

# An Evidence Perspective on Topical Relevance Types and Its Implications for Exploratory and Task-based Retrieval

Xiaoli Huang      Dr. Dagobert Soergel  
 College of Information Studies, University of Maryland  
 4105 Hornbake Library, College Park, MD 20742-4345  
 {xiaoli, dsoergel} @umd.edu

**The common view of topical relevance is limited to *topic matching*, resulting in IR systems' failure to detect more complex topical connections which are needed to respond to diversified user situations and tasks. To reveal the complex evidential relationships involved in topical relevance, we analyzed relevance assessments in the domain of history that used four types of topical relevance– direct, indirect, context, and comparison. Each of these plays a special role in reasoning, making a conclusive argument, or performing a task. Incorporating these relevance types into IR systems allows users more flexibility and a better focus on their tasks.**

## Introduction

This paper explicates topical relevance from an evidence perspective, where one asks: How does a piece of information fit into the structure of an argument or the reasoning needed to accomplish a task? The concept of relevance lies at the heart of intellectual access and information retrieval, indeed of reasoning and communication in general; in turn topical relevance lies at the heart of relevance. Although topical relevance is commonly recognized and widely used as an important selection criterion, with few exceptions it is treated as an atomic notion and remains vague and unexplicated.

In contrast, we treat topical relevance as an umbrella concept subsuming more specific topical relevance types. Based on the role a piece of information plays in the overall structure of an argument or in building an understanding of a situation, problem, or issue in the receiver's mind, or the role a piece of information plays in the process of answering a question, deriving a generalization, building an argument, making a decision, applying information to a work task (in the following just "task"), etc., we have identified four topical relevance types (Huang & Soergel, 2004):

- **Direct relevance**  
 Direct evidence for what the user asks for  
*Example: Talks about food available to Auschwitz inmates, given the topic "Food in Auschwitz"*
- **Indirect relevance**  
 From which one can infer something about the topic  
*Example: Talks about seeing emaciated people in Auschwitz*
- **Context relevance**  
 Provides background/context for topic  
*Example: Talks about physical labor of Auschwitz inmates*

- **Comparison relevance**  
Provides information on a similar/contrasting situation  
*Example: Talks about food in the Warsaw ghetto*

Within this general perspective, in this paper we focus on the evidentiary connection between a piece of information and a user's question, topic, or task. Detecting topical relevance is not just matching meaning or matching words, but determining whether the information at hand can serve as evidence for deriving conclusions of concern. The essence of topical relevance is reasoning from evidence (a piece of information) to a conclusion (or an answer to the user's question). The subject of Evidence gains its coherence from "inferential reasoning"; it has been a sustained focus of multi-disciplinary attention and hence been constantly advanced by confluent efforts in, just to name a few, Law, History, Mathematics, Logic, Witness Psychology, Forensic Science, Intelligence Services, Inferential Communication (especially the Relevance Theory of Communication), Evidence-Based Medicine (EBM), and Evidence-Based Policy (Twining, 2003). Types of topical relevance can and should be specialized and adapted to these different domains. Connecting the concept of Topical Relevance with the notion of Evidence not only brings in new thoughts from other disciplines for improving our understanding of topical relevance, but also takes the discussion into the broader context of human thinking, reasoning, drawing conclusions, building arguments, and, most generally, building understanding and deriving meaning.

## Background

It is generally agreed that relevance is a property of the relationship between a piece of information and a user with a given question or task in a given situation. This concept has many dimensions, such as topicality, appropriateness to the user's background, recency, or authority, to name but a few of over 80 relevance criteria (Barry, 1993; Schamber, 1990; Schamber, 1994; Wang & Soergel, 1998; Lawley, *et. al.*, 2005). Saracevic (1975) gives a review and a framework for an analysis of relevance with emphasis on topic relevance; the framework is extended in Saracevic (1996). For full bibliographies of relevance see, for example, Mizzarro (1998) or Saravevic (2006). This paper focuses on topical relevance and its explications; we limit our background discussion to seminal works that make an original contribution to the specific purpose of this paper.

Our work builds on the definitions of logical relevance (Cooper, 1971) and situational relevance (Wilson, 1973). Cooper defined *logical relevance* based on a strict logical deduction relationship between a statement and a sought-after answer. This definition can rarely be applied literally in the practice of IR but in a broadened view it is fundamental for system design and evaluation. Wilson defined *evidential relevance* and *situational relevance*. A piece of information is evidentially relevant if it either increases or decreases the confirmation or probability of a conclusion through deductive or inductive reasoning, including plausible and probabilistic reasoning. Situational relevance is evidential relevance to any question a user is concerned about, a question of importance in the user's situation; all the possible answers to a question constitute one of the user's "concern sets". Furthermore, Wilson defined *Direct Situational Relevance* and *Indirect Situational Relevance*:

If an item of information  $I_j$  is itself a member of a concern set, we shall say that it is directly relevant situationally; if it is relevant but not a member of a concern set, we shall say that it is indirectly relevant situationally.

Our work is related to Green & Bean (1995) who discussed more specific types of semantic relationships between a piece of information and a topic, primarily from a different perspective, so there is little overlap.

Our work is also informed by the literature in the area of evidence as stated above, an aspect we plan to expand on in the future.

## Data and methods

The broad set of topical relevance types listed in the introduction emerged from thinking about how historians would use primary materials. We trained relevance assessors (Graduate students in history) in applying these relevance types to an oral history collection using real user requests. We analyzed the resulting examples from assessors' relevance judgments, using a grounded theory approach (Lindlof & Taylor, 2002) as a starting point for refining the types of relevance and developing subtypes.

Data we collected in the MALACH project (Gustman *et. al.*, 2002), which aims at improving speech retrieval through automatic speech recognition (ASR) and subsequent information retrieval assisted by NLP techniques. MALACH uses the archive of 52,000 Holocaust survivor interviews assembled by the Shoah Visual History Foundation (VHF). 4,000 of these testimonies have been divided into topical *segments*. From these we selected a speech retrieval test collection, a subset of 400 interviews, totaling about 20,000 segments, paired with 75 topics based on real requests to VHF (CLEF-CLSR 2006).

Our assessors made 27,000 topical relevance assessments between interview segments and topics, using the four relevance types, each on a scale of 0 – 4. They recorded justifications for their assessments. They also kept *Topic Notes*: notes on their interpretation of the topics, their rules defining the application of each relevance type to the topic, and typical segment examples (205 Topic Notes since many topics were done by two or more assessors).

We analyzed 46 Topic Notes (selected for length and variety of assessors and topics) as follows:

- The initial coding scheme consisted simply of the four first-level relevance types;
- as a starting point, read through ten Topic Notes to identify second-level topical relevance subtypes and refined the coding scheme accordingly;
- applied the refined coding scheme to the same ten Topic Notes making further refinements;
- looked at all examples under each second-level subtype to derive third-level subtypes;
- applied the resulting coding scheme to all 46 Topic Notes making further refinements.

The result was a fine-grained classification of topical relevance types and associated examples. We added some topical relevance subtypes based on logical analysis and knowledge from other domains even though they did not happen to occur in our data.

## Findings and Discussion

In the following, we give definitions of our four types of topical relevance as refined by the analysis of many examples, and we develop more specific subtypes, followed by examples. With the limited space, we cannot list a full range of examples to manifest all the subtleties involved. The relevance types presented here are not always clearly distinct.

### (1) Direct Relevance

Directly relevant evidence is a direct answer to a question of interest (is a member of the *concern set* as defined by Wilson) and is exactly, explicitly on topic. It is the most straightforward and intuitive relevance type, with minimal, if any, inferential reasoning involved. Good direct evidence has a wealth of specific details about a topic and has the most significant impact on overall topical relevance.

**Topic:** Strengthening Faith by Holocaust Experience

**Evidence:** A survivor talks about how an elderly Salonikan Jew helped strengthen their religious faith during their incarceration; "we called him grandfather. He always said to us 'you must say Kaddish every night.' I was forced to dispose of corpses in the camp at the time. One day I came back from work and said to him 'Are you crazy?' He said: 'No, something good will happen one day after this. We have to pay a very dear price but we're gonna have our own state of Israel.' And it happened. I survived with my faith and went to Israel."

Direct relevance is the most recognized and emphasized topical relevance type in both research and practice. It has become the central meaning of topical relevance and is even mistaken as the only meaning of topical relevance. This misperception restricts the attention to a very narrow focus of topical relevance. The effort we made in this study is to broaden our vision to the often ignored non-direct evidential relevance.

### (2) Indirect Relevance

Making inference about a topic is the central feature of indirect relevance. Indirect evidence, or inferential or circumstantial evidence, is implicit information on a topic. While direct evidence *is* the answer, indirect evidence can be used to *infer* the answer; it is one or more inferential steps away from the answer. After "joining the dots" it contributes as much to understanding a topic as direct evidence. Indirect or circumstantial evidence is often used in court to establish facts. Both direct and indirect evidence are valid for establishing a fact but they may differ in the level of certainty: the inferential relationship between A and B may be subject to uncertainty, that is, given A we can infer that B with a given probability, also known as inferential strength (Kadane & Schum, 1996). To be recognized and further used to draw a conclusion, the inferential strength of indirect evidence needs to be sufficiently high.

Indirect relevance plays an important role in gleaning relevant information from the Holocaust survivor interviews. Survivors usually go into great detail of their personal experience and feelings, without clearly describing or even explicitly mentioning the events or phenomena asked by a topic (request). Direct discussion is rare for many topics, especially for those looking for information on particular phenomena rather than specific events. In these cases, indirect evidence, which is relatively more available, is usually very

helpful for leading us to conclusive points. We found three specific subtypes of indirect evidence:

### ***(2.1) Generic Indirect Relevance***

This category is almost like direct evidence, missing only a specific piece of information. The argument stays implicit and uncertain only because a survivor's description is not precise enough or the survivor forgets to mention a key name. However, everything else said by the survivor strongly points at a fact that is right on topic. It is characterized by high inferential strength.

**Topic:** Stories of Varian Fry and the Emergency Rescue Committee who saved thousands in Marseille

**Evidence:** The survivor mentions obtaining a false name and being rescued from France but does not specifically mention Fry.

**Reasoning:** Varian Fry created an underground operation to smuggle over 2000 Jews out of France from 1940-1941. Using a false name and being in France constitute strong hints for smuggling associated with Fry.

### ***(2.2) Backward Inference (abduction)***

Both backward inference and forward inference are causal reasoning. Backward inference is “tracing back” or “backward chaining”, reasoning from effect to cause, or from what is acted upon to actor. In general, backward inference has higher inferential strength than forward inference. As we are tracing backward from the effect or consequence to what has happened before, there is a closed or much restricted reasoning space and thus lower probability to go wrong.

#### ***(2.2.1) Inferring an event or phenomenon from its consequence***

The evidence itself does not mention a particular event (or phenomenon) directly, but the consequences caused by the event (or phenomenon) lay out substantial clues for us to trace back to the event (or phenomenon).

**Topic:** Materials that support or rebuff the claim that Bulgaria saved its Jews from Nazism

**Evidence:** A survivor comments about the quality of life being better in Bulgaria.

**Reasoning:** It does not explicitly address the Bulgarian government's policy to its Jews, but better living quality in Bulgaria is definitely one important effect resulting from the leniency of the government.

#### ***(2.2.2) Inferring an event or phenomenon from events that happen later***

The event (or phenomenon) asked by a topic is missing from a survivor's life story narration but the survivor discusses some other events that happened following the particular event. Sometimes those later events can lead us to a conclusive argument about the earlier target event (or phenomenon).

**Topic:** Nazi theft and expropriation of family property and assets

**Evidence:** Segments describe forced labor of sorting clothes, Jewels, and Jewish ritual objects.

**Reasoning:** The intensity of sorting labor and the details of sorting process indirectly demonstrate the severity of seizure of properties/ valuables by Nazis that happened earlier.

#### ***(2.2.3) Inferring an action or phenomenon from reaction***

Reactions towards experiences and attempted experiences rather than those that actually occurred fall into this subcategory. The target event is not mentioned or may not have happened at all, but reaction, perception, feeling, attitude, or attempt is a good mirror to

reflect what has gone on before.

**Topic:** Nazi theft and expropriation of family property and assets

**Evidence:** Segments discuss Jewish efforts to hide property.

### **(2.3) Forward Inference** (deduction)

Forward inference is “looking ahead” or “forward chaining”, reasoning from cause to effect, or from actor to what is acted upon. Looking forward is essentially making predictions. Just as any kind of prediction, it deals with an open reasoning space and the possibilities of a predicted event are many if not infinite. Thus, we can infer only with a low or medium level of certainty.

#### *(2.3.1) Inferring an event or phenomenon from its cause*

#### *(2.3.2) Inferring an event or phenomenon from events that happened earlier*

If the probability of the association of an early event A and a later event B is high, and if the actual occurrence of early event A is known, we can surmise that the later event B, the one of interest, did also occur.

**Topic:** Stories of children hidden without their parents and of their rescuers

**Evidence:** A survivor tells of his sister’s absence on the day of the roundup and explains that she had been delivering food to extended family members already in city X.

**Reasoning:** Not being at the roundup may lead to later hiding experience of his sister, however faintly.

#### *(2.3.3) Inferring reaction or feelings from action or phenomenon*

When a topic asks for reactions or feelings about an event rather than the event itself, this type of indirect evidence is particularly important.

### **(2.4) Inference from Cases** (induction)

There is a fourth type of indirect evidence, not observed in our data: examples from which one can induce a pattern that answers the question.

## **(3) Context Relevance**

Contextual information helps us to better understand or describe a central event by seeing the general picture where the central event fits in. It can be the setting or environment, the factors or effects, something allowing or hindering an event, something happening behind the scene, etc. To conclude, context evidence is information not specifically *on* a topic, but *surrounding* the topic. Context evidence is something we use to backup an argument but not to base an argument on. There are four major subtypes of “surrounding” a topic: by *scope*, by *causal sequence*, by *time sequence*, and by *place*.

### **(3.1) Context by Scope**

Doing research on a specific event is similar to using a camera to take pictures. If we focus only on the event, we collect directly or indirectly relevant information that is right on target. By adjusting the lens, we start to see the background and gain a broader view on the target event. In this sense, context evidence is something happening in the background that enriches our understanding of what is going on in the foreground. It sets up a big picture on the physical, political, social, and cultural level.

### (3.1.1) Context as environmental setting

Descriptions of the physical setting and environmental factors, such as the general camp life, living conditions, medical conditions, etc.

### (3.1.2) Context as social/political/cultural background

This subtype addresses more intangible elements of social, political, and cultural aspects at the time when the target event happened. Most historical events did not happen in a vacuum; instead, they are better described as the highlights of some ongoing trend at the time. We better understand a historical event or phenomenon if we look at it in its broad context.

**Topic:** Stories of Varian Fry and the Emergency Rescue Committee who saved thousands in Marseille

**Evidence:** A survivor details the political situation in France in 1940-1941 regarding refugees and explains the changes in emigration regulations that made fleeing France difficult.

### (3.1.3) Context as other supplemental information

Including supplemental information such as statistics, something happening in parallel with the target event, etc.

**Topic:** Descriptions of Nazi medical experiments

**Evidence:** General discussion of medical care in the concentration camps where medical experiments were conducted.

## (3.2) Context by Causal Sequence

This dimension suggests the causal information surrounding an event of interest. It situates our understanding of a target event into a causal network which helps illuminate relationships among events. However, the causal network constructed by context evidence only tells us what affects or what is affected in a broad way. It is different from the causal relationships involved in direct and indirect relevance. Direct and indirect evidence provide a restrictive evidential space that leads to a specific answer (or fact) while context is much more open to different possibilities and thus does not necessarily lead to *one* answer or any conclusive argument at all. As described in the following example, “the authorities often raided the convent” is one factor that hinders children from hiding in a convent, but it does not sufficiently *cause* hiding or not hiding.

### (3.2.1) Context as helping or hindering factor

It includes both helping and hindering factors behind an event or phenomenon. These factors are *affecting* but not sufficiently *causing* a target event to happen or not to happen.

**Topic:** Stories of children hidden without their parents and of their rescuers

**Evidence:** Mentions of factors hindering hiding, such as “the authorities often raided the convent”; some children were hidden in convents during the war.

### (3.2.2) Context as effect/ influence

Discussions of how situations or conditions were affected by a target event. It indicates the *influence* but not the logical *consequence* of a target event.

**Topic:** Descriptions of Nazi medical experiments

**Evidence:** A survivor describes his long-term conditions related to the medical experiments in which he was forced to participate.

## (3.3) Context by Time Sequence

This dimension is concerned with the things that happened close to the target event in time. While the first dimension describes the background *at* the time, this dimension describes something that happened immediately *before* or *after* a target event. The preceding and following experience (or event) links the isolated descriptions of events together and provides a more continuous view over the events' development. The target event is either the starting point or the ending point of the context evidence. Unlike forward- and backward-inference evidence, their relations to a target event are certain and explicitly stated.

### (3.3.1) *Context as preceding experience/event*

**Topic:** Descriptions of Nazi medical experiments

**Evidence:** The prisoner selections conducted by Dr. Mengele in concentration camps that are related to medical experiments.

### (3.3.2) *Context as following experience/event*

It provides follow-up information and gives an idea of what was going on in survivors' lives following the time the target event or phenomenon took place.

### (3.4) *Context by Place*

This dimension is concerned with the things that happened close to the target event in place. (Not observed in the data.)

## (4) **Comparison Relevance**

Comparison relevance is driven by perceived similarity, identifying both analogous and contrasting persons, places, events or phenomena that can help in understanding a topic; it is related to analogical reasoning. It is not *on inferring* nor *surrounding* the target event or phenomenon; it is another event or phenomenon. That is why we do not use comparative evidence as valid proof in court cases. Its evidential value in terms of establishing a fact is even fainter than contextual evidence, which is at least remotely related to the *exact* event. But when it comes to justify a judicial decision, comparison relevance becomes useful in identifying comparable precedents. To recognize similarity among seemingly discrete facts is at the heart of human thinking and reasoning; it establishes connections, inspires thinking, generates perspectives, and improves distinction among similar facts. On the one hand, by looking at similar cases, we obtain supplemental details, develop a comprehensive view on the same sort of events, and know better of something *unique* about the target event; on the other hand, by looking at contrasting cases, we see the other side of the coin and gain alternative perspectives about the target event. Moreover, in cases where little material on the *exact* event is available, comparable cases can also *induce* some arguments but just not as strong and conclusive.

Comparative evidence can be defined as information about a topic that shares characteristics of the topic but differs from the topic in one or more aspects. A topic usually consists of several aspects or facets. A typical MALACH topic can be described by three major facets: external factors: time and place; participants; and the event/ experience/ phenomenon. Varying values of one or two of these facets, we obtain similar or contrasting cases. (Varying all three at once leaves no similarity or basis for comparison.) Varying values of the first two topical facets, we get the same or comparable event/ experience/ phenomenon happening in a different place, at a different time, in a different situation, or with a different person; varying



values of the last facet, we get an opposite event/ experience/ phenomenon happening in the same time-space or involving the same participant(s). The three major topical facets define three specific types of comparative evidence:

#### ***(4.1) Comparison: Varying External Factor(s), Time/Place***

Many topics specify the place or time of target events or phenomena, such as *The Jews of Shanghai*, *The Postwar Reception of Holocaust Survivors by the American Jewish Community 1945-1954*, etc. Varying elements on these two facets, we get to see similar or comparable events (or phenomena) that happen at a different time and/or a different place that shed light on the topic.

##### *(4.1.1) Comparison: Happening at a difference place*

**Topic:** The Postwar Reception of Holocaust Survivors by the American Jewish Community 1945-1954

**Evidence:** Mention of reception in Netherlands, 1945-1954.

##### *(4.1.2) Comparison: Happening at a different time*

#### ***(4.2) Comparison: Varying the Participant(s)***

Some topics ask for events (or experiences) involving specific participant(s), such as a specific person, a specific group, a specific government, a specific community and so on:

- in the topics such as *Nazi Theft and Expropriation of Family Property and Assets*, the actor who takes an active role in the target event is specified;
- in the topics such as *Children Removed from Their Parents*, the subject who is subject to or acted upon by the target event or experience is specified.

##### *(4.2.1) Comparison: A different actor*

Varying the actor in a comparable event or activity, we get to know the attitude, policy, action taken by a different government, a different resistance group, a different community, etc.

**Topic:** Nazi theft and expropriation of family property and assets

**Evidence:** Seizure of property by Axis governments other than the Nazis, e.g., Hungary pre-1944.

##### *(4.2.2) Comparison: A different subject being acted upon*

A subject can be different in terms of age, gender, attitude, behavior, religion, race, skill, situation, etc.

**Topic:** Treatment of the disabled during the Holocaust

**Evidence:** The Nazi experimentation and extermination of gypsies, homosexuals, twins, and the elderly. These non-disabled provoked the same type of cruelty from Nazis during the Holocaust.

#### ***(4.3) Comparison: Varying the Act/Experience***

Variations on this topical facet often provide contrasting evidence for a topic. It enables us to understand how similar situations engender different actions or events, and how similar participants make different decisions or go through different experience. For instance, under the topic *Strengthening Faith by Holocaust Experience*, a survivor speaks of his loss of faith

in Auschwitz and his antipathy to the religious rationalization of the Holocaust. Seeing the other side of the coin usually enriches our thinking of a topic.

#### (4.3.1) Comparison: A different act

Different attitude, feeling, policy, treatment, activity, or experience taken by comparable actor.

**Topic:** Strengthening faith by Holocaust experience

**Evidence:** A survivor speaks of his loss of faith in Auschwitz. He stopped believing in the form religion, be it Jewish, Catholic, Protestant, or else. He saw a religious fellow teach the young people that the Jews of Poland and Czech and Romania sinned but the Jews of Florida and of New York did not sin. He cannot believe there is a God who could see what was happening in Auschwitz and permit it.

#### (4.3.2) Comparison: A different experience

Different experience, activity, or event that acts upon a similar subject.

**Topic:** The Postwar Reception of Holocaust Survivors by the American Jewish Community 1945-1954

**Evidence:** Segments discuss reception of other displaced ethnic groups.

## Conclusion

This exploration of topical relevance from the evidence perspective shows the highly complex and dynamic relationships involved in the concept. By sorting out the rich topical nuances discovered in the MALACH test collection, we identified and explored four types of topical relevance that play different roles in making topical connections between a piece of information/evidence and a topic:

- Direct evidence explicitly *gives* an answer to a user's question;
- indirect evidence lets us *infer* an answer to the question;
- contextual evidence provides peripheral or background information *surrounding* an answer; and
- comparative evidence provides a basis for *interpretation* or *inspires* some answer through perceived similarity to the question.

The common interpretation of topical relevance as a matching relationship is too limited and downplays the importance of detecting other topical connections. As a consequence, present IR systems limit the user in the types of information/evidence they can easily find when exploring a topic.

### *The Four Types of Topical Relevance*

Direct, indirect, contextual, and comparative evidence play different roles in reasoning, making a conclusive argument, deriving a rich understanding, or performing a task. Direct and indirect relevance provide a relatively restrictive evidential space leading to a specific answer/fact while context is much more open to different possibilities and does not necessarily lead to *one* answer at all. For fact-establishing tasks (for example, in law), direct (seeing the suspect shoot at the victim) and indirect (seeing the suspect throwing a gun away) evidence may differ in inferential strength but both can serve as valid proof. In this situation, comparative evidence cannot be used to establish a fact. However, when the judge looks at what laws and precedents apply to the case at hand, comparison becomes the essential

reasoning tool. Only by looking at the similarities and discrepancies in various dimensions can the judge establish critical connections between cases and make appropriate decisions. Likewise, in the history domain, if the goal is to recover a historical event, direct, indirect, and sometimes context evidence is more useful; but if the goal is to generate interpretation, looking at similar/contrasting historical events can provide much insight. Another example is case-based reasoning in weather forecast where one gathers and compares similar cases and uses *inductive* reasoning to predict future events. Comparative evidence is only a last resort for establishing facts but it is useful for deriving perspectives and inspiring informed guesses.

It may be argued that IR systems with their limitation of finding just direct evidence have served us well; why do we bother to have other types of evidence? This view is oversimplified and fails to take into account the large variety of information needs, user situations, and user tasks:

- In many situations, direct evidence is simply not available. For instance, in court cases, direct witnesses are preferred but often not available. Much trial reasoning is based on indirect (circumstantial) evidence. In the history domain, when direct evidence on an event is missing, historians collect indirect, contextual, or comparative evidence to make hypotheses and establish arguments.
- For many tasks direct evidence alone is not sufficient. The need for comparable cases in law is discussed above. For another example, in Evidence-based Medicine a direct quick diagnostic answer is not sufficient; what matters is how the physician gathers and uses evidence to arrive at the answer, how well she understands the background and problem of an individual patient, and how much she is aware of comparative treatments or tests. EBM puts a focus on contextual and comparative evidence into clinical information seeking.

### ***Implications for Information Seeking and Information Retrieval Systems***

Knowledge of the types of reasoning equips the user to search from different angles and to apply the information found to a task (Huang & Soergel, 2004). With present IR systems which provide only direct evidence on a query topic but fail to effectively respond to non-direct types of user requests, users must think of the types of information that could be relevant indirectly or context or by comparison and conduct separate searches for each, ask each question in a “direct” way. For example, consider a search for contextual information pre-9/11, such as cultural conflicts before 9/11. Instead of submitting a query of “context pre-9/11”, the user needs to recall a particular event or person that she knows to be relevant to start the search. In this way, she makes a new “target” for the search. First, this stops users who are curious but know little about the background; second, even it is feasible for some knowledgeable users, it still requires extra mental effort. Instead of focusing on the target topic, users must shift their attention to formulate their questions in a “direct” manner and constantly orient themselves to one new target after another. Users are doing extra work of preprocessing their non-direct requests before they can interact with the direct-oriented IR systems. Systems capable of detecting information that is relevant in a non-direct way allow users to keep their focus and stay at the center of their task, yet support them in thinking about the task more comprehensively.

We can incorporate new capabilities into IR systems by indicating topical relevance relationships in indexing: The indexer not only identifies the topic to which a piece of information is related but also indicates in what way they are related. This may substantially improve the systems' response to different types of tasks in different user situations. Indexing topical relationships to achieve the desired flexibility is just a starting point. A more general solution is to equip the system with reasoning power so that it can detect what information in the collection is relevant directly, indirectly, in context, or by comparison. Such a facility may be supported by an ontology that stores type-of-evidence relationships at the concept level. For a system that works along these lines, using the relevance types implied by the PICO (Problem, Intervention, Comparison, Outcome) frame in Evidence-Based Medicine as precision device, see Niu & Hirst (2004) and Lin (2006).

From the perspective of task-based retrieval (Vakkari, 2003), the relevance types indicate ways in which a piece of information supports a task. Thus, to identify a user's information needs and then find and present appropriate information one should

- identify the type of the task/question and the corresponding argument structure;
- fit available pieces of information into the argument structure;
- determine gaps in the information needed to complete the argument, and the roles these missing pieces play in the argument;
- find pieces of information that are in the proper relationship to the question (type of relevance);
- present the results organized by type of relevance or, even better, organized by the structure of the argument (possibly shown graphically).

### **Outlook**

The role of different types of evidence must be understood in the context of a particular task/question type *in a particular domain*. So a broader analysis of the types of relevance should proceed from an analysis of the types of tasks/ questions and the types of reasoning used in the argument structure in different domains. Such analyses can be found in writings devoted to methods from different perspectives: philosophy, psychology (human thinking), decision making, and research methods. There will be many question- and domain-specific ways in which a piece of information relates to a task/ question, and thus many nuanced types of relevance. The analysis should look for commonalities and differences across domains and types of tasks/ questions.

Exploration of user tasks and understanding the full complexity of topical relevance are intimately linked. We need more effort to advance the understanding of both and ultimately to improve IR system performance by creatively incorporating an enriched concept of user tasks and of topical relevance.

### **Acknowledgements**

This work has been supported in part by NSF award IIS 0122466. We thank Douglas Oard and Craig Murray for early discussions on the types of topical relevance in the context of MALACH, David Doermann, Liliya Kharevych, Josiane Alagbe, and Ayelet Goldin who created and maintained the assessment system, Ryen White and Firouzeh Jalilian for the

scripts extracting and reformatting segment assessments from the database, and most of all our assessors for their elaborate notes and insightful comments.

## References

- Barry, C. (1993). *The identification of user criteria of relevance and document characteristics: Beyond the topical approach to information retrieval*. Unpublished doctoral dissertation, Syracuse University, Syracuse, NY.
- CLEF-CLSR (2006). Cross-Language Evaluation Forum – Cross-Language Speech Retrieval. Retrieved April 30, 2006 from <http://clef-clsr.umiacs.umd.edu/data.html>
- Cooper, W. S. (1971). A definition of relevance for information retrieval. *Information Storage and Retrieval*, 7, 19-37.
- Green, R. & Bean, C. A. (1995). Topical relevance relationships. I. Why topic matching fails. II. An exploratory study and preliminary typology. *JASIST*, 46(9), 646-653, 654-662.
- Green, R. (1997). The role of relational structures in indexing for the humanities. *Knowledge Organization*, 24(2), 72-83.
- Gustman, S. *et al.* (2002) Supporting access to large digital oral history archives. *JCDL 2002*, 18-27.
- Huang, X. & Soergel, D. (2004). Relevance judges' understanding of topical relevance types: An explication of an enriched concept of topical relevance. *ASIST 2004*, 156-167
- Kadane, J. B. and Schum, D. A. (1996). *A Probabilistic Analysis of the Sacco and Vanzetti Evidence*. New York: John Wiley & Sons, Inc.
- Lawley, K., Soergel, D. & Huang, X. (2005). Relevance criteria used by teachers in selecting oral history materials. *ASIST 2005*, 421 – 448
- Lindlof, T. R. & Taylor, B. C. (2002). *Qualitative Communication Research Methods* (2nd ed.). Thousand Oaks, CA: Sage.
- Lin, J. (2006) The role of knowledge in conceptual retrieval: A study in the domain of clinical medicine. *SIGIR 2006* (in print.).
- MALACH project. Retrieved 29 January, 2006 from <http://www.clsp.jhu.edu/research/malach/>
- Mizzarro, S.(1997), Relevance: The Whole History. *JASIST 48 (9)*, 810-832.
- Niu, Y., & Hirst, G. (2004). Analysis of semantic classes in medical text for question answering. *Workshop on Question Answering in Restricted Domains, 42nd Annual Meeting of the Association for Computational Linguistics*.
- Saracevic, T. (1996). Relevance reconsidered. *Proceedings of the 2nd conference on conceptions of Library and Information Science (CoLIS 2)* (pp. 201-218).
- Saracevic, T. (1975). Relevance: a review of and a framework for the thinking on the notion in information science. *JASIST*, 26, 326-343.
- Saracevic, T. (2006). Relevance: A review of the literature and a framework for thinking on the notion in information science. Part II. *Advances in Librarianship 30* (in print).
- Schamber, L., *et. al.* (1990). A re-examination of Relevance: Toward a dynamic, situational definition. *Information Processing & Management*, 26(6), 755-776.
- Schamber, L. (1994). Relevance and information behavior. *ARIST 1994*, 3 - 48
- Shoah Visual History Foundation. Retrieved 29 January, 2006 from <http://www.usc.edu/schools/college/vhi/>
- Stong, J. W., *et. al.* (1999). *McCormick on Evidence* (5th. ed., Sec. 185). St. Paul, Minn.: West Group.
- Twining William (2003). Evidence as a multi-disciplinary subject. *Law Probability and Risk*, 2, 91 - 107.

Vakkari, P. (2003). Task-Based Information Searching. *ARIST* 37, 413-464

Wang, P.& Soergel, D. (1998). A cognitive model of document use during a research project. Study I. Document selection. *JASIST*. 49(2): 115-133

Wilson, P. (1973). Situational relevance. *Information Storage and Retrieval*, 9, 457-471.