval, event-date-date	held_from	held_from(concerts,"1896","1943")
time-point, event-date	in	in("Railroad Riot", "1877")
time-interval, space-time span	in	in("Pittsburgh","1840s-1850s")
time-point, physical object, date	in_and_around	in_and_around("Civic Arena",1970)
time-interval, span of time-photograph	IntervalInTime	IntervalInTime_of("1900s up to the 1970's",image)
time-interval, span of time-person	lifespan_of	lifespan_of("March 1, 1807 - December 18, 1882","James Laughlin")
time-interval, physical object-time span	of	of("Ninth Street Bridge","1890-1925")
time-point, event-date	of	of("Liberty Avenue Flood","1936")
time-point, event-date	on	on(show,"May 2, 1970")
time-point, physical object-date	prior_to	prior_to('production facilities', "1892")
time-point, event-date	published_in	published_in("Pittsburgh Survey","1910")
time-point, event-date	taken_in	taken_in(picture,"1910")
time-interval, time span-space	time_span_of	time_span_of("1840s-1850s","Pittsburgh")

Relationship Type, Conceptual Entities	**Subtype**	**instance**
time-interval, time span-physical object	time_span_of	time_span_of("1890-1925","Ninth Street Bridge")
time-interval, time span-event	time_span_of	time_span_of("1896-1943",concernts)
time-point, event-date	until	until(housed,"2008")
time-point, space-date	until	until(home,"1920")
time-point, event-date	year_of	year_of(snowstorm,"1950")

A diagrammatic view of the ontology of spatial relationships is presented in Figure 36.

Figure 36: Ontology of temporal relationships

5.6 SUMMARY

The primary aim of this chapter has been to characterize types of relational expressions—that is, the classes to which relationships belong by virtue of their meanings and the properties they posses. The inherent complexity of how humans express relationships should be apparent by now. The range of relationship phenomena being studied is essentially unbounded; there are potentially an infinite number of variations between relationship types and some differences are considered important, while others are not. Ultimately, the level of qualitative resolution (the number of subclasses belonging to a class) will depend on the resolution requirements of a specific application domain. The current study is grounded in Library and Information Science and is interested in the inferences that can be derived from the types of relationships distinguished in this chapter.

In the domain of photographic image descriptions, the resolution problem is magnified by the existence of multiple layers of meaning in different contexts. Most significant are the distinctions drawn between the scene prior to image capture and the scene depicted in the photographic image. In other words, a thing behaves one way as a physical object in 3-dimensional space and another way when it is depicted as a picture element lying on the surface of a geometrical plane. However, the Corpus demonstrates that humans describe these two situations using many of the same expressions. The listener or reader of the image description draws necessary distinctions based on context.

Classifying instances of relationships into families of relationship types resulted in: 1) Semi-formally defining the top-level relationship types, 2) Determining the set of characteristic properties or constraints that help distinguish one relation type from another, 3) Generalizing relationship arguments to semantic categories, and 4) Providing semantic networks showing diagrammatic views of the ontology of relationships and their association with the entity Image Description{ImageDescr}.

The semantic category analysis provided in each table still has problems. For a proper representation of primitive semantic categories joined together by relationships, a classification of conceptual entities is needed. The core ontology formalized in Scone was helpful, but to build a comprehensive computational lexicon representing image-description entities joined by arguments, it would be necessary to introduce additional, finer-grained categories.

The essential characteristic of the ontology of image description relationships (OIDR) became most apparent in the section describing case relationships. No longer seen as adequate, exhaustive representations of reality, image descriptions humans verbalize are regarded as incomplete, fragmentary views expressing only those aspects of image that are salient for the viewer. There is a need in the ontology of the image description relationships to emphasize the context-dependent and event-oriented nature of the image description and account for continuities between experienced events and the background knowledge viewers bring to descriptions. Case relations partially fill this role acting as schemas in what is described in this section as *situation-templates*—structures that enable the researcher to use a single event concept represented in an image description to understand and represent a larger system.

Characterizing an image description by embedding concepts and relationships into this event space introduces intrinsic and extrinsic meaning.[26] An ontology of relationships might include these two concepts to further refine the properties of relationships expressed in the domain of image descriptions. The relationships of the intrinsic level are primarily used to

---

[26] J. Michael Dunn (1990) states, "Metaphysically, an intrinsic property of an object is a property that the object has by virtue of itself, depending on no other thing" (1990, p. 178). An extrinsic property depends on a thing's relationship with other things.

describe states of affairs and objects that are part of the depicted image and the physical carrier of the image. Extrinsic relationships join together entities existing outside the image and physical manifestation of photographic prints themselves.

The contrasting of intrinsic and extrinsic aspects of relationships in descriptions has not been considered in LIS schemas, especially in a system of image descriptions represented as a propositional account of events, objects, and states of affairs. The Corpus provides evidence that these contrasting layers of meaning are integrated into the description itself.

Finally, it is important to be clear about the mixed ontological status of relationships. In the first place, the associated factual assertions, schemas, and semantic categories, which are called situation-templates, are types of information objects that take on the same ontological presuppositions as concepts, entities, objects, and the like.

In the second place, each of the instances of a schema is another information object. As seen earlier, the situation-templates for the photograph-v event is made up of a set of seven schemas. The expression *has_agent(take,"Corsini")*, for example, is an instance of the agent-action schema and an information object—a proposition. The term "instance" is used here and throughout this dissertation as a relation term for a relation certain abstract entities bear to certain classes of abstract entities. The ontological status of a situation-template (abstract entity) is the same as that of a photograph-v event and the ontological status of a relationship type is the same as an instance of a type.

6.0 CONCLUDING REMARKS

With a methodological framework (depicted in Figure 22, p. 101), a Corpus of Relationships, and the ontology of relationship types in hand, it is time to draw some conclusions from the forgoing analysis of how humans express relationships in English when describing images. The section immediately following this introduction looks at how effective content analysis was for capturing different kinds of knowledge—implicit and explicit. Section 6.2 explores the benefits of carrying out analysis in the context of the Corpus as a whole versus one relationship at a time. Section 6.3 extends Green's (2001) definition of relationship to take in relationship instances in the domain of image descriptions. Section 6.4 brings into focus the role of factual assertions in conveying meaning and that the meaning conveyed is context dependent. Section 6.5 contrasts the ontological, multi-contextual view of image descriptions with the more traditional stand that views textual descriptions as flat, unitary database records. Section 6.6 suggests viewing image descriptions as mental situations where relationships can be assigned four different levels of meaning. Section 6.7 offers a new relational view of image attributes. Finally, the two concluding sections discuss the limitations of this study and offer suggestions for future research.

6.1 ALLUSIVE NATURE OF BACKGROUND KNOWLEDGE

One conclusion that can be drawn from this study is an ontological one pertaining to what exists in the domain of image descriptions and that they can be understood in terms of relationships inducing coherence between concepts. As it was explained in chapter three, the task of the content analyst in this study was to recognize and capture the essential features of relationships in real world instances of image descriptions, and then to regard these features as meeting certain conditions for relationship types to be associated with image descriptions. This project is regarded as only partially successful because image descriptions failed to provide precise, orderly, structured semantic patterns. They take a different route. Humans deviate from ideal conditions and text gets misshapen giving it depth, which in turn requires deep analysis. The conditions for predicting relationships using content analytic techniques are inflexible and in this sense are unlike real world image descriptions.

In everyday circumstances, image descriptions also differ in regard to their completeness. The evidence shows that humans verbalize just enough to convey to the reader or listener the meaning of an image in a given context. The full meaning is never communicated, especially in the case of image researchers who are oftentimes attempting to describe unknown, possibly non-existent photographs. It is typical for the image searcher to describe an image using only sentence fragments and to offer lists of subject terms; the curator relies on background knowledge to unravel the implications of the queries; catalogers' descriptions not only have interpretive, aesthetic dimensions, they also have meanings cued in part by description standards and local processing procedures. Content analytic techniques do not systematically account for these background influences. There is a need, therefore, to look for ways of ascertaining and capturing this knowledge in a systematic way and for making valid and replicable inferences from the background knowledge.

This is not a conclusion that should disappoint researchers in Library and Information Science or discourage further research into the relationship problem using content analysis techniques. One should expect content analysis to be effective for identify things like *well-formed relations*—explicit statements that include relational terms and arguments—describing spatial locations of objects on the surface of photographs, dates when photographs were captured, and a photograph's membership in a named collection of photographs, but not for identifying unnamed individuals in image searcher's descriptions such as who is the mother in, "I want a photograph of Andrew Carnegie's mother." Nor should content analysis be expected to identify and make explicit the viewer's or camera's viewpoint expressed in the curator's statement, "I found a few photos of the stage area and "audience" from the stage," or to formulate unnamed places in spatial references such as, "Arrival and encampment of the PA militia, led by General Snowden in July, 1892," even though it is apparent by viewing the reference incident in its entirety that the Pittsburgh borough named "Homestead" is the place where General Snowden arrived. The image description's resistance to this sort of ontological analysis is the ground for further questions about how to ascertain and capture not only the dominant voice in a description, but also the implicit, inner voice of background knowledge and inference when it takes precedence.

There are research frameworks that could be useful in soliciting and representing domain-specific background knowledge about entities, attributes, states, and events in different contexts. Minsky presented the *frame* as a "data-structure representing a stereotyped situation" (1974, p. 212); Schank and Abelson's scripts (1977) refer to knowledge structures for sequences of events—a framework that could be applied to, for example, a photographing event; and Fahlman's (2006) Scone knowledge base and inference engine that supports frame-like structures. In adopting semantic frames called *situation-templates* for describing events in chapter five, for example, a single concept representing a particular situation brought oblique background knowledge to the surface. For example, in reference incident 035, a researcher asks for an image of the "Phipps Conservatory interior." By applying a 'viewpoint' semantic frame to this subject, the content analyst examines a larger set of semantically related entities linked to the frame that brings to the surface what takes place in an interior scene, that the image is captured by a photographer in a particular location on a given date, and so on.

6.2 TAKING THE BROADER VIEW

Putting aside the complexities of real world image descriptions and the roles played by background knowledge and inference, the research brought new insight to the data analysis process itself and suggests a slightly different approach next time. It was discovered that after the analysis of the sample was complete and viewed as a large collection of instances, new, domain-specific families of relationships emerged. Significant evidence came in recognizing and capturing kinship relationships. Kinship relationships were not easily recognized when parsing over 1,600 factual assertions, one assertion at a time in linear order. This suggests that important questions concerning content analysis include to what extent can relations be predicted from formal axioms and relation elements alone and can certain classifications of relationships be predicted only during post analysis of instances? In other words, are some relationships only properly understood in terms of their function within the domain of interest? The *domain of* *interest*, necessarily abstract in this study, is represented in the Corpus in such hypothetical terms as semantic units, factual assertions, arguments, and relational terms. In one respect, looking for answers in the domain of interest and not in formal axioms makes sense; otherwise instances of relationships would have the same meaning in every domain of interest. This domain analytic approach still leaves open the possibility that, within a particular domain of interest, classificatory decisions may be determined from ontological rules applied during Corpora construction.

6.3 RELATIONSHIP DEFINED

Before moving on to a detailed look at the role context plays in understanding relationships, it will be useful to stand back and reconsider what a relationship is. In chapter one, this investigation began with Green's (2001) most basic definition of relationship as, "an association between two or more entities or classes of entities" (p. 3). Based on this study's analysis, a different definition of relationship is offered as it pertains to its various instantiations in image description knowledge.

A relationship instance in an image description is the element or set of elements that an association among two or more objects has, that distinguishes one relationship instance from another, that identifies it as a member of the same family that the same kind of relationships are members of. The meaning of a relationship instance is the element or set of elements that an agent who interprets visual information attributes to that instance that distinguishes that relationship from another.

This brings into focus the nature and the formal properties of relations. The *associations* Green refers to are neither simple terms nor pure conceptual entities, but word senses that correspond to conceptualized uses of a word. The implications for representing relationships in computational environments are that relationship concepts must be shared between agents—humans and machines—even if it is only implicit.

6.4 THE MEANING DILEMMA AND THE ROLE OF CONTEXT

The research described here shows that an adequate representation of the kinds of entities that inhabit the domain of image descriptions must take into consideration the context in which processes, activities, events, and states of affairs are involved. Brachman and Levesque (2004) state, "the concern of semantics is to explain what the expressions of a language mean," and they maintain that in order to understand what it means to believe something, the claim a sentence makes about the world must be specified (2004, p.18). The problem in carrying this out in the current study was that semi-formal analysis represents facts in non-logical expressions and the words occurring in factual statements are context-dependant. For example, the constants "Winnie Mae" and "Monongahela River" (see spatial expressions, Table 36) are used in a photograph description to mean one tugboat on one river. Someone else may use the name "Winnie Mae" to mean another tugboat, or more likely, someone's aunt. In this sense, a general claim cannot be made about the meaning of the photograph description, "The tug Winnie Mae on the

Monongahela River," a statement made about an individual photograph in the Pittsburgh Photographic Library collection with accession code 3865. [27]

The claim that can be made is that the assertions distilled from the natural language statements follow a particular form of interpretation in the context of the domain being investigated, that meaning is derived from this interpretation, and that words in the sentence—not logical symbols—convey the meaning.

In the Corpus, the above sentence fragment is re-presented as four factual assertions:

instance_of("Winnie Mae",tugboat)
on("Winnie Mae","Monongahela River")
accession_code_of("3865",photograph)
subject_of((on("Winnie Mae","Monongahela River")),"3865")

The words in these factual assertions are not atoms. An atom is considered to be an absolute simple that cannot be further understood. For instance, the term "tugboat" is complex and can mean a specific type of boat, of which there are many types and while tugboats may share similar attributes, shared attributes may exhibit different values (e.g., different colors, gross weight, captains, and so on). What can be said is that the instance relationship

instance_of("Winnie Mae",tugboat)

---

[27] There is an implicit relationship that holds between factual statements and semantic units in the Corpus. It was explained in the conclusion to chapter four that the connection is made through the numbering system. The semantic unit "The tug Winnie Mae on the Monongahela River," is numbered 037SEMNL**05** and the three associated facts are numbered consecutively 037FACTS**05**13-15. The relationship, there for, is *part_of(037FACTS**05**13-15, 037SEMNL**05**).*

makes the claim that the entity going by the name “Winnie Mae,” whatever that is, is an instance of “tugboat,” whatever that might be. The factual assertions, therefore, specify the meaning of sentences as a function of the interpretations of predicates.

Specifying the meaning of sentences based on this view of the world relies on these simple working assumptions: First, for a given sentence describing a photograph, some arguments will satisfy a particular relation or function and others will not. An interpretation of the semantic unit decides whether an argument does or does not have the necessary properties. For example, in the description

011SEMNL12 *“…negotiations between Frick and strikers”*

the *between* relation is a ternary predicate (arity 3) and an interpretation of the semantic unit determines which objects stand as which arguments in this factual assertion:

between(negotiations,Frick,strikers)

Second, the meaning of ‘between’ in a given context (here the choices are referent, depiction of referent, object showing depiction, linguistic, and conceptual) is also a function of the interpretation of the predicate. For any predicate of arity 3, some objects will satisfy the predicate and some will not. Negotiation in the context of this image description describes an event that is taking place *between* referents—human beings (Frick and strikers)—not objects pictured in a photograph.[28]

Third, for any function f of arity 2, some of the objects will satisfy f and some will not. For instance, an interpretation of a function of arity 2 decides on which pair of objects map to arguments. For example, the meaning of date_of in the current domain of interest could be a specific mapping from dates to an individual photograph. The interpretation, however, is also influenced by the context in which the statement is made. A cataloger assigns specific dates to specific photographs as capture dates, so the predicate is refined to capture_date_of. An image searcher asking for photographs based on subject, however, may only specify a general time period, for instance, pictures of steel mills during the 1950s. A time interval in this context would not satisfy the capture_date_of as applied by a cataloger, but would satisfy the more general date_of relation used to specify the date of an event, object, or state of affairs a viewer interprets to exist in an image. The formula derived here, therefore, is: The meaning of function f (date_of) is a specific mapping from x (dates of events, objects, states) to y (individual photographic images).

To summarize, once again it is clear from the results of this dissertation that relationships cannot be predicted on axioms alone. At the very least, there is also a close relationship between the type of group doing the describing and the type of description that results because each group

[28] This goes straight to the heart of Gruber's Thematic Relations Hypothesis (1965). Applied here, the *spatial* relationship *between* helps explain the *event* relationship *between*. {Negotiation} fills the located entity slot; {Frick} and {striker} fill the reference entity slot; *between* fills the role played by location in the semantic field of spatial expressions. If an event E is referred to in a photograph description involving kinds of objects, as it is here, then the pre-photographic referent domain character of E is described. That is, E is described in terms of the objects it involves, their properties, and the relevant relations that hold between instances of those kinds when the process is instantiated.

brings to the description different background knowledge based on different assumptions. Post analysis discovers the domain-specific relationships and properties of those relationships.

Thus, there are distinct contexts in which references can occur and these are crosscut by the meanings different describers (contexts) bring to the description. This, in effect, results in three different corpora: one describing catalogers' viewpoints, another image searchers', and still another describing curators' viewpoints.

6.5 NEW ACCOUNT OF IMAGE DESCRIPTION KNOWLEDGE

Surprisingly perhaps, images were not examined during this investigation. Instead, this study brought into focus a view of image descriptions as types of information objects worthy of explanation. From one perspective—what might be called the traditional viewpoint—memory institutions have been looked upon as contexts in which images are described and classified. Museums, archives, and libraries build associations between photographs and other image objects through a variety of means, including subject cataloging, descriptive cataloging, and by organizing collections of photographs into files, series, and fonds. Local rules and guidelines, international standards, thesauri, and the catalog records themselves define these roles. In this traditional view, catalogers describe images in terms of attributes. The description is sometimes standards-based and entails a process that can be traced from early picture card catalog practices, through Shatford Layne, to present day cataloging models.

Followers of Shatford (1986) treat images as primitives, as visual objects that explain associations with other objects, but the image itself is not analyzable in terms of relationships. Shatford extended Erwin Panofsky's iconological method to the categorization of subjects enabling catalogers to more systematically explain what a picture was *of* and *about* both generically and specifically. These attributes were not analyzed or decomposed into smaller parts.

The role of ontologies is not to replace current controlled vocabularies and description practices. Ontological constructs are capable of being intuitive and can assist in aligning the underlying assumptions of different thesauri. A critical role of ontologies, therefore, is to collect together, utilize, and make interoperable the terms and semantics that already exist in thesauri such as *Art and Architecture Thesaurus (ATT)* and *Thesaurus for Graphic Materials (TGM)*.

What this dissertation does propose is a shift in understanding how descriptions come about and the role of relationships in describing images. This alternative view of image descriptions and in general, of *description knowledge*, manifests itself in a number of ways. First, textual descriptions of images are considered as themselves subjects for explanation whereas in the traditional approach it is the image that is interrogated for meaning, not the description.

Second, it has been argued that relationships reside within image description texts—real-world entities—and by adopting content analytic methods relationships can be recognized, analyzed, sorted, and categorized, and then used to express the meaning of descriptions in computational environments using symbolic representation. In the traditional approach, a small set of semantically overloaded relations amount to *a prior* entities residing in thesauri, headings lists, and other structured vocabularies. Their primary role is to control what terms catalogers use when constructing records.

Third, relational inferences and multiple inheritance emerge when engaging ontology's role in representing and structuring relationships into families of relations. The traditional bibliographic record, on the other hand, is flat and connects to other records through surface structures matching brittle strings of text.

The fourth and final way in which the traditional and alternative views can be compared and contrasted is that in the traditional model, image descriptions are viewed as unitary records housed in institutional catalogs. This dissertation examines a multidimensional context in which constellations of visual interpreters bring meaning to images. When viewed this way, image descriptions are not the unitary catalog record, but envelope a discursive space that includes catalog records and dialogues between curators and image searchers. This model, which could be called *relational description*, recognizes the assumptions and background knowledge implicit in how different describers bring meaning to descriptions. According to the unitary catalog approach, subject terms describing photographs are determined authoritatively by ascertaining whether they are listed in structured vocabularies and it is within that structure where terms are connected through labeled associations.

6.6 LEVELS OF MEANING

What emerged from this study's description analytic view is the recognition that relations in visual information reside in image descriptions, consist of more elementary parts (predicates and arguments) and can be explained, organized, and represented in a semi-formal ontology.[29] Making relationships and their properties explicit in a knowledge base makes reasoning possible.

[29] It is important to restate why the ontology presented in this dissertation is regarded as semi-formal. First, it does not attempt to specify relationships in a formal language. Relationships are expressed in natural language. Second, the ontology as such is language-dependent, which brings with it certain ambiguities, some of which were pointed out during analysis.

The default values with exceptions are made possible by representing photography events using frame-like structures or *situation-templates* and then formalizing these structures in knowledge-based systems that support these features.

There is a deep irony here. The common, unquestioning appeal to structured vocabularies and headings lists within LIS is simply not mirrored within the empirical evidence collected in this dissertation. Indeed, modern LIS descriptive and subject cataloging evolved in the understanding that machine-readable descriptions must be reduced to lists of words. The claim made here is not that the very idea of words is inconsistent with human understanding of image meaning. The contention, rather, is that the simple-minded notion of accounting for image meaning via strings of text alone is undermined by ontological considerations. A "word" is something meaningful and might have more than one spelling or more than one meaning or *word sense*. There is a lot more to using a word than just using the string that spells it.

Another main conclusion of this study, therefore, is that image descriptions represent image content as abstract events or states of affairs—that is, *mental situations* in which mental objects are described and associations are built between relational items. These entities lie beneath the surface structure of words. This should not be surprising. The semantic structures of sentences have often been analyzed as describing events and states of affairs. See, for example, Schank (1973) and Norman and Rumelhart (1975). When image descriptions are viewed in this light, it is possible to assign various relationships to different events and states at a variety of levels of meaning.

First, the Corpus provides evidence that there are mental objects described as referents that exist before now in an original photographing context—that is, people, places, things, and events that are viewed in the original scene. Relationships are described in this context as associations between elements in the scene, such as the caption “The farmer is leaning on the fence post while the cows are walking to the barn to be milked.” The key distinction here is that individuals describe situations as if the event is taking place in the present moment.

Second, there are *situations* where mental objects being described are depictions of referents. That is, images of people, places, events, and states of affairs that are pictured in known photographs or are described as being ‘desired’ components of yet-to-be discovered photographic images (See Ch. 5, § 5.5.2.4). An example is an image searcher stating, ‘I am looking for a photograph that shows a mother with her children engaged in some activity that depicts motherhood’ or ‘P-7080 is a picture of the Homestead Steel Mill in Pittsburgh taken during the year 1920.’ The key element here is that references are made to existing pictures and photographic images or desired photographic images.

Third, the mental objects humans make reference to in image descriptions can be analyzed as textual entities. Take, for example, the proposition:

The photograph’s subject is industries.

Case describes the kinds of relationships that hold between nouns and pronouns with other words in a sentence that generally describes an activity or event. In this example, a relationship is observed within the sentence itself where the head noun ‘photograph’ is in the genitive case showing possession of the noun ‘subject’ by ‘photograph’. In the next example:

Picture of Eugene V. Debs.

The preposition "of" expresses composition or substance and that the picture is made of this entity's visual information, that is, Eugene V. Debs. ‘Of’ directly follows a noun, indicating the material of which it is made. Or maybe it is an example of near-apposition, relationship between two nouns "Picture" and "Eugene" placing them close together. In a sense, we are talking about two abstractions here placed close together in conceptual space. Explore as a spatial relation.

Fourth and final, the mental concepts humans describe may be conceptual entities that have no counterpart in the world of substance and are not part of the original context or scene being photographed. For instance, in reference incident 045, semantic unit 94 includes this statement in the image description: “Andrew Carnegie built his one-man rule on a competitive system that rewarded the shrewdest and strongest with partnership.”

These observations help bring into focus a constant tension between concepts and relationships. In order to properly define a relationship, one needs to classify the conceptual entities the relationship joins together. Without an ontology of the conceptual space of image descriptions, it is a difficult task deciding whether a given relationship may be asserted about two (binary relation) or three (ternary relation) given concepts. This problem was further complicated by the treatment of arguments as primitives, even though many arguments in the Corpus are complex semantic structures in themselves—a concern briefly touched upon in Ch. 4, § 4.2.

6.7 ATTRIBUTE AS RELATION

There has to be evidence that there are at least two items (a dyad of items) for a relationship to exist, not just a property being said of one item. In this study, the relationality of attributes was brought into question because an attribute does not bring two specific items into relation. The subject attribute “Steel Industry,” for example, does not hold as a relationship because there are not at least two items being related. The subject attribute adheres in “Steel Industry,” but there is no relationship between subject and “Steel Industry.” Rather, it is a monadic predicate that can be represented as subject(steel industry) and read as “has subject steel industry.”

This study provided a solution for representing attributes that act like a relational predicate in the domain of image descriptions—a domain where attributes hold together two entities: an attribute and the concept {photograph}, the latter being either an individual instance of photograph or the concept {photograph}. The view offered here, therefore, treats attributes in image descriptions as being dependent on the physical manifestation of the photograph.

6.8 LIMITATIONS AND SUMMARY

There are definite limits to the claims made in this study. Recognizing and capturing semantic relationships in image descriptions is neither straightforward nor simple. There are contributing factors that potentially threaten consistency in coding relationship instances including other possible readings of the data and other viewpoints in which the Pittsburgh Photographic Library Corpus data could function as well. In some cases, theoretical constructs are applied to reduce risks, for example, Hood’s Criteria of Relation in re-presenting factual assertions.

The content analysis demanded a lot of small, systematic steps, weaving together observations from explicit facts and inferential knowledge. Identifying an instance of a particular relationship type in text is a non-trivial task demanding an either/or-type of inference about class membership. For instance, to identify a description as expressing a projective spatial relationship type, the content analysts has to first circumscribe the boundaries of what sort of thing a projective spatial relationship is and then show that the instance in question possesses the applicable properties. This was further complicated by the limitations of the theoretical constructs and the multiple possible interpretations of a basic factual assertion. Axiomatic definitions of ontological relations (Bunge, 1977) are important but difficult to unravel and natural language definitions of relationships are at times inconsistent and ambiguous. Some of these concerns are alleviated by making assumptions clear throughout the investigation and by attempting in chapter three to define the context in which the Pittsburgh Photographic Library Corpus data is used.

It became evident during analysis of relationship instances and modeling of the ontology that there are prototypical cases of relationship types and borderline cases—regions of relationship types that are problematic. This should not be surprising considering the fuzzy nature of flesh and blood reality. There may be special cases that are properly understood as relationships in linguistic expressions but not textual expressions of visual information. Whatever borderline cases were considered, the intention was to determine which relationships were ontological and which were instances particular to photograph collections and libraries. Eventually the argument was made for the existence of an ontology or *family tree* of relationship types and instances and a representational form—a type of construction in the form **(?X{thing} <relationship> ?Y{thing})** to specify the semantic categories of arguments. Specifying what a relationship type's generalized semantic categories are is an attempt to attain the essential quality that all such relations share in common. Constructions are, in this respect, not definitions, but constraints on the use of relationships.

Finally, by using real life incidents as the data set, it was difficult to control or suppress what might likely be highly unstructured, unwieldy variations in form, which is a reflection of the fact that ordinary people—curators, catalogers, and archives users—behave in different ways when describing and searching for images. Whereas surveys and structured interviews offer participants predefined choices that are easy to tabulate, the view held in the current investigation is that data becomes interesting after it is created and that real-world descriptions like those examined in this study need explanation.

In summary, this dissertation was essentially a feasibility study evaluating the possibilities and the difficulties of capturing relationships humans express. It was a first pass at providing empirically grounded evidence that instances of semantic relations can be extracted from natural language descriptions of photographs and that these relationships can, with some difficulty, be identified and classified according to categories and structured as an ontology of relationship types. Criteria for validating this first account of relationship types is grounded in the theoretical writings of Aristotle, Hume, Chaffin, Herrmann, Winston, and others. Throughout the dissertation, the main justification has been that the content analytic technique and ontological tools and methods handle relationship phenomena that other analyses do not. However, there were other justifications. These approaches formed the network of correlations, or a substantial portion of them, necessary to connect the Pittsburgh Photographic Library Corpus data to the answers to research questions posed in this dissertation.

It is impossible within the scope of a single dissertation to herd together every interesting theoretical construct explaining relationships in the world. Explaining the finer-grained semantic roles of individual items within relationship instances will be the subject of future research. Now it is critical to leave the theoretical territory of this dissertation and begin testing in a computational environment the types and subtypes of relationships and the interesting *image* *description constructions* developed here and determine to what extent they provide guidance for structuring visual information germane to Library and Information Science.

6.9 FUTURE LINES OF RESEARCH

The relationship problem is not new to LIS. However, the line of inquiry applied here sheds new light on the problem by engaging content analysis techniques and ontological tools and methods. It is a first pass or first approximation of this problem in library and information science that engages content analysis and ontological tools and methods. This approach has consequences for relationship studies in LIS with long-term interests that span across many research disciplines concerned with ontology-based representation of visual information and natural language processing.

The goal is not to find a perfect ontology, but to consider many potential ontologies concerning different domains and capturing different views of the same reality. The ontology of image description relationships put forth in this study are a mental product of this researcher's empirical observations. Other researchers may decide that the proposed ontology holds a shared view of reality, or that the conceptualizations are not shared. One could attempt negotiating different meanings from the same semantic units captured in the Corpus of Relationships, or begin with other sets of data based on samples taken from the domains of, for instance, visual ethnography, art history and criticism, or from practicing photographers in fine art photography, photo journalism, and documentary photography to name a few. It is likely that different types of relationships will be discovered in these other visual information contexts. The same is true of relationship studies that might focus on other, unrelated domains of interest, including sound, motion pictures, and bibliographic entities.

Additional, in-depth studies are needed for the individual relationship types that emerged in this study and these should be grounded within the domain of Library and Information Science if practical solutions are to be found for improving description, organization, and retrieval in library information systems using ontology-based representations. For example, in the area of attribution relationships where the format_of relationship was considered semantically overloaded, studies are needed on specific types of format relationships that emerge in visual information.

It is not clear from the brief application of *situation-templates* and *schemas* used for explaining events what they amount to beyond specifying and structuring primitive meanings of event participants. They do not, for instance, account for the manner in which activities are carried out and they do not place constraints on what event relations might be taken to mean. The Corpus suggests that the meaning of the event in the image description domain can be explained by the sum of its parts, so further study is needed to examine the characteristics of *templates* and the role of *schemas* in structuring relationships. Event analysis ought to be tied to a system of semantic frame analysis. Perhaps some of the problems in determining case relations would disappear.

The notion that a relationship derives its meaning from the concepts it joins together is a new puzzle for LIS to explore. From the standpoint of knowledge-based systems, the proposed ontology of relationships is only a partial representation of the whole conceptual framework needed for establishing a knowledge base. What is needed is an ontology of the conceptual entities that characterize the domain of image descriptions.

Finally, the ultimate goal is to understand how the rich set of relationships that emerged in this study can benefit image descriptions in computational environments. This investigation intends to move forward with a systematic analysis of the relationship types specified in Scone representation language. There is a need to know how Scone's inference engine reasons over image descriptions in a knowledge base environment.

APPENDIX A

RELATIONSHIP CORPUS: CODING INSTRUCTION BOOK AND FORM

The *Relationship Corpus Coding Instruction Book* sets out key criteria for identifying, capturing, and recording relations that occur in textual descriptions of visual information. The object of this coding process is to quantify relationship types and categories of relationships in English reported to hold between concepts in written image queries and responses and in any corresponding photograph card-catalog descriptions. There are several variables being analyzed and recorded. Data quality plays a key role in supporting subsequent ontological analysis, so it is critical for coded image descriptions to be collected in a complete, consistent, and accurate manner.

Coder's qualifications

Recommended training for coders consists of the head coder going through this instruction book with all other coders line-by-line explaining each variable and its appropriate response patterns. Trainee should complete one sample incident and then review with head coder any issues that arise. When the initial training is complete, perform a test-retest on appropriate random sample to determine stability and inter-coding reliability.

Syntax and semantics

The coding book explains syntax and semantics of the data language, including cognitive procedures coders must apply in order to code image descriptions reliably. The desired format of the 'response pattern' is given in the coding tables and the definitions of variables are provided in the instruction book.

Nature and administration of records to be produced.

Obtain written permission from appropriate parties to make copies of and analyze samples. Scan the original samples and save to disk as unmarked copies and then printout for coding. Organize and record data collected from the samples in digital spreadsheets. Blacken out personal names in the hardcopy samples used for analysis. Retain digital and paper backup copies of all samples and spreadsheets.

Column Number	Question Number	Variable Name	Response Pattern	Code
1	1	Record Number (RECNO)	First sample record number	001

			Second sample record number	002
			Third sample record number	003
2	2	Incident Type (INTYP) **Choose One**	Can't determine	0
			Unilateral	1
			Bilateral	2
3	3	Total number of Semantic Units (SEMUN)	Enter the total number of Semantic Units indicated in the incident. (Format this spreadsheet cell for numbers as opposed to text.)	0,…,n
4	4	Semantic Unit Number (SEMNL#)	Enter Semantic Unit number. For example, **001SEMNL07** would be interpreted as semantic unit *number 07, Incident 001*.	Use format <RECNO>SEMNL<01,…,n)
5	4	Natural language semantic units (SEMNL)	Enter one natural language Semantic Unit per row. This will consist of direct quotes from the incident record. Write "can't determine" if unable to identify Semantic Unit(s).	Sentence, clause, or word

Column Number	Question Number	Variable Name	Response Pattern	Code
6	5	Total number of factual assertions or statements (FACTN)	Enter the total number of factual assertions. (Format this spreadsheet cell for numbers as opposed to text.)	0,…,n
7	6	Factual assertion statement number (FACTS#)	Enter incident number <RECNO> followed by the variable name FACTS, and then enter the 4-digit assertion number. The first two digits of the FACTS# should match its corresponding Semantic Unit in Column 5. For example, FACTS# **003FACTS0204** would be interpreted as Factual Assertion *04, corresponding to Semantic unit, SEMNL02 in Incident 003.*	Use format <RECNO>FACTS<FACTS#>
8	6	Enter all assertions that apply.(FACTS)	Enter one factual assertion expressing a single basic structural unit per row.	Sentence, clause, or word
9	7	Recording Unit Type (RECTP) **Choose one**	Can't determine	0
			Searcher's Query	1
			Curator or Reference Librarian Correspondence	2
			Cataloger's Description	3

Column Number	Question Number	Variable Name	Response Pattern	Code
10	8	What relationship type is expressed? (RELTP) **Choose one**	Can't determine	0
			Attribution	1
			Case	2
			Inclusion	3
			Meronymy	4
			Spatial	5
			Synonym	6
			Temporal	7
11	9	The Concepts: Argument 1 (the subject that is prehending) (ARGU1) **Enter One.**	Does not apply	0
			Record the noun or noun-phrase	Noun or noun-phrase
12	10	The Concepts: Argument 2 (the datum that is prehended) (ARGU2) **Enter One**	Does not apply	0
			Record the noun or noun-phrase	Noun or noun-phrase
13	11	The Concepts: Argument 3 in tryadic predicate. (ARGU3).	Does not apply	0
			Record the noun or noun-phrase	Noun or noun-phrase

Column Number	**Question Number**	**Variable Name**	**Response Pattern**	**Code**
14	12	The Realtion, the subject form in a prehension (RELAT). **Enter One**	Record the predicate	Verb or verb-phrase
15	13	If Case relationship, specify category (CASER). **Choose one**	Does not apply or can’t determine	0
			VS-Agent-Action	1
			VS-Agent-Instrument	2
			VS-Agent-Object	3
			VS-Action-Recipient	4
			VS-Action-Instrument	5
16	14	If Inclusion relationship, specify category (INCLR). **Choose one**	Does not apply or can’t determine	0
			I5-Functional-Inclusion	1
			I5-State-Inclusion	2
			I5-Activty-Inclusion	3
			I5-Action-Inclusion	4
			I5-Perceptual-Inclusion	5
17	15	If Meronymic relationship, specify category (MEROR). **Choose one**	Does not apply or can’t determine	0
			WCH-Component/Integral Object	1
			WCH-Member/Collection	2
			WCH-Portion/Mass	3
			WCH-Stuff/Object	4
			WCH-Feature/Activity	5
			WCH-Place/Area	6

Column Number	Question Number	Variable Name	Response Pattern	Code
18	16	If Temporal relationship, specify category (TEMPR). **Choose one**	Does not apply or can’t determine	0
			Time-point	1
			Time-Interval	2
19	17	If Spatial relationship, specify category (SPATL). **Choose one**	Does not apply or can’t determine	0
			Static locative	1
			Directional locative	2
			Both static and directional	3
			Preposition followed by another preposition	4
			Ambiguous	5
			Preposition not followed by an argument	6
			Spatial expression applied in other semantic field	7
20	18	Notes Field (NOTES)	Enter notes	

BIBLIOGRAPHY

Adams, A. (1983) *The Print*. Boston: Little, Brown.

Allen, J. F. (1984). “Towards A General Theory Of Action And Time.” *Artificial Intelligence, 23,* 123-154.

Allen, J. F., & Hayes, P. J. (1985). A common-sense theory of time. *Proceedings AAAI-85,* 528-531.

Anelo, C. G. (2003-2004). The changing role of ‘support’ and ‘contiguity’. The hidden facet of the preposition “on’ in Old English. *Journal of English Studies*, *4*, 89-105.

Anelo, C. G. (2005). Some fundamental issues in the semantic analysis of prepositions. *Estudios Ingleses de la Universidad Complutense, 13*, 5-21

Armitage, L. H., & Enser, G. B. Analysis of User Need in Image Archives. *Journal of Information Science, 23*(4), 287-299.

Armstrong, D.M. (1978). *Universals and scientific realism*, Vol. II, A theory of universals. New York, NY: Cambridge University Press.

Back, A. (2000). *Aristotle’s theory of predication*. Leiden; Boston: Brill.

Badler, Norman. (1975). The conceptual description of physical activities. *American Journal of Computational Linguistics*, Microfiche 35, 70-83.

Barrett, T. (1986). A Theoretical Construct for Interpreting Photographs. *Studies in Art Education, 27*(2), 52-60,

Barthes, R. (1977). *Image/Music/Text*. Trans. Stephan Heath. New York: Hill and Wang.

Barthes, R. (1981). *Camera lucida: reflections on photography*. New York, NY: Hill and Wang.

Bartlett, Sir F. (1932). *Remembering: A study in experimental and social psychology*. Cambridge: The University Press.

Batchen, G. (1997). *Burning with desire: the conception of photography*. Cambridge, Mass.: The MIT Press.

Batchen, G. (2001). *Each wild idea: writing, photography, history.* Cambridge, Mass.: The MIT Press.

Bauer, M. & Aarts, B. (2000). Corpus construction: a principle for qualitative data collection. In M. Bauer & G. Gaskell (Eds.), *Qualitative researching with text, image and sound* (19-37). London: Sage.

Bean, C. A. (1996). Analysis of non-hierarchical associative relationships among medical subject headings (MeSH): Anatomical and related terminology. *Advances in Knowledge Organization, 5,* 80-86.

Bean, C. A., & Green, R. (Ed.). (2001). *Relationships in the Organization of Knowledge*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Bean, C. A., & Molholt, P. (1996). A relational model of clinical anatomy: Symbolic knowledge base design. Submitted to *1996 AMIA Annual Fall Symposium, Washington, DC, October 1996.*

Beaudoin, J. E. (2009). *An investigation of image users across professions: A framework of their image needs, retrieval and use*. Doctoral Dissertation, Drexel University.

Bennett, D. C. (1975). *Spatial and temporal uses of English prepositions: An essay in stratificational semantics.* London, England: Langman Group.

Benson, A. C. (2009). The archival photograph and its meaning: formalisms for modeling images. *Journal of Archival Organization, 7*(4), 148-187.

Benson, A. C. (2010). Killed negatives: The unseen photographic archives. *Archivaria, 68,* 1-37.

Benson, A. C. (2011a). Capturing relationship in image descriptions: a preliminary investigation. *BiD: textos universitaris de biblioteconomia I documentació,* University of Barcelona (Under review).

Benson, A. C. (2011b). *OntoPhoto* and the role of ontology in organizing knowledge. *Knowledge Organization,* 38(2).

Biber, D. (1993). Representativeness in corpus design. *Literary and Linguistic Computing, 8*(4), 243-257.

Kreyer, R. (2003). Genitive and *of*-constructions in modern written English. Processability and human involvement. *International Journal of Corpus Linguistics, 8*(2), 169-207.

Bowker, G. C., & Star, S. L. (1999). *Sorting things out: classification and its consequences*. Cambridge Mass.: MIT Press.

Brookes, B. C. (1986). Jason Farradane and relational indexing. *Journal of Information Science, 12*,15-18.

Bunge, M. (1977). *The furniture of the world*. Dordrecht, Holland: D. Reidel Publishing.

Byrd, R. (1989). Discovering relationships among word senses. Proceeding of the Fifth Annual Conference of the University of Waterloo Centre for the New Oxford English Dictionary, Waterloo, Canada, 67-79.

Byrd, R. J., Calzolari, N., Chodorow, M., Klavans, J., Neff, M., & Rizk, O. (1987). Tools and methods for computational lexicology. *Computational Linguistics, 13*, 219-240.

Carnegie Library of Pittsburgh, Pittsburgh Photographic Library (PPL) Preliminary Inventory. (2000).

Casagrande, J. B., & Hale, K. (1967). Semantic relationships in Papago folk-definitions. In D. H. Hymes & W. E. Bittle (Eds.), *Studies in Southwestern Ethnolinguistics* (165-193). The Hague, The Netherlands: Mouton.

Chaffin, R., & Herrmann, D. J. (1984). The similarity and diversity of semantic relations. *Memory & Cognition, 12*(2), 134-141.

Chaffin, R., & Herrmann, D. J. (1987). Relation element theory: A new account of the representation and processing of semantic relations. In D. S. Gorfein & R. R. Hoffman (Eds.), *Memory and learning* (221-245). Hillsdale, NJ: Lawrence Erlbaum Associates.

Clarke, G. (1997). *The photograph.* New York: Oxford University Press.

Coles, L. S. 1968. An on-line question-answering system with natural language and pictorial input," in Proceedings -- -WM National Conference (157-167).

Collins, K. (1998). Providing subject access to images: a study of user queries. *The American Archivist, 61*(1), 36-55.

Cook, T.D., & Campbell, D.T. (1979). *Quasi-experimentation: design and analysis for field settings*. Chicago, IL: Rand McNally.

Cooper, G. S. (1968). *A semantic analysis of English locative prepositions* (Bolt, Beranek, & Newman, Report No. 1587). Springfield, VA: Clearing House for Federal, Scientific, and Technical Information. Available http://handle.dtic.mil/100.2/AD666444

Crombie, W. (1985). *Process and relation in discourse and language learning*. New York, NY: Oxford University Press.

Cruse, D. A. (1979). On the transitivity of the part-whole relation. *Journal of Linguistics, 15*, 29-38.

Denofsky, Murray E., How Near is Near?: a "near' specialist, AX Memo No. 344, Cambridge: Artificial Intelligence Laboratory Massachusetts Institute of Technology, February 1976.

Denzin, N. K., & Lincoln, Y. S. (Eds.). (1994). *Handbook of qualitative research*. New York, NY: Sage Publications.

Derrida, J. (2010). *Copy, archive, signature: a conversation on photography*. Stanford, CA: Stanford University Press.

Dick, J. P. (1991). *A conceptual case-relation representation of text for information retrieval*. Doctoral Dissertation, Computer Systems Research Institute, University of Toronto.

Dorr, C. (2005). What we disagree about when we disagree about ontology. In M. Kalderon (Ed.), *Fictionalism in Metaphysics* (234-286). Oxford: Oxford University Press.

Dunn, J. M. (1990). Relevant predication 2: Intrinsic properties and internal relations. *Philosophical Studies, 60,* 177-206.

Ellenbogen, J. (2008). Camera and mind. *Representations, 101*, 86-115.

Technology Applications Programme, JTAP Reports, University of Northumbria at Newcastle, pp. 1-67.

Enser, P.G.B. (1993). Query analysis in a visual information retrieval context. *Journal of Document and Text Management, 1*(1), 25-52.

Enser, P.G.B. (1995). Progress in documentation: Pictorial information retrieval. *Journal of Documentation*, *51*(2), 126-170.

Enser, Peter. (2000). Visual image retrieval: seeking the alliance of concept-based and content-based paradigms. *Journal of Information Science, 26*(4),199-210.

Enser, P.G.B. (2008). Visual image retrieval. In B. Cronin (Ed.), *Annual Review of Information Science and Technology* (pp.3-42). Medford, NY: Information Today.

Enser, P. G. B., & McGregor, C. G. (1993). *Analysis of visual information retrieval queries.* British Library R&D Report 6104. The British Library Board.

Enser, P.G.B., Snadom, C. J., & Lewis, P.H. (2005). Surveying the reality of semantic image retrieval. In S. Bres & R. Laurini (Eds.), *Visual information and information systems: 8th International Conference, revised selected papers* (Lecture Notes in Computer Science, 3736) (pp. 177-188). Berlin: Springer-Verlag.

Evans, M. (2002). Thesaural relations in information retrieval. In R. Green, C. A. Bean, & S. H. Myaeng (Eds.), *The semantics of relationships: an interdisciplinary perspective*. Dordrecht: Kluwer Academic Publishers.

Fahlman, S. E. (n.d.). The Scone Knowledge-Base Project. Retrieved on September 24, 2010: http://www.cs.cmu.edu/~sef/scone/.

Fahlman, S. E. (1979). *NETL: A System for Representing and Using Real-World Knowledge*. Cambridge, Massachusetts: MIT Press.

Fahlman, S. E. (1982.). Common Lisp. *Annual Review of Computer Science, 2*, 1-19.

Fahlman, S. E. (2006). Marker-passing inference in the Scone knowledge-base system. In Lang, J., et al., eds., *KSEM 2006, LNAI 4092*, 114-126.

Fahlman, S. E. (2008a). In defense of incomplete inference. In *Knowledge Nuggets*. Retrieved on August 10, 2009: http://sef-linux.radar.cs.cmu.edu/nuggets/?m=200807

Fahlman, S. E. (2008b). Mini-Nuggets: Natural Language Understanding. In *Knowledge Nuggets: Scott Fahlman's Notes on AI and Knowledge Representation*. Retrieved on March 20, 2011: http://scone1.scone.cs.cmu.edu/nuggets/?p=31

Farradane, J. (1980a). Relational indexing, Part I. *Journal of Information Science, 1*, 267-276.

Farradane, J. (1980b). Relational indexing, Part II. *Journal of Information Science, 1*, 313-324.

Fidel, R., Hahn, T. B., Rasmussen, E.M., & Smith, P.J. (Eds.). (1994). *Challenges in indexing electronic text and images.* Medford, NJ: Learned Information.

Fillmore, C. J. (1968). The case for case. In E. Bach, R. T. Harms (Eds.), *Universals in Linguistic Theory*, 1-88. New York, NY: Holt, Rinehart and Winston.

Fillmore, C. J. (1976). Frame semantics and the nature of language. *Annals of the New York Academy of Sciences: Conference on the Origin and Development of Language and Speech*, *280*, 20-32.

Fillmore, C. J. (1977). Scenes-and-frames semantics. In A. Zampolli (Ed.), *Linguistic structures processing*, 55-81. New York, NY: North-Holland Publishing.

Flavell, J. H., & Flavell, E. R. (1959). One determinant of judged semantic and associative connections between words. *Journal of Experimental Psychology, 58*, 159-165.

Flusser, V. (1983). *Towards a philosophy of photography.* London: Reaktion Books.

Fox, E. A. (1988). Improved retrieval using a relational thesaurus for automatic expansion of extended Boolean logic queries. In M. Evens (Ed.), *Relational Models of the Lexicon*, 199-210. Cambridge, UK: Cambridge University Press.

Forbus, K., Nielsen, P., and Faltings, B. 1991. Qualitative spatial reasoning: the CLOCK project. *Artificial Intelligence, 51*, 1-3.

Frake, C. O. (2009). Notes on queries in ethnography. *American Anthropologist, 66* (3), 132-145.

Gehrke, B. (2006). On directional readings of locative prepositions, *Proceedings of ConSOLE XIV*.

Gene Ontology. (2010). Extended GO ontology relations. Retrtieved on 1 November 2010: http://www.geneontology.org/GO.ontology-ext.relations.shtml#relations

Gibbs-Smith, C. H. (1950). The Hulton Picture Post Library. *Journal of Documentation*, *6* (1), 12-24.

Green, R. (1996). Development of a relational thesaurus. *Knowledge Organization and Change: Proceedings of the 4th International ISKO Conference, 15-18 July 1996, Washington, DC*, Rebecca Green, ed. Frankfurt/Main: INDEKS Verlag.

Graesser, A. C., & Goodman, S. M. (1985). Implicit knowledge, question answering, and the representation of expository text. In B. K. Britton & J. B. Black (Eds.), *Understanding expository text: a theoretical and practical handbook for analyzing explanatory text*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Greisdorf, H. (2000). *Relevance thresholds: a conjunctive/disjunctive model of end-user cognition as an evaluative process*. Doctoral Dissertation, University of North Texas, Denton, TX.

Greisdorf, H., & O'Connor, B. (2002). Modeling what users see when they look at images: a cognitive viewpoint. *Journal of Documentation, 58*(1), 6-29.

Grice, H. P. (1957). Meaning. *Philosophical Review*, *66*, 377-8.

Grice, H. P. (1975). Logic and Conversation. In P. Cole and J.L. Morgan (Eds.), *Syntax and Semantics*, Vol. 3. (41-58). New York: Academic Press.

Griffith, R. L. Three principles of representation for semantic networks, *ACM Transactions on Database Systems*, *7*(3), 417-442.

Gruber, J. S. (1965). Studies in lexical relations. Doctoral Dissertation, MIT, Cambridge; Indiana University Linguistics Club, Bloomington, Ind. Reprinted as part of *Lexical structures in syntax and semantics*, North-Holland, Amsterdam, 1976.

Guarino, N. (1992). Concepts, Attributes, and Arbitrary Relations: Some Linguistic and Ontological Criteria for Structuring Knowledge Bases. *Data and Knowledge Engineering*, *8*(3), 249-261.

Guizzardi, G. (2005). *Ontological foundations for structural conceptual models*. Doctoral Dissertation, Telematica Instituut, Enschede, The Netherlands. Retrieved from http://doc.utwente.nl/50826/1/thesis_Guizzardi.pdf

Habel, C. & Eschenbach, C. (1997). Abstract structures in spatial cognition. *Foundations of Computer Science*, *1337*, 369-378.

Harry Ransom Center. 2007. The University of Texas at Austin. Exhibitions: The First Photograph. Retrieved on 12 December 2009: http://www.hrc.utexas.edu/exhibitions/permanent/wfp/

Herskovits, A. (1986). *Language and spatial cognition: an interdisciplinary study of the prepositions in English*. New York, NY: Cambridge University Press.

Hobbs, J. R., & Moore, R. C. (Eds.). (1985). *Formal theories of the commonsense world.* Norwood, NJ: Ablex Publishing.

Hood, P. M. (2004). *Aristotle on the category of relation*. New York, NY: University Press of America.

Huhns, M. N. and Stephens, L. M. (1989).Plausible inferencing using extended composition. In *IJCAI'89: Proceedings of the 11th international joint conference on AI*, Sanfrancisco, CA.

Hying, C. (2009). Semantics of projective locative expressions: an empirical evaluation of geometrical conditions. Doctoral Dissertation, Institut für maschinelle Sprachverarbeitung Universität Stuttgart.

J. Paul Getty Trust (2000). *Art and Architecture Thesaurus (AAT).* Retrieved on 22 August 2010: http://www.getty.edu/research/conducting_research/vocabularies/aat/

Jackendoff, R. (1983). *Semantics and cognition.* Cambridge, Mass.: MIT Press.

Jackendoff, R., & Landau, B. (1992). Spatial language and spatial cognition. In R. Jackendoff, *Languages of the mind*, (99-124). Cambridge, Mass.: MIT Press.

Jackendoff, R. (1996). The architecture of the linguistic-spatial interface. In P. Bloom, M. A. Peterson, L. Nadel, & M. F. Garrett (Eds.), *Language and space* (1-30). Cambridge, MA: MIT Press.

Jaimes, A., & Chang, S-F. (2000). A conceptual framework for indexing visual information at multiple levels. *IS&T/SPIE Internet Imaging*, 3964, (San Jose, CA).

Jones, Karen Spärck. (2004). A statistical interpretation of term specificity and its application in retrieval. *Journal of Documentation, 60* (5), 493-502.

Jörgensen, C. (1995). Image attributes: an investigation. Doctoral Dissertation, Syracuse University.

Jörgensen, C. (1998) Attributes of Images in Describing Tasks. *Information Processing & Management, 34* (2/3), 161-174.

Jörgensen, C. (2003). *Image Retrieval: Theory and Research*. Lanham, MD: Scarecrow Press.

Jörgensen, C., Jaimes, A., Benitez, A. B., & Chang, S-F. (2001). A conceptual framework and empirical research for classifying visual descriptors. *Journal of the American Society for Information Science and Technology*, *52*(11), 938-947.

Jörgensen, C. & Jörgensen, P. (2005). Image querying by image professionals. *Journal of the American Society for Information Science and Technology, 52*(12), 1346-1359.

Keister, L. H. (1994) User types and queries: Impact on image access systems. In R. Fidel et al. (Eds.), *Challenges in indexing electronic text and images*, (7-22). Medford, NJ: Learned Information.

Klijn, E. (Ed.). (2003). *SEPIADES: Recommendations for cataloguing photographic collections*. Amsterdam, The Netherlands: European Commission on Preservation and Access.

Kochen, Manfred. (1969). Automatic Question-Answering of English-Like Questions About Simple Diagrams. *Journal of the ACM*, *16*(1), 26-48.

Kosslyn, S. M., Chabris, C. F., Marsolek, C. J., & Koenig, O. (1992). Categorical versus coordinate spatial relations: computational analyses and computer simulations. *Journal of Experimental Psychology: Human Perception and Performance*, *18* (2), 562-577.

Krause, M. (1988). Intellectual problems of indexing picture collections. *Audiovisual Librarian, 14*(2), 73-81.

Krippendorff, K. (2004). *Content analysis: an introduction to its methodology*. Thousand Oaks, CA: Sage Publications.

Kuntz, J. (1987). Some remarks on case relations. Proceeding EACL ’87 Proceedings of the third conference on European chapter of the Association for Computational Linguistics.

Langer, S. K. (1957). *Philosophy in a new key*. Cambridge, Mass.: Harvard University Press.

Larson, M. L. (1984). *Meaning-based translation: A guide to cross-language equivalence.* Lanham, MD: University Press of America

Levesque, H. J. & Lakemeyer, G. (2000). *The logic of knowledge bases*. Cambridge, Mass.: MIT Press.

Levinson, S. C. (1983). *Pragmatics.* Cambridge, England: Cambridge University Press.

Levinson, S. C. (1996). Language and space. *Annual Review of Anthropology, 25*, 353-382.

Li, P., & Gleitman, L. (2002). Turning the tables: language and spatial reasoning. Cognition, 83, 265-294.

Library of Congress. Prints and Photographs Division. (1995). Thesaurus for Graphic Materials. Washington, D.C.: Library of Congress, Cataloging Distribution Service.

Lockwood, K. (2009). Using Analogy to Model Spatial Language Use and Multimodal Knowledge Capture. Doctoral Dissertation, Northwestern University.

Longacre, R. E. (1983).*The grammar of discourse.* New York: Plenum.

Lorant, S. (1964). Pittsburgh; the story of an American city. Garden City, NY: Doubleday, 1964.

Lyons, J. (1977). Semantics. London, England: Cambridge University Press.

Mann, W. C., & Thompson, S. A. (1985).*Assertions from discourse structure.* Marina del Rey, CA: Information Sciences Institute.

Margolis, E. & Laurence, S. (Eds.). (1999). Concepts: core readings. Cambridge, Mass.: MIT Press.

Markey, K. (1983). Computer-assisted construction of a thematic catalog of primary and secondary subject matter. *Visual resources III*, 16-49.

Mealy, G. H. (1967). Another look at data. Proceedings of the Fall Joint Computer Conference, November 14-16, Anaheim, California (AFIPS Conference Proceedings, Volume 31, 525-534). Washington, DC: Thompson Books, London: Academic Press.

Michael, E. (1974). Peirce's early study of the logic of relations, 1865-1867. *Transactions of the Charles S. Peirce Soceity, 10* (2), 63-75.

Miller, D. R. (2007). Principia bibliographica? Balancing principles, practice, and pragmatics in a changing digital environment. In R. P. Holley (Ed.), *Cataloger, editor, and scholar: essays in honor of Ruth C. Carter* (pp. 281-305). Binghamton, NY: Haworth Information Press.

Miller, G. A. (1969). The organization of lexical memory: Are word associations sufficient? In G. A. Talland & N. C. Waugh (Eds.), *The pathology of memory*. New York: Academic Press.

Minsky, M. (1974). "A framework for representing knowledge." Artificial Intelligence Memo No. 306. Cambridge, MA: Massachusetts Institute of Technology A.I. Laboratory.

Mitchell, W.J. (1992). *The reconfigured eye: Visual truth in the post-photographic era.* Cambridge, MA: MIT Press.

Mitchell, W.J.T. (1994). *Picture Theory.* Chicago: The University of Chicago Press.

Mitchell, W.J.T. (2005). *What do images want?: the lives and loves of images.* Chicago: University of Chicago Press.

Moholy-Nagy, L. (1991). *Moholy-Nagy: an anthology*. New York, NY: Da Capo Press.

Molholt, P. (1996). Standardization of inter-concept links and their usage. *Knowledge Organization and Change: Proceedings of the 4th International ISKO Conference, 15-18 July 1996, Washington, DC,* Rebecca Green, ed. Frankfurt/Main: INDEKS Verlag.

Moravcsik, J. M. (1991). What makes reality intelligible? Reflections on Aristotle's theory of *aitia*. In L. Judson (Ed.), *Aristotle's physics: A collection of essays*, 31-47. Clarendon Press.

Murphy, L. (2003). *Semantic relations and the lexicon*. New York, NY: Cambridge University Press.

Neelameghan, A., & Maitra, R. (1978). Non-hierarchical associative relationships among concepts: Identification and typology. *International Federation for Documentation Committee on Classification Research Report No. 18*, FID Publ. No. 405, Bangalore, India: FID/CR Secretariat, 25 pp.

Newcombe, N. S., & Huttenlocher, J. (2000). *Making space: the development of spatial representation and reasoning.* Cambridge, Mass.: MIT Press.

Newhall, B. (1982). *The history of photography*. Boston, Mass: Little, Brown and Company.

Nolt, J., Rohatyn, D., Varzi, A. (1998). *Schaum's outline of theory and problems of logic*. New York, NY: McGraw-Hill.

Oliver, P., & Gapp, K-P. (Eds.). (1998). *Representation and processing of spatial expressions*. Mahwah, NJ: Lawrence Erlbaum Associates.

Oxford English Dictionary Online. March 2011. Oxford University Press. Retrieved on April 28, 2011: http://www.oed.com/view/Entry/41873?redirectedFrom=corpus

Panofsky, E. (1939). *Studies in iconology: humanistic themes in the art of the Renaissance*. New York, NY: Oxford University Press.

Parkinson, G. H. R. (1965). Logic and Reality in Leibniz's Metaphysics. Oxford: Clarendon Press.

Perakath C. B. et al. (1994). *IDEF5 Method Report*. Knowledge Based Systems, Inc.

Perfetti, C. A. (1967). A study of denotative similarity with restricted word associations. *Journal of Verbal Learning and Verbal Behavior, 6,* 788-795.

Pitt Plans Photograph Library. (June 18, 1950). *Pittsburgh Post-Gazette*.

Pittsburgh Photographic Library (PPL) Preliminary Inventory, Revision 19. (March 23, 2000). Retrieved on April 12, 2009: http://216.183.184.20/exhibit/ppl_plan.html

Pollack, P. (1969). The picture history of photography from the earliest beginnings to the present day. New York, NY: Harry N. Abrams.

Preston, D. R. (2004). Folk metalanguage. In A. Jaworski, N. Coupland, & D. Galasinski (Eds.), *Metalanguage: Social and ideological perspectives* (75-101). Berlin: Walter de Gruyter GmbH & Co.

Proposal for the Establishment of the Civic Photographic Center under the Sponsorship of the Allegheny Conference on Community Development. (December 1949, revised February 20, 1950). Archives Service Center, University of Pittsburgh, Box 56, Folder File 1, A. W. Mellon Education and Charitable Trust, Pgh., PA Records 1930-1980.

Pustejovsky, J. (1996). *The generative lexicon*. Cambridge, MA: MIT Press.

Riegel, K. F., & Riegel, R. M. (1963). An investigation into denotative aspects of word meaning. *Language and Speech, 6*, 5-21.

Ritchen, F. (1990). *In our own image: the coming revolution in photography*. New York: Aperture.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology, 104*, 192-233.

Rothkegel, R., Wender, K. F., & Schumacher, S. (1998). Judging spatial relations from memory. In C. Freska, C. Habel, & K.F. Wender (Eds.), *Spatial cognition: An interdisciplinary approach to representation and processing of spatial knowledge* (p. 79-105). Berlin: Springer-Verlag.

Russell, L. (1938c). *Removing packages of rice from conveyer to place them in fiber cartons. State rice mill, Abbeville, Louisiana.* (Library of Congress, Prints & Photographs Division, FSA/OWI Collection, [LC-USF34-031383-D (b&w film neg.)] Retrieved on 15 January 2010: http://www.loc.gov/pictures/collection/fsa/item/fsa2000012549/PP/

Saint-Dizier, P. (Ed.). (2006). *Syntax and semantics of prepositions.* Dordrecht

Saussure, Ferdinand de. (1966). *Course in General Linguistics*. C. Bally & A. Sechehaye, in collaboration with A. Riedlinger (Eds.). Translated by W. Baskin. New York, NY: McGraw-Hill Book Company.

Schmied, J. (1990). Corpus linguistics and non-native varieties of English. *World Englishes, 9*(3), p. 255-268.

Schulz, C. B., & Plattner, S. W. (Eds.). (1999). *Witness to the Fifties: The Pittsburgh Photographic Library, 1950-53*. Pittsburgh, PA: University of Pittsburgh Press.

Shatford, S. (1986). Analyzing the subject of a picture: A theoretical approach. *Cataloging & Classification Quarterly*, *6* (3), 39-62.

Shatford Layne, S. (1994). Some issues in the indexing of images. *Journal of the American Society for Information Science, 45*(8), 583-586.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin, 86*(2), 420-428.

Sim, J., & Wright, C. C. (2005). The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Physical Therapy 85* (3), 257-268.

Smith, B. (2004). Knowledge Representation Hypothesis. In R. J. Brachman & H. J. Levesque, *Knowledge Representation and Reasoning* (6). San Francisco, CA: Morgan Kaufmann Publishers.

Smith, B., Ceusters, W., Klagges, B., Köhler, Kumar, A., Lomax, J., Mungall, C., Neuhaus, F., Rector, A. L., & Rosse, C. (2005). Relations in biomedical ontologies. *Genome Biology*, *6*(5), R46.

Somers, H. L. (1987). Valency and case in computational linguistics. Edinburgh: Edinburgh University Press.

Sowa, J. F. (2000). *Knowledge representation: logical, philosophical, and computational foundations*. Pacific Grove, CA: Brooks/Cole.

"Speaking of Pictures." (n.d.). *Life*, pp. 22–24. Archives Service Center, University of Pittsburgh, Box 56, Folder File 1, A.W. Mellon Education and Charitable Trust, Pgh., PA Records 1930–1980.

Stasio, T., Herrmann, D. J., & Chaffin, R. (1985). Relation similarity as a function of agreement between relation elements. *Bulletin of the Psychonomatic Society, 23* (1), 5-8.

Stemler, S. (2001). An overview of content analysis. *Practical Assessment, Research & Evaluation*, *7*(17). Retrieved on January 15, 2010: http://PAREonline.net/getvn.asp?v=7&n=1

Story, V. C. (1993). Understanding semantic relationships. *VLDB Journal, 2*, 455-488.

Stryker, R., & Seidenberg, M. (1959). *A Pittsburgh album.* Pittsburgh, PA: Pittsburgh Post-Gazette.

Svenonius, E. (1994). Access to nonbook materials: the limits of subject indexing for visual and aural languages. *Journal of the American Society for Information Science, 45*(8), 600-606.

Svenonius, E. (2000). *The intellectual foundation of information organization*. Cambridge, Mass.: MIT Press.

Szarkowski, J. (1989). *Photography until now*. New York, NY: Museum of Modern Art.

Talbot, William Henry Fox. (1968). *The pencil of nature*. New York, NY: Da Capo Press. (First edition of *The Pencil of Nature* was published in London between 1844 and 1846 in six separate fascicles.)

Taylor, L. (Ed.). (1994). *Visualizing theory: selected essays from V.A.R., 1990-1994.* New York, NY: Routledge.

Thompson, L. (1959, November 22). Photographer's champion. *The Pittsburgh Press*, 32, 34-35.

Tribble, Alice. (2010). Textual inference for retrieving labeled object descriptions. Doctoral Dissertation, Carnegie Mellon University.

Wand, Y., Storey, V., & Weber, R. (1999). An ontological analysis of the relationship construct in conceptual modeling. *ACM Transactions on Database Systems, 24* (4), 494-528.

Weber, R. P. (1990). *Basic content analysis*. Newbury Park, CA: Sage Publications.

Whitehead, A. N. (1978). *Process and Reality*. New York, NY: Free Press.

Whorf, B. L. (1940/1956). Science and linguistics. In J. B. Carroll (Ed.), Language, thought, and reality: selected writings of Benjamin Lee Whorf. Cambridge, MA: MIT Press. (Reprinted from Technology Review, 42 (6), 229-231, 247-248.

Winograd, T. (1972). *Understanding natural language*. New York: Academic Press.

Winston, M., Chaffin, R., & Herrmann, D. (1987). A taxonomy of part-whole relations. *Cognitive Science, 11,* 417-444.

Wittgenstein, L. (1992). *Tractatus logico-philosophicus*. New York, NY: Routledge & K. Paul.

Wittgenstein, L. (2001). *Philosophical investigations* (G. E. M. Anscombe, Trans.). Oxford: Blackwell. (Original work published 1953)

Wood, J. M. (2007). Understanding and computing Cohen's Kappa: a tutorial. *WebPsychEmpiricist*. Retrieved on 14 May 2010: http://wpe.info/papers_table.html

Woods, W.A. (1975). What's in a link: foundations for semantic networks. In D. G. Bobrow & A. Collins (Eds.), *Representation and understanding: studies in cognitive science* (35-82). New York, NY: Academic Press.