Developing a Cross-Disciplinary Typology of Topical Relevance Relationships as the Basis for a Topic-Oriented Information Architecture

Xiaoli Huang

PhD, University of Maryland

xiaoli@umd.edu; www.terpconnect.umd.edu/~xiaoli

Abstract

This submission reports on a cross-disciplinary inquiry into topicality and relevance, involving an in-depth literature analysis and an inductive development of a faceted typology (containing 227 fine-grained *topical relevance* relationships arrayed in three facets and 33 *types of presentation* relationships). This inquiry reveals a large variety of topical connections beyond *topic matching* (the common assumption of topical relevance in the field), renders a closer look into the structure of a topic, and induces a generic topic-oriented information architecture that is meaningful across topics and domain boundaries. The findings from the analysis contribute to the foundation work of information organization, metadata development, intellectual access / information retrieval, and knowledge discovery.

The typology of topical relevance relationships is structured with three major facets:

- Functional role of a piece of information plays in the overall structure of a topic or an argument;
- Mode of reasoning: How information contributes to the user's reasoning about a topic;
- Semantic relationship: How information connects to a topic semantically.

This inquiry demonstrated that topical relevance with its close linkage to thinking and reasoning is central to many disciplines. The multidisciplinary approach allows synthesis and examination from new angles, leading to an integrated scheme of relevance relationships or a system of thinking that informs each individual discipline. The scheme resolving from the synthesis can be used to improve text and image understanding, knowledge organization and retrieval, reasoning, argumentation, and thinking in general, by people and machines.

Keywords: Knowledge Organization, Knowledge Structures, Knowledge Representation, Information Architecture, Knowledge Management, Qualitative Research

Introduction

Topical relevance is the central concept of information seeking and information retrieval; yet our understanding and research of topical relevance is not matching up with its importance. Most recent relevance research in information science studies user behavior and criteria applied by users in assessing relevance and usefulness, often focusing on non-topical criteria rather than analyzing the structure of topics and in what ways the relevant information relates to a topic. With very few exceptions, topical relevance is treated as a single relationship type, matching topic, without further discussion or analysis of this complex concept. A major goal of this research is to draw attention to this worrying gap and revive the discussion on the "topical layer" of the puzzle, especially under the current circumstance where information overload and content management have become such pressing issues.

Using *qualitative content analysis*, the inquiry focuses on meaning and deep structure. It consists of two phases:

- Phase 1: develop a unified theory-grounded typology of topical relevance relationships through close reading of literature and synthesis of thinking from communication, rhetoric, cognitive psychology, education, information science, argumentation, logic, law, medicine, and art history;

- Phase 2: in-depth qualitative analysis of empirical relevance datasets to examine manifestations of the theory-grounded typology in various contexts and to further refine the typology; three relevance datasets in oral history, clinical question answering, and art image tagging were used for analysis to achieve variation in form, domain, and context.

This article focuses on reporting Phase 1 of the research and the applications of the derived typology.

Cross-Disciplinary Literature Analysis

Relevance lies at the heart of human cognition; in turn, *topical relevance* lies at the heart of relevance. The concept of relevance and topicality is so fundamental that it becomes an inevitable subject for discussion in any field that is concerned with human thinking, reasoning, and learning, even though these fields and literatures may not label it as such. For example, rhetoric labels the topical connection by "rhetorical functional role"; cognitive psychology calls it "cognitive mechanism".

Relevance and topicality is a central notion of human cognition shared and enriched by thinking and theories from many disciplines. The disciplines selected for analytical review can be divided into those that emphasize human thinking, reasoning, communication, and learning in a *general context*: argumentation & logic, cognitive psychology & education (learning theory), communication (relevance theory), rhetoric (rhetorical structure theory), and information science, and those that focus on thinking and reasoning in *specific subject domains*: legal reasoning, history, clinical medicine, and art theory and art history.

Using *qualitative content analysis*, the literature analysis involved two steps:

1. *Identifying, collecting,* and *extracting* types of topical relevance relationships, definitions, associated examples, and use contexts from the literature. Although ongoing comparisons took place all along the reading and coding, this phase focused on the idiosyncratic, or the differences.

2. *Comparing* and *integrating* the relationships identified from different domains into a unified typology of topical relevance relationships. This phase focused on the representative, or the convergences.

See specific coding examples in Appendix A.

The literatures reviewed approach the intangible notion of topicality and relevance from many angles and contribute to elaborating its substance. The analysis focuses on what is *generically* true about the concept instead of going into details of domain-dependent specifics. The analytical literature review identified fine-grained topical relevance relationships and organized them into a typology of three

2

facets plus an additional presentation facet, a *theory-grounded* typology of topical relevance relationships, as summarized in Table 1 in the next section.

Among all literature reviewed, the major contributions to the structure and specific relationship types of the typology come from

- Mann & Thompson's (1988, 2006) Rhetorical Structure Theory (RST)
  (the 31 RST relationships become the "building blocks" to the function-based facet of the typology);
- Toulmin's argumentation theory (1958; 1984)
  (main source for the reasoning-based facet); and
- Green & Bean's (1995) semantic-based topical relevance relationship inventory.

Rather than incorporating these schemes and their relationships directly, the study selected and re-organized them into a systematic framework; in some cases the relationships are given more generic definitions. In particular, RST provides a rather comprehensive framework for investigating relational propositions based on *functional role*. It is an inclusive inventory of rhetorical relations that has a wide range of applications in text annotation and discourse analysis. During the review, it became clear that, from an information perspective, the inventory of RST relationships is a mixture containing relationships related to

- The substance of information, e.g., *Purpose, Evaluation*, *Means*;

- Forms of presentation, e.g., Elaboration, Definition, Summarization, Reference; and

- Emphasis on rhetorical use, as in *Concession* (Ex: Tempting as it may be, we shouldn't embrace every popular issue that comes along.) and *Antithesis* (which implies the substance-based relationship *Contrast*; what distinguishes *Antithesis* is its rhetorical use of *Contrast*)

Substance-related RST relationships deal with the essence of the given information or message, which is the focus of the present inquiry. **Presentation-related** RST relationships also differentiate types of relevant information on a topic, e.g., the relevant information can be a *definition* or a *summary*, but they account for the differences in presentation rather than in substance; they do not address the issue of in what way the given information relates to the topic, e.g., *definition* does not specify if it is matching topic, or delivers context, or provides comparisons. Presentation is a secondary aspect; relevance relationship types combine with forms of presentation. **Rhetorical-use-related** RST relationships account for differences in rhetorical devices used rather than in substance. These three types of RST relationships are orthogonal to each other. Recognizing these nuances may better structure the RST relationships and improve its applications in text analysis.

This is just one example of how the study brought in relationships from original inventories, sorted them out, and put them together under the current framework of topic-oriented information. Examining these schemes from multiple perspectives going beyond their original purpose led to new insights and frameworks that might not have been discovered otherwise. These insights inform the original theories and inventories by suggesting more thought-out structures and opening new angles for applications.

Empirical Manifestation Study

To examine manifestations of the typology in various contexts and to further refine it, subsequent qualitative analyses of empirical relevance datasets in oral history, clinical question answering, and art image tagging were conducted. Following the same rationale as the literature analysis (Phase 1), the analysis of relevance data in Phase 2 ensures that the scope of examination is comprehensive and the findings are inclusive and not limited to an individual domain. Three kinds of empirical relevance data were used to achieve considerable variations in "forms", "domains", and "contexts" (as illustrated in Table 2):

Table 2. A Summary of the Three Empirical Relevance Datasets Used for Analysis

Dimension	Dataset A	Dataset B	Dataset C
Subject domain	Oral history	Clinical medicine	Fine arts
Setting	Relevance assessment	Question answering	Subject indexing/Tagging
Information type	Audio (transcribed)	Text	Image
Participant	Graduate students in history or information science	Expert physicians	Art historians or art librarians
Sample	41 detailed relevance assessment notes on 40 topics by 8 assessors	26 pairs of clinical questions and answers	11 art images and 768 unique descriptors/tags assigned by 13 indexers

The findings provide rich examples to illustrate the large variety of topical connections between a topic or question and an information object or between two information objects. Examples of an information object are: a Holocaust survivor testimony or a passage from it, an evidence-based clinical answer or a passage from it, a tag assigned to an art image. The analysis also highlights the domain effects on refining the typology.

Result: A Typology of Topical Relevance Relationships

The primary result of the inquiry is a theory-grounded and empirically-refined typology of topical relevance relationships that deal with the substance of information. The typology consists of three facets and total 227 fine-grained topical relevance relationships:

- functional role (function-based): 151 relevance relationship types

- mode of reasoning (reasoning-based): 30 relevance relationship types; and

- semantic relations (semantic-based, developed by Green & Bean (1995)):
  56 relevance relationship types.

The secondary result is a scheme of 33 "presentation" types that can be combined with the topical relevance relationships.

The top-level topical relevance relationships characterized by the three facets are presented in Table 1.

Table 1. Top-Level Structure of the Topical Relevance Typology

Relevance facet	Definition	Top-level relationships characterized by the facet
Function-based	What functional role a piece of information plays in the overall structure of a topic.	**Matching topic**: manifestation/symptom, image content, image theme;
		Context: scope, framework, environmental setting, social background, time and sequence, assumption/expectation, biographic information;
		Condition: helping factor/condition, hindering factor/condition, unconditional, exceptional condition;
		Cause and effect: cause, effect/outcome, explanation (causal), prediction;
		Comparison: by similarity, by difference (contrast), by factor that is different;
		Evaluation: significance, limitation, criterion/standard, comparative evaluation;
		Purpose/Motivation: purpose, motivation;
		Method/Solution: method, approach, instrument, technique, style, solution.
Reasoning-based	How information contributes to users' reasoning about a topic.	**Generic reasoning;**
		Reasoning by analogy;
		Reasoning by contrast;
		Rule-based reasoning (deduction);
		Generalization (induction);
		Causal-based reasoning: forward/backward inference
Semantic-based (Green & Bean, 1995)	How information connects to a topic semantically.	**Class – Member;**
		Whole – Part (partonomy): process – step, etc.;
		Object – Attribute: adjectival, adverbial;
		Class – Subclass (taxonomy).
Secondary aspect		
Presentation types	In what form or style information is presented; it can be combined with the topical relevance facets.	**Reference;**
		Definition;
		Restatement: paraphrase, clarification, translation, representation;
		Summarization: abstraction
		Elaboration: amplification, extension, specialization/specification, object – attribute;
		Interpretation: organization, concretization, humanization, transformation;
		Emphasis / Drawing attention.

This study focuses primarily on the function-based facet and secondarily on the reasoning-based facet.

- Functional role: the role a piece of information plays in the overall structure of a topic or an argument, by taking into account its relations with other parts of the given information passage or the argument. Adopting the rhetorical structure theory (Mann & Thompson, 1988) perspective, "for every part of a coherent text, there is some function for its presence, evident to readers".

- Mode of reasoning (Evidentiary connection): logic- and inference-based relationships that link pieces of information and a topic; it can be seen as the inference chain between information and topic. This perspective is concerned with how pieces of information can be identified through an inference chain and how specifically they relate and contribute to a receiver's reasoning about a topic.

- The inquiry did not study the semantic facet on its own, since Green & Bean (1995) have provided a thorough explication on this facet in their study. Some relations from the semantic facet, such as class – member, class – subclass, whole – part (including process – step), and object – attribute (including adjectival and adverbial), were combined with the function-based facet and the presentation facet to facilitate the empirical analysis.

The typology of topical relevance relationships is a work in progress and open to further developments (especially in specific domains).

A Generic Topic-Oriented Information Architecture

The function-based facet or a subset of the facet can serve as a basis for a generic topic-oriented information architecture that organizes and structures the topic space, filling a gap in knowledge organization and content management; see Figure 1.
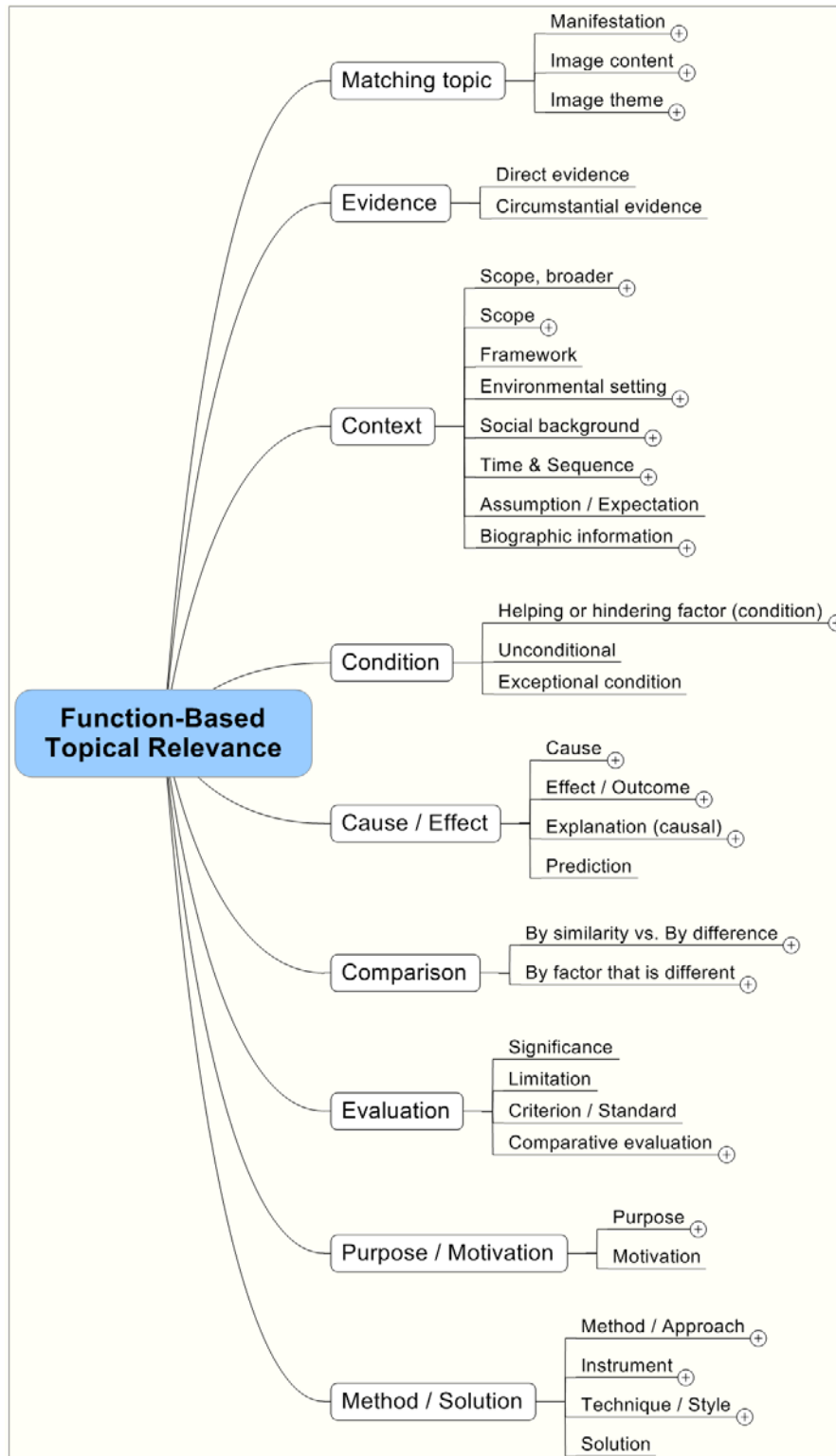
6

Figure 1. Function-Based Topic-Oriented Information Architecture

Based on the empirical analysis, the function-based topic-oriented information architecture has the followings features:

- The function-based topic framework provides the overall structure for organizing a topic. Particular domains and topics may use only some branches of the architecture and may instill domain-specific meanings to these branches, but the overall framework remains stable and meaningful across multiple domains analyzed (oral history, clinical medicine, and art images).

- The function-based topic-oriented information architecture can be easily customized to a subject domain through domain-specific definitions and extensions. For example, [Method / Solution] is used to characterize "Medical treatment" and "Diagnostic method" in clinical medicine, whereas in fine arts it is used for describing "Technique" and "Style / Genre".

- The function-based topic-oriented information architecture is a multi-level topical structure. To fully represent the complicated structure embedded in the very rich information on a topic (such as an evidence-based clinical answer), the function-based relationships and framework need to be applied on multiple levels (See Figure 2 in the following for an example). Note that in the example the topical relevance relationships are applied at each level and the presented information relates to the central topic only through "steps" of connection. For example, "A large random trial" does not directly connect to the central topic of "ADHD in children": It is not the "Evaluation" of "ADHD in children"; instead, it is the "Evaluation" of "Stimulant medication therapy", which in turn is the "Medical treatment" of "ADHD in children".
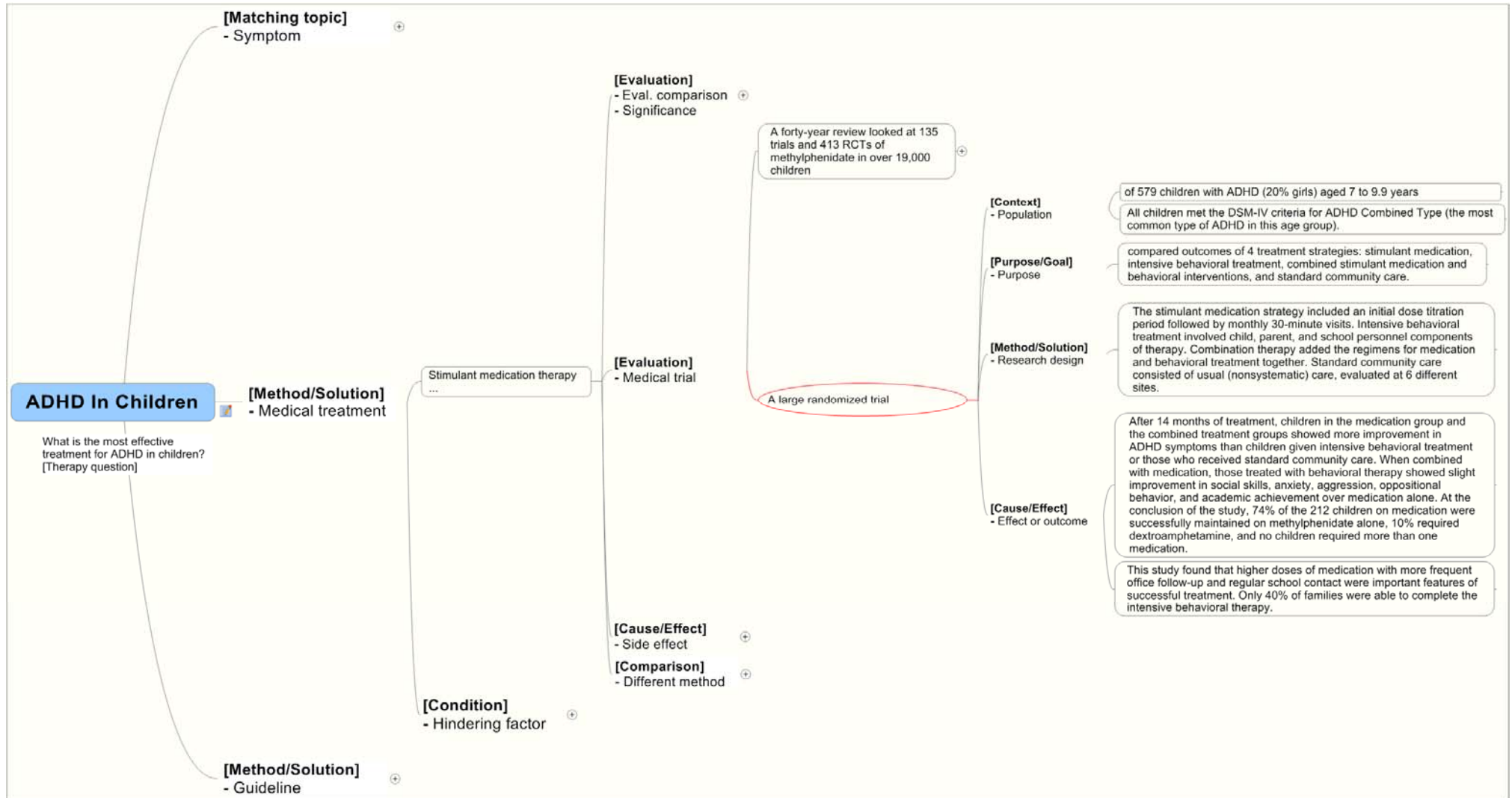
Figure 2. Zooming In On One Branch of the ADHD Therapy Question: "A Large Randomized Trial"

Implications

Content Management

The topic-oriented information architecture contributes to the important "topic layer" of content management models. More and more IT decision makers in government and in business recognize the value of "information structure" for effective content management and enterprise search, which becomes critical to improve operational efficiency and competiveness of the organization. To meet this increasingly pressing need in enterprises, solutions have been suggested with faceted metadata profiles and faceted search output for content management and access (e.g., Autonomy, or MOSS search embedded in SharePoint). Currently, the facets available in such products rely largely on document attributes and "Dublin Core" type of metadata elements: document / content type (e.g., html file, spreadsheet, blog, memo, report, and announcement), creator / author, source (e.g., division, department), creation date, version / edition, and so on. In other words, the present faceted solutions focus on form attributes. The topic component or the topic layer that plays the most significant role in meaningfully organizing content is yet missing from the foreground. The topic-oriented information architecture presented here fills this gap and provides a generically meaningful framework for organizing content. It is an explicit relational framework parallel to the non-topical metadata framework, except that it directly addresses the content whereas the other manages the properties closely associated with the content.

Topic Navigation

Structured with function-based topical relevance relationships, the topic-oriented information architecture can serve as a useful navigation tool for the user to explore a new topic space more efficiently and more systematically; it also allows the user to easily pin down specific branches of information that are precisely tailored to his/her task. Under the current context of information overload, simply saying the information is relevant but not specifying how it is related to a topic gives limited assistance to the user. It would not help the user from getting overwhelmed or lost in the sea of relevant information. Topical relationships and topical structures built upon these relationships are essential for the user to quickly make sense of a topic space, particularly when the user is not familiar with the topic.

The function-based topical information framework serves the same purpose as the Relation Browser (Gary Marchionini) and the Flamenco search interface framework (Marti Hearst), both of which are topic navigation tools based on content-oriented category metadata. This inquiry approaches the same goal from a different perspective by proposing an explicit relational structure that is discipline- and dataset-independent. The intention is to leverage topic knowledge structure and thinking across domain boundaries.

Figure 3 shows an example of a topic space structured with the topic-oriented information architecture. It is a topic map for "ADHD in Children" that combines the answers to three clinical questions that share the same "base concept" of *ADHD in Children* in the QA dataset:

- Therapy question: *What is the most effective intervention for ADHD in children?*

- Diagnosis question: *Does a short symptom checklist accurately diagnose ADHD?*

- Etiology question: *Does maternal smoking cause ADHD?*

Given the limited space, the topic map displays only the high-level structure not all the details ("+"indicates the branch can be further expanded).

In addition to structuring a single topic space, the topical relationships can also be used to link different topics and connect them into a knowledge network.
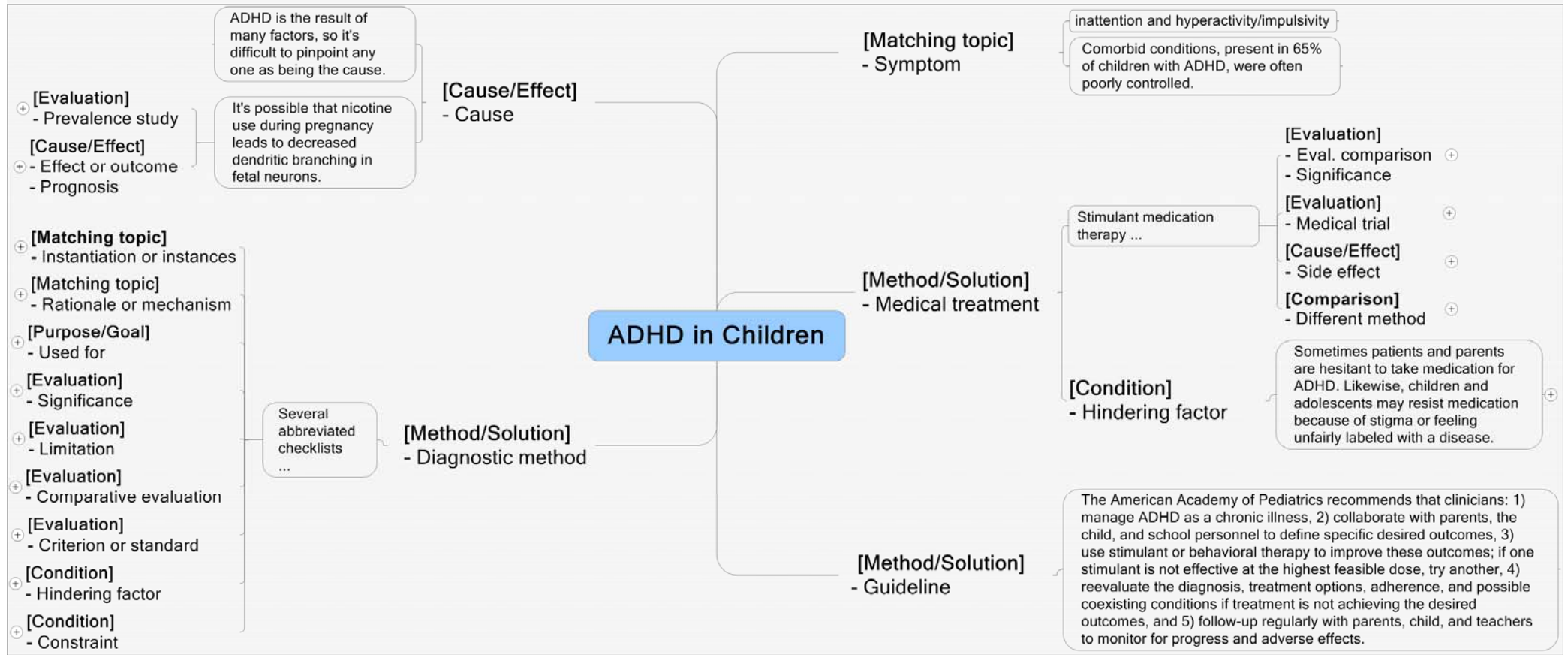
10

Figure 3. A Combined Topic Map of "ADHD In Children" From ADHD_Therapy, ADHD_Diagnosis, And ADHD_Etiology

Organize Search Output

The topic-oriented information architecture can be used to organize the search results for a topic. For example, on the result interface, instead of having a long unordered list, we can have expandable categories indicating "Social background", "Contrasting cases", "Circumstantial evidence", "Definitions", "Cause / Effect", etc. With this faceted output, the user can quickly get an overview of the topic s/he is researching. It also substantially cuts down on the time and energy to navigate to the specific aspects of interest, instead of digging through the hodgepodge of one million results. It provides similar functionality of "faceted filtering" describe in Padilla (Padilla, 2008) by allowing users to actively whittle down the information space. Figure 4 gives a sample of search results of "ADHD children" in PubMed in which the article titles are arranged by relevance categories/facets.

[Matching Topic]	**Symptom** ADHD and Working Memory: The Impact of Central Executive Deficits and Exceeding Storage/Rehearsal Capacity on Observed Inattentive Behavior. Timing of conditioned eyeblink responses is impaired in children with attention-deficit/hyperactivity disorder. **Prevalence** Prevalence of attention deficit disorder among preschool age children.
[Cause/Effect]	**Cause/Pathology** Gene-environment interactions in attention-deficit/hyperactivity disorder. Serum protein profiling and proteomics in autistic spectrum disorder using magnetic bead-assisted mass spectrometry. **Effect/Outcome** Attention deficit hyperactivity syndrome throughout the life span.
[Condition]	**Provoking factors** Short-Term Effects of Playing Computer Games on Attention. A Systematic Review of Parenting in Relation to the Development of Comorbidities and Functional Impairments in Children with Attention-Deficit/Hyperactivity Disorder (ADHD).
[Context]	**Social expectation** Potential Impact of ADHD With Stimulant Medication Label on Teacher Expectations.
[Comparison]	**Different population** Treatment seeking adults with autism or ADHD and co-morbid Substance Use Disorder: Prevalence, risk factors and functional disability. **Related disease** Disorder-specific dysfunction in right inferior prefrontal cortex during two inhibition tasks in boys with attention-deficit hyperactivity disorder compared to boys with obsessive-compulsive disorder.
[Method/Solution]	**Medical Treatment** School-based interventions for students with attention-deficit/hyperactivity disorder. Evaluation of atomoxetine for first-line treatment of newly diagnosed, treatment-naïve children and adolescents with attention deficit/hyperactivity disorder. **Side effect** Safety of stimulant treatment in attention deficit hyperactivity disorder: part I. **Helping condition** Parental Attributions for Success in Managing the Behavior of Children With ADHD. **Diagnostic Method** New insights into attention-deficit/hyperactivity disorder using structural neuroimaging.

Figure 4. PubMed Search Results (partial) on "ADHD children"

12

Empower Relational Subject Indexing

In conventional indexing, we only indicate what topic terms are relevant and assign them to the information object without differentiating the topical connections held between the terms and the information object. With this fine-grained topical relevance typology, we can also specify in what specific ways the topic terms relate to the information object. Ultimately, it helps to fine-tune the correspondence between an information object and a user's request (topic), allowing users pin down the desired information much faster and easier.

The topical relevance relationships are structured by facet; the three facets complement each other to suit different needs. This allows more flexibility for the indexer to reveal topical connections from different perspectives and to better adjust to the target audience.

Organize User Tags and Structure Social Tagging

The typology can be used to organize user tags and structure social tagging. Web 2.0 provides an open and dynamic environment for social collaboration and sharing of information (knowledge) on a scale never imagined before. Among all, social tagging empowers and engages users to organize information through generating and sharing their own metadata. No longer passive receivers at the door, individual users can now come in and contribute. They actively index the Web, ranging from texts, blogs, websites, to music, photos, fine arts, and cultural artifacts.

There has been an increasing awareness of viewing information as institutional assets to create value by aggregating knowledge across departments and discovering hidden knowledge links in the vast amount of enterprise information. Under this context, user tagging has also rapidly gained its popularity as an innovative solution to content management in government sectors and corporations, such as the fast-spreading *TagXchange* project at the IMF (International Monetary Fund).

As a tag is assigned to an information object, a relevance relationship to a particular topic (feature) is established, shared, and preserved for future recall. As tags accumulate every second, we can picture every tag weaving into this large invisible "web" of relevance relationships, rapidly grouping knowledge, meanings, ideas, and opinions for easy access. While the "social layer" of the Web has been heavily discussed among researchers, we need to focus attention to this "relevance layer" which engages users' relevance judgments and cultivates a meaningful bottom-up information structure onto the Web.

Containing rich fine-grained topic relevance relationships, the derived typology equips users to make sense of the ever-growing "tag cloud", to achieve a better understanding of the relationship propositions underlying the "relevance web", and to better facilitate its development. In particular, we can directly apply the typology to arrange user tags into a more useful presentation, using the following Painting (Figure 5) and Flickr image (Figure 6) and associated tags as examples:

13

Figure 5. A Young Woman and Her Little Boy. Artist: Agnolo Bronzino. 1540. Florentine.

Matching topic
- **Image content: Focal**
  - *Reference*
    - Boy,
    - Child {children},
    - Mother and child,
    - Woman {women},
  - *Elaboration (Adj.)*
    - Elegance,
    - Elongation,
    - Lavish,
    - Opulence,
    - Wealth,
  - *Elaboration (Adv.)*
    - Embraces,
- **Image content: Peripheral**
  - *Reference*
    - Costume,
    - Head coverings,
    - Gloves,
    - Textiles,
    - Jacquard,
    - Jewelry,
  - *Elaboration (Adj.)*
    - Red,

Effect
- **Reaction or feeling**
  - Cold,
  - Detached,
  - Hardness,

Method
- **Style or genre**
  - Figure,
  - Portrait {portraits, portraiture},
  - Group portraits,
  - Mannerism,
  - Florentine painting,
  - Italian art—16th Century {Italian art},
- **Style or genre: Detail**
  - Raking light,

Context
- **Biographic information: Artist**
  - Agnolo Bronzino {Bronzino},
- **Social background**
  - Medici court {Medici}

Figure 6. Flickr Image Example: Wasteland

***Tag*s:** stairs
stadium
bcplace
distagon35
railing
empty
alone
vancouver
day
black and white (bw)
film
bathroomdarkroom
rodinal
ilford
fp4
contax
rtsii
stand
development
standdevelopment

Image Tags Organized With the Topic-Oriented Information Architecture

Matching topic

- **Image theme: Title**
    - Wasteland
- **Image content**
    - *Reference*
        - stairs,
        - stadium,
        - railing,
    - *Elaboration (Adj.)*
        - empty,
        - alone,

Context

- **Environmental setting: Physical location**
    - bcplace,
    - vancouver,
- **Time & sequence: Time / Period**
    - day,

Method / Solution

- **Style / Genre**
    - blackandwhite (bw),
- **Style / Genre: Design or composition**
    - standdevelopment {stand, development}

- **Technique**
    - bathroomdarkroom,
- **Instrument**
    - distagon35,
    - contax rtsii,
- **Instrument: Material or medium**
    - film,
    - ilford fp4,
    - rodinal

15

References

Green, R. (1995). Topical relevance relationships. I. Why topic matching fails. *Journal of the American Society for Information Science, 46*(9), 646-653.

Green, R., & Bean, C. A. (1995). Topical relevance relationships. II. An exploratory study and preliminary typology. *Journal of the American Society for Information Science, 46*(9), 654-662.

Mann, W. C., & Thompson, S. A. (1988). Rhetorical structure theory: Toward a functional theory of text organization. *Text, 8*(3), 243-281.

Padilla, M. (2008). User Interface Implementations of Faceted Browsing Digital Web Magazine [Electronic Version] Retrieved from http://www.digital-web.com/articles/user_interface_implementations_of_faceted_browsing/

Taboada, M., & Mann, W. C. (2006). Rhetorical Structure Theory: Looking Back and Moving Ahead. *Discourse Studies, 8*(3), 423-459.

Toulmin, S. (1958). *The Uses of Argument*. Cambridge, England: Cambridge University Press.

Toulmin, S., Rieke, R. J., & Janik, A. (1984). *An introduction to reasoning*. New York: Macmillan Publishing Co., Inc.

Appendix A:
Examples of the Conceptual Analysis of the Relevance Literature

Example 1: Law of Evidence

Source: *State v Famber*, 358 Mo 288, 214 SW2d 40.

Direct evidence is testimony or other proof which expressly or straight-forwardly proves the existence of a fact. It is different from circumstantial evidence, which is evidence that, without going directly to prove the existence of a fact, gives rise to a **logical inference** that such fact does exist.

Direct evidence is evidence which, if believed, proves the existence of the fact in issue **without inference or presumption**. It is evidence which comes from one who speaks directly of his or her own knowledge on the main or ultimate fact to be proved, or who saw or heard the factual matters which are the subject of the testimony. It is not necessary that this direct knowledge be gained through the senses of sight and hearing alone, but it may be obtained from any of the senses through which outside knowledge is acquired, including the senses of touch or pain.

Circumstantial evidence
Evidence of an indirect nature which implies the existence of the main fact in question but does not in itself prove it. That is, the existence of the main fact is **deduced** from the indirect or circumstantial evidence by a process of **probable reasoning**. The introduction of a defendant's fingerprints or DNA sample are examples of circumstantial evidence.

Some people believe that all evidence is circumstantial because -- some observers think (and some thoughtful judges agree) -- no evidence ever directly proves a fact.

. Direct Relevance
. . **Direct evidence**
. . . Example
...

. Comparing Diff. Relevance
. . Direct vs. Circumstantial
  evidence
. . . **Inference**
...

. Inference
. . Logical Inference
. . . **Deductive**
. . **Probable inference**
...

. Indirect Relevance
. . **Circumstantial evidence**
. . . **Example**
...

Example 2: Information Science

Source: Cooper, W. S. (1971). A Definition of Relevance for Information Retrieval. *Information Storage and Retrieval, 7*(1), 19-37.

Logical relevance: A stored sentence is logically relevant to (a representation of) an information need if and only if it is a member of some minimal premiss set of stored sentences for some component statement of that need.

That is to say, the relevance of a sentence to a need is dependent entirely upon its membership in a minimal stored set from which an answer to the need can be **deduced**. This is the basic definition of logical relevance, restricted of course to the special conditions 1-3. As an example of this definition, if two of the sentences stored were "An atom of hydrogen contains just one proton" and "No atom of any halogen element contains less than six protons' they would each be considered relevant to the need expressed in the question "Is hydrogen a halogen element?" (p.24-25)

. Indirect Relevance
. . Circumstantial evidence
. . . Example
. . **Logical relevance**
. . . **Example**
. . **Evidential relevance**
. . Example
...

Source: Wilson, P. (1973). Situational relevance. *Information Storage and Retrieval, 9*, 457-471.
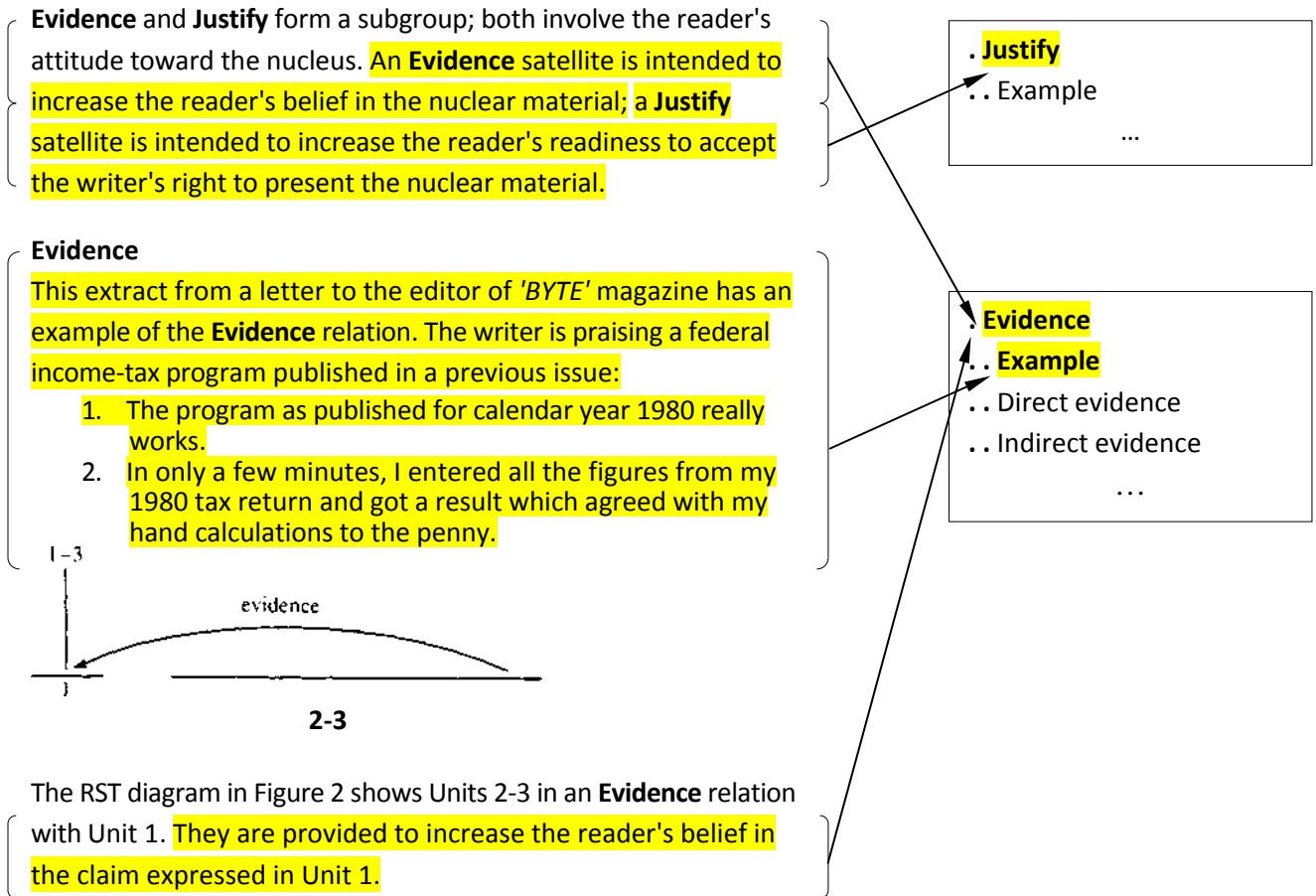
Evidential relevance: This is a sort of relevance: evidential relevance, we shall call it. It is a sort of relevance that cannot be completely understood in terms of the notion of logical consequence, but requires the notion of the degree of **confirmation**, or **probability**, of conclusions in relation to given premises. The simplest, no doubt inadequate, explanation of evidential relevance is this: an item of information I_j is relevant to a conclusion **h** in relation to premises **e** if the degree of confirmation, or probability, of **h** on evidence e and I_j is greater or less than the degree of confirmation, or probability, on e alone. (p.460)

. Inference
. . Logical Inference
. . . **Deductive**
. . . **Inductive**
. . **Probable inference**
...

Example 3: Communication & Rhetoric –Rhetorical Structure Theory

Source: Mann, W. C., & Thompson, S. A. (1988). Rhetorical structure theory: Toward a functional theory of text organization. *Text, 8*(3), 243-281.

Evidence and **Justify** form a subgroup; both involve the reader's attitude toward the nucleus. An **Evidence** satellite is intended to increase the reader's belief in the nuclear material; a **Justify** satellite is intended to increase the reader's readiness to accept the writer's right to present the nuclear material.

Evidence

This extract from a letter to the editor of *'BYTE'* magazine has an example of the **Evidence** relation. The writer is praising a federal income-tax program published in a previous issue:
1. The program as published for calendar year 1980 really works.
2. In only a few minutes, I entered all the figures from my 1980 tax return and got a result which agreed with my hand calculations to the penny.

1 – 3

evidence

2-3

The RST diagram in Figure 2 shows Units 2-3 in an **Evidence** relation with Unit 1. They are provided to increase the reader's belief in the claim expressed in Unit 1.

.Justify
.. Example
...

Evidence
.. **Example**
.. Direct evidence
.. Indirect evidence
. . .

19

Comparison and Integration among the Three Examples

All the three examples of literature analysis focus on *evidence* as one type of topical relevance relationship. Different disciplines reveal different aspects of the issue:

- **Rhetorical structure theory**: defines evidence as relevant information that "increase the reader's belief in the nuclear material", which is often an argument or conclusion (as shown in the example). The emphasis is placed on *increasing the receiver's belief* in an argument (conclusion). Therefore, evidence as a relevance relationship, regardless of its specific type, is always tied to some argument, e.g., a diagnostic hypothesis, a belief, a truth statement.

- **Law of evidence**: defines two types of evidence in the context of court cases—direct evidence and circumstantial evidence. To differentiate the two, *inference* is introduced, including both logical inference and probable inference. The major difference between direct evidence and circumstantial evidence, if any, is whether or not the evidence involves inferences. Rather, direct evidence and circumstantial evidence can be seen as two ends of a continuous inferential scale.

- **Information science**: Cooper's and Wilson's work directly relate logic and inference to the conceptualization of topical relevance. Taken together, they further specify two major types of evidence that involves inference: Cooper's logical relevance follows deductive logic and is based on logical consequence, which is inferring backward. Wilson's evidential relevance follows inductive logic and probable reasoning, which is inferring forward and involves uncertainty.

Taking the three perspectives together, an enriched understanding of evidence as one type of topical relevance relationship is developed. The integrated notion of evidence as relevance derived from the three examples can be represented as follows:

Evidence as relevance relationship

      By degree of inference:
          Direct Evidence
          Indirect/Circumstantial Evidence
               By type of inference:
                    Logical inference
                        Deductive/Backward inference (e.g. logical relevance)
                        Inductive/Forward inference (e.g., evidential relevance)
                        Probable inference (e.g., evidential relevance)

These examples help to illustrate the process of literature analysis.

The reading of rhetorical structure theory brings in the viewpoint of convincing a *reader* –a somewhat different angle from the common sense *user* who actively seeks relevant information for her tasks. Therefore, I need to consider broadening the scope and definition of "user" as discussed early on in Chapter 1 to include this notion.