Meta-Information in Conversational Search

JOHANNES KIESEL*, Bauhaus-Universität Weimar, Germany
LARS MEYER*, Bauhaus-Universität Weimar, Germany
MARTIN POTTHAST, Leipzig University, Germany
BENNO STEIN, Bauhaus-Universität Weimar, Germany

The exchange of meta-information has always formed part of information behavior. In this paper, we show that this rule also extends to conversational search. Information about the user's information need, their preferences, and the quality of search results are only some of the most salient examples of meta-information that are exchanged as a matter of course in a search conversation. To understand the importance of meta-information for conversational search, we revisit its definition and survey how meta-information has been taken into account in the past in information retrieval. Meta-information has gone by many names, about which a concise overview is provided. An in-depth analysis of the role of meta-information in search and conversation theories reveals that they provide significant support for the importance of meta-information in conversational search. We further identify conversational search datasets are suitable for a deeper inspection with regard to meta-information, namely, SCS and MISC. A quantitative data analysis demonstrates the practical significance of meta-information in information-seeking conversations, whereas a qualitative analysis shows the effects of exchanging different types. Finally, we discuss practical applications and challenges of meta-information in conversational search, including a case study of VERSE, an existing search system for the visually impaired.

CCS Concepts: • **Information systems → Users and interactive retrieval**; *Web searching and information discovery*; • **Human-centered computing → Interaction paradigms**.

Additional Key Words and Phrases: conversational search, information retrieval, information seeking, meta-information

1 INTRODUCTION

Conversational search marks a paradigm shift in information retrieval. The most salient difference to previous retrieval paradigms is the possibility to "talk to a search engine" instead of just demanding results. What does it mean to have a conversation with a search engine? Probably most important is the fact that the system becomes an equal conversation partner that can also take the initiative in the interaction [2, 42]. Conversational search can be seen as a major milestone in the ongoing process of enlarging the action space of information retrieval systems, as illustrated in Figure 1.[1] With the shift from classical information retrieval to interactive information retrieval, which happened sometime in the 1990s, the foundations of today's AI-enabled shift towards conversational search have been laid. Key benefits of enlarging the action space in conversational search are said to be a more natural and thereby intuitive form of interaction, the possibility to express and tackle even vague and complex information needs, and maintaining a dialog until ambiguities are resolved, choices are confirmed, and the like. Moreover, the flexible interface allows the conversation partners to bring up everything that might be relevant to the current search task, thereby making available to each other a wealth of meta-information that may be critical to tackle the task.

We thus argue that another important aspect of a search conversation is *the exchange of meta-information between user and system*, and the contributions of this paper shed light on this aspect.

*Both authors contributed equally to the paper.
[1]The next milestone may be "predictive retrieval systems" that provide answers as soon as our information needs become apparent to a passively observing system, even before questions are asked.

Authors' addresses: Johannes Kiesel, Bauhaus-Universität Weimar, Bauhausstraße 11, Weimar, Germany, johannes.kiesel@uni-weimar.de; Lars Meyer, Bauhaus-Universität Weimar, Bauhausstraße 11, Weimar, Germany, lars.meyer@live.de; Martin Potthast, Leipzig University, Augustusplatz 10, Leipzig, Germany, martin.potthast@uni-leipzig.de; Benno Stein, Bauhaus-Universität Weimar, Bauhausstraße 11, Weimar, Germany, benno.stein@uni-weimar.de.

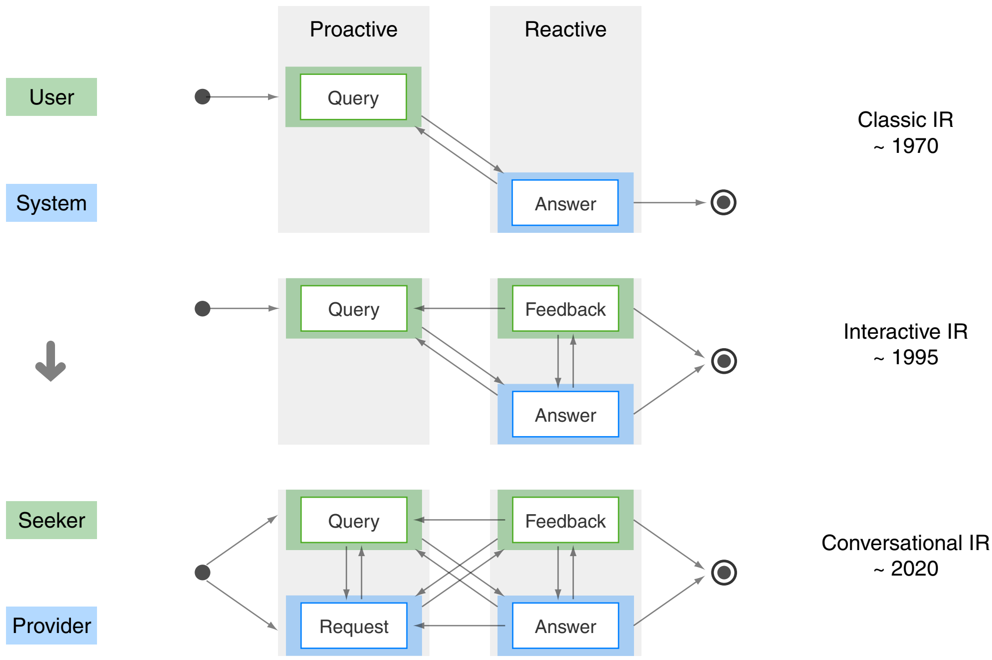

Fig. 1. Paradigm shifts in information retrieval (IR) illustrated through the QRFA-model [67] (bottom image). Where classic information retrieval focuses on answering a query (top), interactive information retrieval brings user feedback into the focus (middle), while conversational information retrieval allows a dynamic back-and-forth of actions between *equal partners* (say, seeker and provider [53]). The years are estimates, based on the volume of relevant scientific papers as per Semantic Scholar (https://www.semanticscholar.org/).

We are aware of the fact that meta-information has been exchanged between retrieval systems and their users long before conversational search came into the focus, among others, in the form of query suggestions. Moreover, ideas to exploit various kinds of meta-information in interactive retrieval date back more than 50 years [45]. Nevertheless, the importance of taking the exchange of meta-information into account in conversational search, in particular, has not been widely appreciated as of yet. This is not to say that research on meta-information is lacking—the opposite is true. However, the overarching concept of meta-information rarely becomes apparent.

Meta-information is often aptly defined as "information about information." To be more precise, one should first distinguish *information* from the related concept of *data*, two terms that are often used interchangeably. Floridi [23] provides definitions to distinguish these terms from each other:

- A datum is a putative fact regarding some lack of uniformity within some context.[2]
- σ is an instance of information, if and only if σ consists of one or more data, the data in σ are well-formed, and the well-formed data in σ are meaningful.
- Meta-information, μ, gives indications about the nature of some other information, σ.

Regarding Floridi's definition of a datum, note that the mentioned "non-uniformity" has to be created teleologically, namely for the purpose of becoming observed or retrieved. Consider, for instance, the perforation made by a ticket punch when marking admission tickets or the name of

[2]As an example, Floridi [23] gives a black dot on a white sheet of paper. The white sheet is the otherwise uniform context, broken by the black dot as the difference, making it a datum.

an addressee on an envelope. The persistent or transient carrier of data (the ticket, the envelope) is called *medium*.

We further specify Floridi's definition of meta-information by stipulating three central meta-information properties:

(1) Meta-information is identifiable as such only by its relation to other information.
(2) This (aforementioned) relation is dynamic, i.e., it can be established or broken.
(3) The meaning of meta-information depends on both the referred information and the recipient.

As an example, consider the sentence "There are many websites that discuss the statistical analyses of Michael Jordan.", which is an instance of information, σ. Note that σ may refer to Michael Jordan, the basketball player, or Michael Jordan, the statistician from UC Berkeley. The following instances of information become meta-information if they can be related to σ by a recipient who interprets σ:

μ_1 "Michael Jordan was born in 1956."
At first, observe that this sentence is an instance of information on its own; in relation to σ, it unfolds its meta-information property: Michael Jordan in σ refers to the statistician.[3]

μ_2 "Michael Jordan's middle initial is »J«."
Similarly, when used as meta-information for σ, this instance of information reveals that Michael Jordan in σ refers to the basketball player.[4]

μ_3 "I found this [σ] on a blog of the SLAM Magazine."
For a recipient who does not know the magazine, this sentence may not be meaningful. With the additional information "SLAM is a renowned basketball magazine.", however, it becomes clear that Michael Jordan in σ refers to the basketball player rather than to the statistician. In fact, "SLAM is a renowned basketball magazine." can be understood as *meta*-meta-information, μ_4, since it is in the role of meta-information to the meta-information μ_3.

Finally, observe the subtle difference between the information instances (a) and (b) on the one hand and (c) on the other. While the former two instances relate to an aspect (age, middle name) of the named entity (Michael Jordan) in σ, the latter instance relates to the "abstract role" of σ as a piece of information. We are convinced that the understanding of what meta-information is—and to which information it is linked during the states of an information-seeking process—is key for search conversations to be successful.

1.1 Terminology

We strive to employ consistent terminology in this paper, in particular for the parties involved in a search conversation. Various terms have been used in prior work, the most frequent being "user" for the party requesting the information, which is implicitly assumed to be a human. A notable distinction has been made for the party delivering the information: is it the collection of documents itself, or is it an "intermediary" (e.g., a librarian) who preprocesses the information from the collection before presenting it? Nowadays, it has become rare to interact with collections directly. Traditional search engines fulfill the role of an "intermediary" that provides a snippet to summarize each result, along with various supplementary services. However, it is not inconceivable that future information systems may unify these now separate concepts again and extend their reach to the user's site, for example, by incorporating a private knowledge base in a local appliance.

[3] Michael Jordan, the basketball player, was born in 1963.
[4] The middle initial of Michael Jordan, the statistician, is »I«.

Therefore, we avoid the traditional distinction as it derives primarily from classic information retrieval, and adopt the terms "seeker" and "provider" (cf. Figure 1), as used in Sitter and Stein [53].

Throughout the remainder of this work, the examples given are formalized by using σ to denote a piece of information, and μ to denote a piece of meta-information in relation to σ, or to some other meta-information. We analyze each piece of meta-information with respect to three attributes: (1) The *target* denotes the entity to which the meta-information is primarily related; (2) The *concept* is the group of particular instances of meta-information to which it belongs, either taken from existing literature or from our analysis of theories that incorporate meta-information; (3) The *intent* summarizes the desired effect of the meta-information.

We formally model the use and effect of meta-information in the respective examples. These *formal models* summarize the situations described in the examples and are presented as RDF-like triplets, depicting predicate relations between entities found in pieces of information, σ, and pieces of meta-information, μ, in the given conversations. To make the formalization easier to read, we sometimes use quintuplets rather than triplets, but these can be straightforwardly decomposed into two triplets. By using, we want to point to the possibility of using knowledge graphs as the data-structure for recording meta-information, which may associate meta-information with the knowledge representation in a conversational search system. Knowledge graphs are already being employed to establish the conversational search system property of memory (e.g., see Moon et al. [37]), and we believe that meta-information that is obtained by a system from various sources can be similarly integrated and utilized.

1.2 Exemplary Uses of Meta-Information in Conversational Search

The following three examples illustrate the aforementioned terminology. For brevity, these examples, like all others throughout the paper, show specific pieces of meta-information relevant to the current point of discussion. The formal models given are kept consistent and represent only the relevant portion of what can be extracted from a conversation. Example 1 illustrates time pressure, Example 2 affective signals, and Example 3 facets.

[Example 1] If a seeker starts a conversation with "I don't have much time, but I don't know much about the Spanish flu. Could you give me details?" then it is essential to understand that an answer such as the following is inappropriate: "Sure, here is an in-depth article about the historical relevance of the Spanish flu.", even if the article contains everything the seeker wants to know. To appropriately constrain the conversation's level of detail or scope, the provider needs to adjust how they present their results.

Example 1. Seeker uses meta-information to influence the conversation.

	Information		Formal model
σ	"Could you give me details [about the Spanish flu]?"	(1)	σ *sets* topic
μ	"I don't have much time [for the answer]" Target: Conversation Concept: User context [29, 48] Intent: Constrain duration	(2)	μ *constrains* duration

This observation is not new for retrieval systems. What is new for conversational search systems is both the wealth of expressions that an information seeker can employ to express their information need and that these expressions follow the style of human-to-human interactions.

[Example 2] Consider the query, “Can you give me the key facts about the Spanish flu, now.” Understanding a conversation requires one to reason about plausible interpretations of past and current states. Here, the word “now” can either signal impatience or encode the wish that the seeker wants to switch the conversation. The tone of voice of the spoken message and facial expression are valuable meta-information that helps the system interpret the query correctly. Let us assume the word “now” was said with a happy tone of voice: it then becomes unlikely that the intent was to signal impatience and constrain the conversation’s duration.

Example 2. Reasoning about meta-information.

Information		**Formal model**	
σ	“Can you give me the key facts about the Spanish flu, […]”	(1)	σ *sets* topic
μ_1	“[…] now” Target: Conversation Concept: Word choice Intent: Constrain duration	(2)	μ_1 *constrains* duration
	and/or Target: Conversation Concept: Word choice Intent: Emphasize topic change	(3)	μ_1 *supports* (1)
μ_2	Seeker’s happy tone of voice Target: Seeker Concept: Affective/physiological/behavioral features [5, 6, 38] Intent: Show affective state	(4) (5)	μ_2 *sets* affective state Affective state *opposes* (2)

[Example 3] A touted advantage of conversational search is its relative ease of building complex queries. An analysis for meta-information can break such complex queries down into small parts. Radlinski and Craswell [42] provide the following example of a complex query: “I’m looking for an email that contains a link to a research paper that I got from a student who emailed me right after SIGIR last year. I can’t remember the student’s name, but I had never heard from her before.” While this is a typical example of a known-item search task, the seeker provides only meta-information to retrieve the desired item, which has been shown to be effective for certain tasks [18]. The chance to process such requests successfully depends on the system’s ability to identify meta-information and enable the seeker to reveal further search constraints encoded as meta-information.

Example 3. Meta-information for a complex query.

Information		**Formal model**	
σ	"I'm looking for an email"	(1)	σ *sets* type
μ_1	"[the email] contains a link" Target: Query Concept: Facets [42] Intent: Constrain results	(2)	μ_1 *constrains* content
μ_2	"[the link points] to a research paper" Target, concept: same as μ_1 Intent: Elaborate on constraint	(3)	μ_2 *elaborates on* (2)
μ_3	"I [the seeker] got [the email]" Target, concept, intent: same as μ_1	(4)	μ_3 *sets* recipient
μ_4	"[the email] from a student" Target, concept, intent: same as μ_1	(5)	μ_4 *constrains* author
μ_5	"I [the seeker] got [the email] right after SIGIR last year" Target, concept, intent: same as μ_1	(6)	μ_5 *constrains* date
μ_6	"I [the seeker] had never heard from [the student] before" Target, concept, intent: same as μ_2	(7)	μ_6 *elaborates on* (5)
μ_7	"her [the student]" Target, concept, intent: same as μ_2	(8)	μ_7 *elaborates on* (5)

Compared to known-item search, exploratory search tasks exhibit an even greater reliance on the provider revealing meta-information. Revealing meta-information about the results of an initial, broad query may help seekers gain an overview, learn how the knowledge is structured, and identify directions for more focused requests. Indeed, current search engines already provide some level of support for such scenarios: query suggestions ("Others also searched for...") and similar features, such as Google's "People also ask" box that lists frequent questions related to a seeker's query, reveal new aspects to a topic. Meta-information, such as the number of obtained results or date ranges, can inform seekers about their query's success and the result's topicality. Conversational search can offer significant benefits in this regard, for example, by synthesizing summaries from the obtained results and suggesting further search avenues adapted to the seeker's interests.

Also, for everyday information-seeking tasks, which are currently served sufficiently well by a standard web search engine interface, the effective identification and interpretation of meta-information will become critical. Conversational search introduces "dialog settings with flexible communication channels, such as where a screen or keyboard may be inconvenient or unavailable" [21], where seekers may be unable to skim a search engine results page visually. In such situations, meta-information (be it provided implicitly or explicitly) can help seekers narrow down the result set. In turn, the expectations of seekers towards traditional user interfaces are raised.

In this paper, we systematically review and analyze the concept of meta-information with respect to conversational search. Section 2 reinterprets the related work in non-conversational from the perspective of meta-information. Section 3 does the same for conversation and conversational search, illustrating the variety of theories that fit into a meta-information framework. Section 4 contributes both a quantitative and a qualitative discussion of two conversational search datasets with respect to meta-information. Section 5 demonstrates the potential impact of meta-information on retrieval systems through a case study, analysis of challenges, and example use cases.

2 RELATED WORK

As briefly mentioned in the introduction, meta-information has been exchanged between retrieval systems and seekers long before conversational search came into the focus. Still, so far, meta-information has always been analyzed in the context of specific concepts, like document traits [10] or user context [29, 48]. The paper in hand generalizes these concepts by treating them as categories of particular instances of meta-information. In this view, the paradigm shifts in information retrieval (cf. Figure 1) introduced new targets of meta-information to the retrieval process, which, in turn, allowed the consideration of new concepts. The meta-information analyzed in classic information retrieval (considering just query and answer) has targeted either the collection's documents, their representation as results, the collection itself, or the query. Interactive information retrieval then focused on feedback. This change both facilitated and required to incorporate meta-information that targets the seeker and their information needs into retrieval process models.

The subsequent discussion categorizes the concepts from the literature by their primary target. Table 1 compiles a concise overview of the variety of meta-information concepts that have been operationalized for the two retrieval paradigms of non-conversational search. Though concepts are not strictly bound to a single target, the table illustrates the kind of meta-information to expect for each target. In practice, the provider can employ knowledge on the actual target to judge the reliability of meta-information (e.g., documents with different levels of trustworthiness), its expected scope or life-time (e.g., the target being the seeker vs. the query), or how to explain its use to the seeker (e.g., by echoing back targeted query terms).

2.1 Document-Related Meta-Information

The document has been the basic retrievable unit of information in non-conversational search. The way seekers use meta-information to identify relevant documents has been subject to inquiry for decades [10–12, 19, 34, 39, 61, 70, 71]. For example, even without explicitly being asked for it, seekers frequently mention that they use pieces of information to infer their judgments that are not necessarily part of the examined documents' content. Barry [10] grouped this meta-information into document traits (e.g., publication date, length, or document type), which seekers employed in 10.3% of their relevance judgments, and source traits (e.g., authors, organization, or publication), which seekers employed in 31.0% of their relevance judgments. Barry and Schamber [11] identify ten relevance criterion categories that seekers employ, for example, currency, quality of sources, and verification [11]. While the studies mentioned above focus on scholarly search, Tombros et al. [61] analyze relevance judgments for three kinds of generic web search tasks: search for background knowledge, search to make a decision, and search to compile a list of things. They compiled a list of 24 "document features" and grouped them into five categories, "content" being just one among them. See Saracevic [51] for an extensive treatment of relevance.

But even if a document seems topically relevant, seekers might reject it as they deem it not credible. In a survey of existing literature, Gînsca et al. [25] identify four aspects of credibility: the source's expertise and trustworthiness, as well as the content's quality (fitness to its purpose) and reliability (consistency over time). Conceptually, credibility is a meta-information that can be derived from these aspects that are in turn meta-information on the document. However, different seekers will see different meta-information as pertinent for credibility or make different judgments based on different experiences [44], which is why these aspects are also debated in philosophy, psychology, and sociology. This subjectivity makes credibility estimation a complex problem, and ranking algorithms like PageRank thus used a popularity estimate as a substitute [25]. Moreover, four types of credibility can be distinguished [24]: presumed credibility (based on general assumptions), reputed credibility (based on reports from others), surface credibility (based on superficial inspection), and experienced credibility (based on own experience). It is thus challenging for a retrieval system to learn how a seeker judges credibility. Even if a result scores high due to one type of credibility, the seeker might dismiss it for scoring low due to another type.

As is widely understood, the decision of whether or not a seeker inspects a document does thus not solely depend on the document. In one think-aloud study that illustrates this distinction exceptionally well, Wang and Soergel [70] analyzed the process in which seekers select documents from a results page for closer inspection. During the study, the participants spoke out why they would or would not inspect specific documents from a scientific publication database. Wang and Soergel model these decisions as a cognitive process: document information elements (which include document and source traits, but also the title and abstract, all of which are meta-information on the document) allow the seeker to judge the document on account of several user criteria (e.g., topicality or quality, the preference of which is meta-information on the seeker). In turn, the seeker uses these criteria in combination with the result representation (e.g., the snippet) to judge the document's value. User criteria and document values are described in the sections below for their respective main targets.

2.2 Result-Related Meta-Information

Frequently, seekers have not seen the retrieved documents beforehand and thus have to employ meta-information on search engine results pages to decide whether to inspect a document. In one study, Xie and Benoit [71] find that seekers do indeed employ additional elements from the results page for relevance judgments, namely the URL, highlighted keywords, and the result rank. For example, one participant in their study ruled out several results as their URL indicated them to be part of a dictionary website: the seeker believed that dictionaries only provide simple definitions instead of the detailed discussion they were seeking. Moreover, seekers frequently reformulate their queries based solely on meta-information provided on the search engine results pages. In a different study, Wang and Soergel [70] distinguish five different kinds of high-level values that inform the seeker's decision to inspect the document. The five values are taken from consumer research: The perceived utilities of a result to satisfy a desire for knowledge or information (epistemic value); to contribute to the task at hand (functional value); to be useful under hypothetical circumstances (conditional value); in association with a group or individual (social value); and to arouse feelings or affective states (emotional value).

These insights, however, have not been developed to their fullest potential to assist information seekers, as search engine interfaces are—on purpose and for good reasons—minimal. To facilitate the use of meta-information, Wang and Soergel [70] propose developing highly faceted interfaces that allow specifying as much information as possible. However, today's web search engines limit the presentation so as not to confuse their users. As we discuss in Section 3, conversational search allows for intuitive interfaces that still enable seekers to employ meta-information to full effect. Indeed, the well-known theoretical framework for conversational search describes a system that provides so-called partial items as responses to the seeker: not results, but so-called fields that describe and partition the set of currently retrieved documents [42]. This concept is already implemented in a voice-based search interface for visually impaired seekers, VERSE [68]. The system presents retrieved documents separated by source, additionally informing about the result count for each source. We discuss this system in more detail in Section 5.1. For another example, the args.me[5] argument search engine includes a visualization of the topic space of retrieved arguments. This visualization allows the seeker to see what aspects are covered by retrieved arguments and filter them by specific aspects [1].

[5]https://www.args.me/

Table 1. Concepts from the search literature that encode meta-information for non-conversational search, as discussed in Section 2. Definitions are adapted from the original publication for the sake of consistency.

Target	Concept	Concept definition and examples
Document (Section 2.1)	Document traits [10]	**Definition**: Characteristics that pertain to the physical format or actual publication **Examples**: Document length, document type (article, book, ...), date
	Source traits [10]	**Definition**: Characteristics that pertain to the intellectual source **Examples**: Authors, containing publication (journal, book, ...), editors
	Document features [61]	**Definition**: None provided **Examples**: Non-textual items, physical properties (file size, language, page already seen, subscription/registration), quality (authority/source, content novelty, error on the page, recency), structure (layout, links quality, table data), text (content)
	Credibility [24, 25, 44]	**Definition**: Various provided; judgement of both trustworthiness and expertise **Examples**: Expertise, quality, reliability, trustworthiness
	Document information elements [70]	**Definition**: The basic units that collectively represent a document to provide clues to the users' criteria (not all are meta-information to the document) **Examples**: Authors, affiliation, author's expertise, citation status, document length, document type, geographic location, journal, language, publication date, publisher
Result (Section 2.2)	Elements [71]	**Definition**: The individual components that participants examined during the evaluation of a result list or an individual result **Examples**: Keywords, number of results, pictures, rank, snippet, source, time/date
	Document values [70]	**Definition**: The basis for the document selection decision **Complete list**: Conditional, emotional, epistemic, functional, social
	Fields [42]	**Definition**: Aspects that can be used to describe a relevant item **Examples**: Cuisine, price range; someone I had not met, for a teenager, less gory
Collection (Section 2.3)	Meta-information resources [13, 14]	**Definition**: Describe the structure and content of information objects **Examples**: Classification scheme, thesaurus
	Metadata [32]	**Definition**: Elements or properties for use in resource description **Complete list**: Contributor, coverage, creator, date, description, format, identifier, language, publisher, relation, rights, source, subject, title, type
Query (Section 2.4)	Context terms [46, 47]	**Definition**: A term that is used within a query to describe the wanted document **Examples**: 19th century, album, best, definition, geology, Google, high school, PDF, troubleshooting, tutorial, Wikipedia
	Extra-topical dimensions [7]	**Definition**: Constraint that is independent from the information need's topic **Complete list**: Domain knowledge, viewpoint, experiential, venue location, source location, temporal
	Query strategies for extra-topical preferences [7]	**Definition**: The way people request information matching a specific type of extra-topical preference **Examples**: Absolute time range, all sides, author, complexity, discourse, exclude relative location, genre, landmark, proximity, purpose, state
	Cognitive Search Intents [28]	**Definition**: Needs for the cognitive characteristics of documents to be retrieved **Complete list**: Exhaustive, comprehensible, objective, subjective, concrete, abstract
	Facets [42]	Identical to "Fields [42]" of a result (Section 2.2), but specified by the seeker
Information need (Section 2.5)	Categories [16, 17, 22, 50, 57, 69, 74]	**Definition**: Type of intent of the query **Examples**: Navigating, informing, transacting, re-finding, entertaining, arguing, being creative, making meaning
	Information seeking strategies [14]	**Definition**: Method, goal, mode, and resource considered when seeking information **Complete list**: All 16 combinations of to scan vs. to search, to learn vs. to select, to recognize vs. to specify, with information vs. with meta-information
Seeker (Section 2.6)	User criteria [70]	**Definition**: Used to evaluate alternatives **Complete list**: (Preference for) authority, availability, discipline, novelty, orientation/level, quality, reading time, recency, relation/origin, special requisite, topicality
	User context [29, 48]	**Definition**: Context encompasses any information for defining the user's situation. A situation is an instance of the contextual information available **Complete list**: Task, social, personal, spatio-temporal, environmental
	Affective/physiological/behavioral features [5, 6, 38]	**Definition**: None provided **Examples**: Facial expression, heart rate, skin temperature, skin conductivity, neural activity, dwell time

2.3 Collection-Related Meta-Information

Interactions of the information seeker with a traditional search engine can broadly be categorized into either directly interacting with items (i.e., results) or interacting with collection-related meta-information resources (e.g., a categorization system) [13, 14]. For example, the ACM Computing Classification System,[6] allows for learning about the topics inside a research field without looking at the actual publications. Such resources are precious when browsing a collection (in the example: the ACM Digital Library), as they help the seeker understand the structure of the collection and thus allow for narrowing the browsing scope. In this context, the Dublin Core Metadata Initiative [32] can be seen as a standardization initiative for meta-information resources.

2.4 Query-Related Meta-Information

Several researchers noticed that also queries transport meta-information. Russell [46, 47] notes that queries sometimes contain "context terms," which do not specify terms that should occur in a document for its retrieval but rather describe the wanted document. A query might contain the context term "PDF" to describe the expected result file type. This way of specifying meta-information works well for regular search engines as most context terms do occur in the documents. For example, the context term "Wikipedia," which implies a need for a Wikipedia page, does indeed match the page name on Wikipedia pages.

[Example 4] The following example from the Spoken Conversational Search dataset [62–64][7] demonstrates how a seeker uses context terms to restrict the search results. The seeker wishes to find out what uses old tires can be put to when recycled. Not satisfied with the results so far, because they include many wreckers and tire dealers from which the seeker does not expect to obtain the desired information, the conversation continues as follows:

Dialog 1. SCS dialog excerpt on uses for old car tires.

Turn	Role	Transcript
7	Seeker	Actually, can I add something else to that?
8	Provider	Yeah, sure.
9	Seeker	Can we have "NOT"—in caps—"wreckers", and "NOT sales"... actually, just "NOT wreckers" first and let's see if we can use one of them.
10	Provider	OK, with the "NOT wreckers" in the search we actually get a whole lot of wreckers, yes... so the first five or six results all are for wreckers.
11	Seeker	Surely we should not be getting wreckers though we have "NOT"... uhm, maybe just remove the word "wreckers".

The seeker wishes to modify their query with a context term to exclude results from wreckers. However, the use of the specified context term ultimately fails; perhaps this was due to lack of support for the desired Boolean search operator from the employed search engine, incorrect usage by the provider (who probably should have silently swapped the "NOT" for the more frequently used minus sign: "−wreckers"), or simply a showcase of the hit-or-miss nature of using context terms. Below our formalization of Dialog 1: the meta-information μ constrains the results' contents.

[6]https://dl.acm.org/ccs

[7]Despite its focus on conversational search, the setting in which the data was collected (discussed in Section 4.1) allowed for the use of methods more closely associated with traditional search interfaces, such as context terms.

Example 4. Modeling context terms as meta-information.

Information		Formal model	
σ	Web pages on uses for old car tires	(1)	σ *sets* topic
μ	"not wreckers" Target: Query Concept: Context terms [46, 47] Intent: Constrain results	(2)	μ *constrains* content

However, context terms cannot fully express the range of available meta-information. So how do seekers cope with such constraints? Kato et al. [28] and Arguello et al. [7] analyzed query formulation when facing this problem. Kato et al. investigate how cognitive search intents are expressed in search queries intended to affect the cognitive characteristics of the retrieved documents. However, their study finds that about half of the participants did not use query terms related to cognitive search intents, concluding that this is likely caused by current web search engines' lack of processing capabilities for such terms. Arguello et al., on the other hand, employ a selection of six so-called extra-topical dimensions, which they define as a type of constraint that is independent of the information need's topic; for example, looking for experiential anecdotes or information from a specific point of view. Participants in their study used a variety of strategies to implement these constraints in their queries. However, they report that many of these strategies lead to reduced retrieval effectiveness for regular search engines.

Theoretical results indicate that conversational search will allow seekers to employ meta-information in queries much more effectively. Radlinski and Craswell's [42] widely accepted theoretical framework for conversational search sees faceted elicitation as one primary use case for conversational search: Seekers specify facets, which correspond to meta-information of any kind, to narrow down the results. Example 3 illustrates the variety of meta-information a seeker could use. From the point of view of meta-information, these "facets" are indistinguishable from the "fields" that the search engine indexes (cf. [42] and Section 2.1). Not discussed by Radlinski and Craswell, this observation hints at the possibility—or maybe even the need—of conversational search engines to move their knowledge representation closer to that of human seekers.

2.5 Information-Need-Related Meta-Information

When the seeker engages with an information retrieval system, they do so to gain information. But different from a database lookup, uncertainty is paramount in information retrieval. Probably the most significant uncertainty lies within what the seeker wants to know, as they might not know it at the start, and communicating it to the provider is a complicated process [56]. It is thus helpful to contrast the information need, the "need behind the query" [16], to the query itself, the "compromised need" [56] discussed above.

The literature identifies several different categories for the information need. Categories include navigating, informing, and transacting [16], but also re-finding [57], entertaining [22], arguing [69], being creative [74], or making meaning [50]. Due to increasing capabilities of commercial search engines, recent query logs also show requests for a weather forecast or calculations [17]. Clearly, search systems that support several of these categories must be able to distinguish these. For the classical category of informing, several so-called information-seeking strategies can be further distinguished: does the seeker browse a collection or search it; try to learn about something or to find specific items; try to recognize some item or specify it, and look directly for information or investigate meta-information resources? All 16 combinations of these four binary attributes correspond to one specific information-seeking strategy [14]. Like the categories of needs, information about the employed strategy can allow the provider to better adapt to the seeker's need.

2.6 Seeker-Related Meta-Information

As discussed above, other seekers may judge the same document's relevance to the same query or information need pretty differently as they have other preferences. From observations in their study on scientific search, Wang and Soergel [70] identified eleven such user criteria and linked them to meta-information on the document. For example, the document's recency is informed by its publication date and the quality by its publishing journal, author, and citation status.

In practice, the use of meta-information depends not just on long-term preferences but also heavily on the current seeker context [10, 13], which presumes the repeated collection of personal data about them. In the extreme case, "any information describing a document will, for some users in some situations, contain useful clues about that document." [10]. Contextual information retrieval [48] operationalizes this observation by analyzing which information about a seeker's current situational context renders a document more or less relevant. Contextual information retrieval has become a major branch of information retrieval, investigating the kinds of contextual information about seekers that the provider can exploit.

In Table 1, we list a categorization of sources of contextual information about seekers as per Kofod-Petersen and Aamodt [29]. The task context refers to information about a seeker's goal or task that led to information behavior; for instance, in various TREC tracks, the narratives have sometimes been used to augment the topics' queries to investigate whether this meta-information about the information need is helpful. The social context of seekers refers, for example, to family, friends, and co-workers, with whom one collaborates actively or whose past information behavior logs the provider might exploit to improve future search results in a given context. The personal context of seekers includes physiological and mental contexts. The former refers to meta-information like height, weight, age, and physical ability, the latter to mood, expertise, and personal interests.

[Example 5] Perhaps one of the most widespread contextual information retrieval applications is exploiting the spatio-temporal context of seekers to adapt search results to the occasion. For example, asking "What's the closest place where I can get pizza?" implicitly specifies that restaurants are to be ranked based on their proximity to the seeker.

Example 5. Modeling user context as meta-information.

	Information		**Formal model**
σ	"What's [a] place where I can get pizza?"	(1)	σ *sets* topic
μ_1	GPS coordinates of seeker's device Target: Seeker Concept: User context [29, 48] Intent: Locate seeker	(2)	μ_1 *sets* (seeker) location
μ_2	"the closest place" Target: Query Concept: Extra-topical dimensions [7] Intent: Rank results	(3)	μ_2 *sets* ranking *to* distance to (seeker) location

Other meta-information from environmental context includes both physical properties of the surroundings (e.g., lighting) and social ones, like who is nearby and whether it is appropriate to return specific results when others might overhear them.

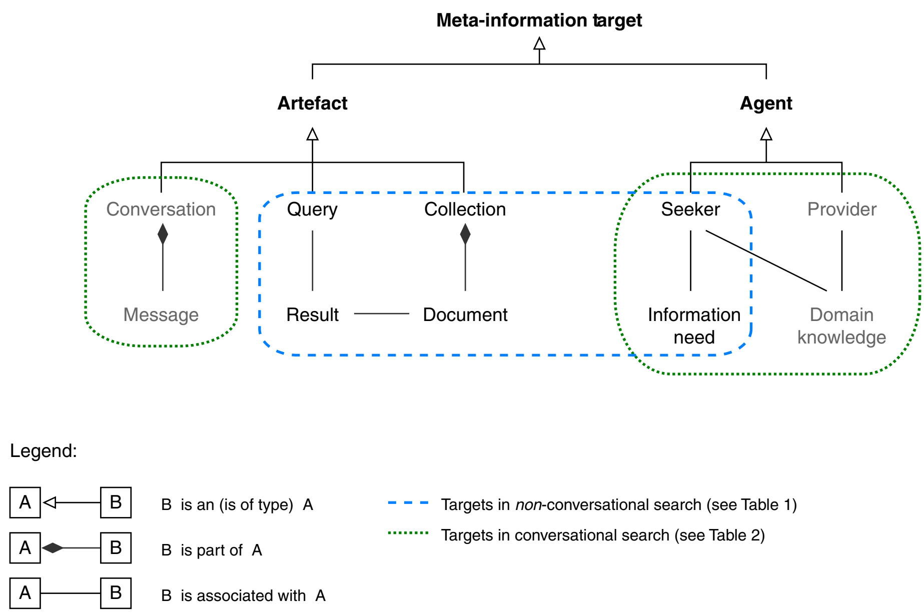

Fig. 2. Taxonomy of meta-information targets identified in Sections 2 and 3.

Most meta-information from the seeker's context is available before or during retrieval. However, important aspects can emerge post-retrieval as part of the seeker's reaction to the presented results. Research into implicit relevance feedback has attempted to use these post-retrieval aspects as meta-information for improving the result set relevance. The efforts expanded from the use of behavioral features, such as dwell time, to include affective features, such as the seeker's facial expression [5, 6], and physiological features, such as heart rate or neural activity [38].

Information retrieval in context has evolved into a widespread branch of information retrieval. Surveying all contextual factors that the literature has been exploring is beyond the scope of this paper. However, it is clear that conversational search, especially when considering audio-only interfaces, requires contextual meta-information as a matter of course.

3 SYSTEMATIZING META-INFORMATION IN CONVERSATIONAL SEARCH

This section reviews and analyzes nine relevant theories of conversation and conversational search from the literature. Our goals are to identify and reconcile the targets pertinent to conversational search, organize them alongside the targets of classic and interactive information retrieval, and derive from the theories the relevant concepts that encode meta-information.

To cut a long story short, Figure 2 shows in a taxonomy of meta-information targets the ones considered in classic and interactive information retrieval (cf. Figure 1 and Table 1), and the additional ones due to the new paradigm of conversational search (cf. Table 2). In classic information retrieval, the query has been the central piece of information for the provider to understand the seeker's information need. In interactive information retrieval, on the other hand, the seeker became the focus of attention. Conversational search places the provider on the same level as the seeker, allowing for a dynamic back-and-forth between equal partners, much like in information-seeking conversations between humans. Collaboration on this level requires a mutual understanding among the partners and thus requires revealing to each other their capabilities, knowledge, and maybe emotions and personality. The seeker describes their need by exchanging messages throughout several turns to verbalize what they do not know—the provider helps form the picture of the information need. Conveniently for the seeker, the provider may directly embed pieces of information into the conversation instead of forwarding seekers to external documents of which only small parts might be relevant. This development blurs the boundaries of results and documents in messages, making them less prevalent as targets for meta-information in the conversation. Similarly, this development blurs the boundaries of queries, making it a challenge to extract them from the conversation, just like in conversational question answering [3].

Meta-information appears to be especially suited for modeling the process of a conversational search, considering the three characteristics of meta-information introduced in the introduction. (1) Meta-information is only identifiable as such by its relation to other information, which matches the concepts described in both Section 2 and this section, which essentially provide meta-information describing observations or predicting events. (2) This relation is dynamic, which matches the dynamic nature of a conversation, where participants interpret what is being said and may re-interpret previous statements as the information-seeking conversation gains clarity. (3) The semantics of meta-information depends on both the referred information and the recipient, which matches a conversation where both partners can take the initiative to ensure mutual understanding.

In what follows, we review existing meta-information concepts that pertain to *conversational* search as per Figure 2. Table 2 compiles a brief overview of the variety of meta-information concepts that have been operationalized for the paradigm of conversational search. As the table shows, Radlinski and Craswell's "Theoretical Framework for Conversational Search" [42] comprises most targets. This observation is not surprising, given that the framework's stated goal is to "capture the desirable properties of conversation specifically from an information retrieval perspective." The concepts introduced in their framework are essential to our discussion, so that we briefly recap them here before analyzing them in more depth below: the *memory* of past messages for a consistent conversation; *mixed initiative* for proactive collaboration from both partners; *set retrieval* for a combination of information items from different sources; *system revealment* for informing seekers of the provider's capabilities and knowledge; and *user revealment* for assistance in formalizing the information need. As discussed below, the presence of these properties either constitutes meta-information or facilitates the use of meta-information in a conversation.

3.1 Message-Related Meta-Information

Conversational search is an exchange of messages: the seeker transmits one message to the provider, who replies with another message. In classic information retrieval, the seeker's messages are queries, and the provider's messages are search engine results pages. A "message" can take many forms, though, including text, sound, and gesture. This setup corresponds to the Mathematical Theory of Communication [52] (cf. Figure 3). Therefore, the theory's three levels of communication problems apply: accurate transmission of symbols, precise conveyance of meaning, and intended effect on the receiver. The first problem is barely discussed in information retrieval, as it has some importance in voice search only. Intuitively, meta-information arises at the second level, where the receiver interprets the message. Yet communication problems at the third level also involve meta-information, when the intent is to have the receiver form some connection in their mind—often the point and purpose of the conversational search interaction.

Messages serve thus several purposes in conversational search. For example, they serve as the collective memory [42], to which both partners can refer later on. Most of the meta-information concepts discussed in Section 2 could serve as the point of reference in such a case. For instance, consider "What was the thing that annoyed me so much?" Such analysis has to deal with the ambiguities of human language and thus benefits from methods like term disambiguation (e.g., [35])

Table 2. Concepts from the search literature that encode meta-information for conversational search, as discussed in Section 3. Definitions are adapted from the original publication for the sake of consistency.

Target	Concept	Concept definition and examples
Message (Section 3.1)	Levels of communication problems [52]	**Definition**: Stages in message reception in which communication problems occur **Complete list**: Inaccurate transmission of symbols, inprecise conveyance of meaning, unintended effect on the receiver
	Memory [42]	**Definition**: References to past statements **Examples**: By affective state, by complexity, by content, by speaker, by time
	Disambiguation clues [27, 35]	**Definition**: Signals that indicate the intended meaning or segmentation **Examples**: Associated entities, pauses, pronounciation, semantically related words
Conversations (Section 3.2)	Mixed initiative [42]	**Definition**: One agent takes initiative from the other **Complete list**: Seeker, provider
	QRFA state [67]	**Definition**: Who sent the message and whether they had initiative at that time **Complete list**: Query (seeker, proactive), request (provider, proactive), feedback (seeker, reactive), answer (provider, reactive)
	Conversational roles [53]	**Definition**: The roles that conversants take, the dialog acts they use, and the states these acts bring the conversation to **Examples**: Asker, answerer; request, reject offer, withdraw commissive; State 1 to 11
	Context spaces [43]	**Definition**: A fine-grained topical unit of the conversation with specific attributes **Examples**: Answer, claim, comparison, illustration, offer, question
	Conversational implicature [26]	**Definition**: An obvious failure to be cooperative in order to imply something else **Examples**: Insufficient message, overly verbose message, incorrect content, irrelevant content, bad manners
Domain knowledge (Section 3.3)	Set retrieval [42]	**Definition**: Reasoning about the utility of sets of complementary items **Examples**: Flight and hotel bookings, how-tos and knowledge requirements, smartphones and headphone plugs, software and operating systems
	System revealment [42]	**Definition**: The provider reveals to the seeker its capabilities and knowledge **Examples**: ACM Digital Library, argumentative reasoning, explanations of own actions, history knowledge, understanding irony, Wikidata
Information need (Section 3.4)	Anomalous state of knowledge [15]	**Definition**: Recognition that the seeker's state of knowledge is inadequate for resolving their problem **Examples**: Missing a connection, lack of the right terms, using the wrong tool
	User revealment [42]	**Definition**: Assistance to express or discover the seeker's information need and long-term preferences **Examples**: Providing examples, explanations, suggestions
Seeker (Section 3.5)	Principle of uncertainty [31]	**Definition**: The seeker's feelings, thoughts, and actions as per the phases of the information search process **Examples**: Initiation (uncertainty, vague thoughts, seeking relevant information), collection (confidence, focused thoughts, documenting)
Provider (Section 3.6)	Conversation action space [42]	**Definition**: Categories of actions taken and accepted by the provider **Examples**: Providing nothing, a partial item, a complete item; responding with a preference, rating, critique, free text
	Explicit persona traits [73]	**Definition**: A verbalized character trait that influenced a specific message **Examples**: "I'm always serious," "I'm better safe than sorry," "I like trash movies"
	Personality manifestations [33]	**Definition**: Perceptions of a system's personality based on its actions **Examples**: Warmth, competence; extroversion, feeling, sensing, judging

and segmentation (e.g., [27]). In some cases, messages can contain clues for the disambiguation of other messages, which, too, can be modeled through meta-information.

[Example 6] Imagine that the seeker asks the provider: "Could you remind me of the technical term you used yesterday?" To answer such a question, providers need to consider the meta-information of past messages: What message did the provider send yesterday containing a technical term?

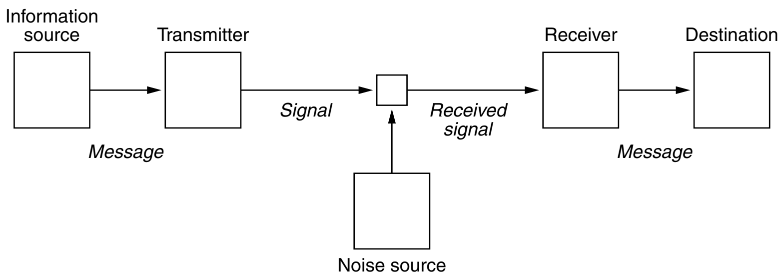

Fig. 3. Symbolic representation of a communication system as per Shannon and Weaver [52]. In conversational search, the seeker and provider take turns on being the information source and destination.

Example 6. Modeling memory as meta-information.

Information		**Formal model**	
σ	"Could you remind me of the [...] term [...]?"	(1)	σ *sets* topic
μ_1	"the technical term" Target: Message Concept: Facets, Memory [42] Intent: Constrain results	(2)	μ_1 *constrains* (message) content
μ_2	"the [...] term you [the provider] used" Target, concept, intent: same as μ_1	(3)	μ_2 *sets* (message) speaker
μ_3	"the [...] term [...] used yesterday" Target, concept, intent: same as μ_1	(4)	μ_3 *sets* (message) date

3.2 Conversation-Related Meta-Information

In conversational search, the term "conversation" refers to an exchange of information between the seeker and the provider over a longer period of time [42]. Concepts that primarily target the conversation carry information on the current "state" of the conversation or are deduced from the conversation even though it is never explicitly said. As a part of its state, the conversation's topic can take the place of the query, and thus the concepts that target the query (cf. Section 2.4) could be said to target the conversation in conversational search. But also other concepts target the conversation. In the following, we first discuss concepts that are already prominently discussed in the conversational search literature, and then highlight examples of concepts that could be adapted for conversational search from the generic conversation and dialog literature. To illustrate these different concepts of meta-information, Dialog 2 shows a short excerpt from the start of a conversation on "airport security" from the Spoken Conversational Search (SCS) dataset [62–64] (cf. Section 4 for a thorough discussion of this dataset). Both seeker and provider are humans. The seeker has a specific information-seeking task and communicates voice-only with the provider, who uses a web search engine to answer the seeker's requests.

Dialog 2. SCS dialog excerpt on "airport security".

Turn	**Role**	**Message**
1	Seeker	Can you type in, uhm, "effective"... "effectiveness of new security measures at airports"?
2	Provider	Australia, or is it just airports?
3	Seeker	Put, uhm, "international"... "international airports".
4	Provider	OK, the first one that comes up is [...]

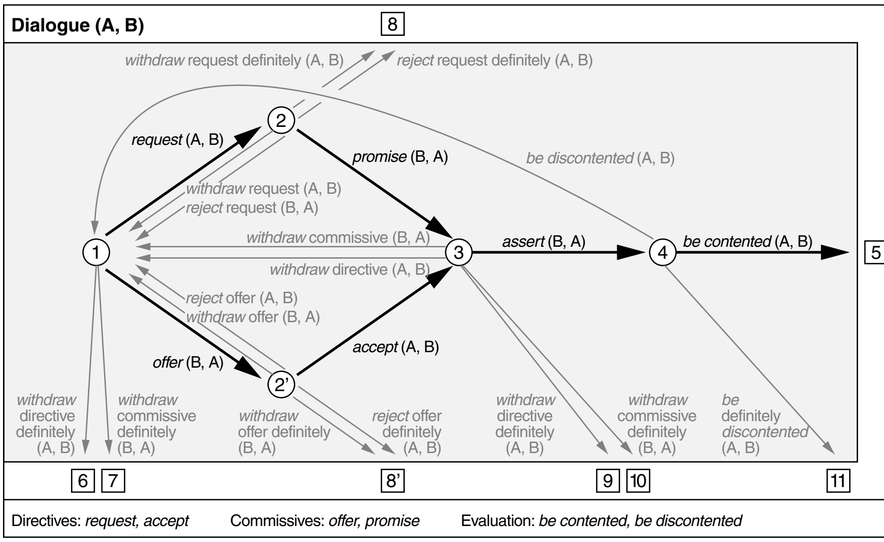

Fig. 4. Conversational roles schema for one question in information-seeking conversations as per Sitter and Stein [53]. The numbers in circles depict states, the ones in squares end states of a dialog cycle, the arrows depict usual state transitions, and the dark arrows expected state transitions.

As Figure 1 illustrates, conversational search enables the provider (the search engine) to take the initiative to request information from the seeker ("mixed initiative" [42]). Both provider and seeker thus have to be aware of the state of the conversation. In Dialog 2, the provider takes the initiative in Turn 2 by not answering the question but asking a question in return. The initiative is then restored to the seeker in Turn 3 when they answer the provider's question.

We here highlight two well-known theories that introduce the concept of a "conversation state" specifically for information-seeking conversations: QRFA [67] and COR [53]. The QRFA model [67], on the one hand, just distinguishes who sent last the message and who had the initiative at that time (Figure 1, bottom), which result in four states (in addition to a start and end state): At the *query* state, the seeker asked the provider for either information or to perform some action; at the *request* state, the situation is the same, but that the provider asked the seeker; at the *feedback* state, the seeker reacts either positively or negatively to what the provider asked or provided; and at the *answer* state, the provider reacts to the seeker by providing an answer or confirmation. Despite its simplicity, QRFA differentiates between well-known interaction patterns (e.g., iterative query or answer refinement), and it indicates odd behavior, such as when the provider makes a request and immediately follows with an answer. The conversational roles model (COR, [53], cf. Figure 4), on the other hand, treats information-seeking conversations as a sequence of dialog acts. These acts are "illocutionary," in that an action is performed through them by speaking, and they let the seeker and the provider dynamically take respective roles (e.g., the role of an asker), and put the other party into the corresponding role (the role of an answerer). The speaker's current role and the conversation's current state constitute meta-information for the information exchanged in the conversation, as they limit the interpretations of consecutive dialog acts. The roles and state are induced from the dialog acts, which are modeled as meta-information on the respective message.

[Example 7] The table below illustrates both QRFA and COR for Dialog 2. Note that the rejections for COR in (9) and (12) are implicit but deductible from the explicit acts that follow them.

Example 7. Modeling conversation state with QRFA and COR as meta-information for Dialog 2.

Information		**Formal model**	
σ	Seeker engages in conversation with provider	(1)	σ *sets* state$_{\text{QRFA}}$ *to* start
		(2)	σ *sets* seeker role *to* A
		(3)	σ *sets* provider role *to* B
		(4)	σ *sets* state$_{\text{COR}}$ *to* 1
μ_1	"Effectiveness of new security measures at airports?"	(5)	μ_1 *sets* state$_{\text{QRFA}}$ *to* query
	Target: Conversation	(6)	μ_1 *sets* dialog act *to* request
	Concept: QRFA state [67], conversational roles (COR) [53]	(7)	(6) *sets* state$_{\text{COR}}$ *to* 2
	Intent: Request information		
μ_2	"Australia, or is it just airports?"	(8)	μ_2 *sets* state$_{\text{QRFA}}$ *to* request
	Target, concept: same as μ_1	(9)	μ_2 *sets* dialog act *to*
	Intent: Offer specification		(reject request), offer
		(10)	(9) *sets* state$_{\text{COR}}$ *to* 2'
μ_3	"International airports"	(11)	μ_3 *sets* state$_{\text{QRFA}}$ *to* query
	Target, concept: same as μ_1	(12)	μ_3 *sets* dialog act *to*
	Intent: Constrain results		(reject offer), request
		(13)	(12) *sets* state$_{\text{COR}}$ *to* 2
μ_4	"OK, the first one that comes up is [...]"	(14)	μ_4 *sets* state$_{\text{QRFA}}$ *to* answer
	Target, concept: same as μ_1	(15)	μ_4 *sets* dialog act *to*
	Intent: Provide answer		promise, assert
		(16)	(15) *sets* state$_{\text{COR}}$ *to* 4

The concept of a "conversation state" exists beyond the search literature, and some theories take a different approach than the one above. To hint at the possibilities for modeling state, we here detail the theory of context spaces by Reichman [43], which uses a stack of such spaces as a model of the current state in a discourse. A context space corresponds to a fine-grained topical unit of the conversation. For an information-seeking conversation, a context space may correspond, for example, to a raised question or an elaboration of that question.[8] At all times during the conversation, a single context space is *active* and thus in the conversation's focus. Other statuses are: *controlling* for the (maximum of one) context space that serves as the reference point for the currently active space (e.g., if a context space that elaborates on another context space is active, the context space being elaborated on would be controlling); *closed* for a context space for which a conclusion is reached; and *open* for a context space that is not yet closed but neither active nor controlling (e.g., if the seeker asks for a definition of some term that the provider just used in an answer, the original question would become open if it is not yet closed). One motivation for the use of context spaces is to comprehend the usage of coreferences in human language [43], which also poses a challenge to conversational search systems: for example, knowing the active context space is essential to differentiate between "this" and "that."

[Example 8] To illustrate the use of context spaces as meta-information, we show a possible formalization for Dialog 2 with context spaces. The context space types in the example are only a first attempt at adapting context space types for information-seeking conversations. A more thorough investigation is necessary, for which the conversational roles model discussed above might serve as a starting point. The dialog acts used there are similar to the "conversational moves" that Reichman used to develop the context space types for argumentative discourse.

[8] Reichman focuses on discourse as a special case, with context spaces corresponding to raised issues or support arguments.

Example 8. Modeling conversation state with context spaces as meta-information for Dialog 2.

Information			**Formal model**
σ	Seeker engages in conversation with provider		
μ_1	"Effectiveness of new security measures at airports?" Target: Conversation Concept: Context spaces [43] Intent: Request information	(1)	μ_1 *creates* C_1 *as* question *with* C_1.status *set to* active *and* C_1.topic
μ_2	"Australia, or is it just airports?" Target, concept: same as μ_1 Intent: Offer specification	(2)	μ_2 *creates* C_2 *as* offer *with* C_2.status *set to* active *and* C_2.suggestion
		(3)	(2) *sets* C_1.status *to* controlling
μ_3	"International airports" Target, concept: same as μ_1 Intent: Constrain results	(4)	μ_3 *sets* C_1.topic
		(5)	μ_3 *sets* C_2.status *to* closed
		(6)	(5) *sets* C_1.status *to* active
μ_4	"OK, the first one that comes up is [...]" Target, concept: same as μ_1 Intent: Provide answer	(7)	μ_4 *creates* C_3 *as* answer *with* C_3.status *set to* active *and* C_3.content
		(8)	(7) *sets* C_1.status *to* controlling

Other theories on conversation focus on different aspects than the state. For example, the cooperative principle [26] states a set of maxims for a cooperative conversation: say as much as you need, but no more; say what is true; say what is relevant; and have good manners. The principle is relevant for conversational search, which one usually assumes to a cooperation of seeker and provider. An apparent failure to be cooperative—to fulfill a maxim—, can often be interpreted as the speaker implying something different than the literal meaning of their utterance. This situation gives rise to conversational implicature [26], where a speaker implies a proposition q through stating another proposition p. However, the speaker has to believe that q is a requirement for them stating p and that the listener knows or can figure out this belief of the speaker. Such implicatures frequently occur in human conversations, not least with irony (which usually violates "say what is true"). As a different example, the seeker may be too lengthy (from the provider's perspective) and talk about what comes to their mind rather than asking a question. Such a situation may imply that the seeker is not yet able to formulate a focused question (cf. meta-information on the information need in Section 3.4), implicitly asking the provided for assistance. Tracking the maxims' fulfillment as meta-information is the first step to understand them. Most likely, however, not all human seekers fulfill these maxims in a "conversation" with an AI provider, especially "have good manners." One plausible reaction of a system that is capable of understanding conversational implicature to such a "failure" on the human's part could be to stop using conversational implicature itself, as this might be closer to the behavior the seeker expects. Search systems that adopt a personality might demonstrate their capabilities and choose a response that implies they understood the impoliteness.

[Example 9] Dialog 2 contains two subtle but common violations of "say what is relevant:" neither the message of Turn 2 nor that of Turn 3 answer the corresponding question. The message in Turn 2 does not answer but offers a specification of the question, thereby conversationally implying that the question is too general for the conversation. The message in Turn 3 does not select one of the offered choices but a different one, thereby conversationally implying that the choices are not suited. These violations are so common in conversations that a human reader might resolve them without further notice, but a machine might have difficulties making these connections.

Example 9. Subtle violations of "say what is relevant" in Turns 1 to 3 of Dialog 2.

Information		**Formal model**	
σ	"Effectiveness of new security measures at airports?"	(1)	σ *requests* answer
μ_1	"Australia, or is it just airports?" Target: Conversation Concept: Conversational implicature [26] Intent: Request specification	(2) (3) (4)	μ_1 *requests* answer (2) *is not relevant to* (1) (3) *implies that* (2) *offers* elaboration *for* (1)
μ_2	"International airports" Target, concept: same as μ_1 Intent: Constrain results	(5) (6) (7)	μ_2 *sets* topic (5) *is not relevant to* (2) (6) *implies that* (5) *closes* (4) *instead*

3.3 Domain-Knowledge-Related Meta-Information

In conversational search, knowledge is exchanged both ways in the messages of the conversation. To live up to the ideal of human conversation, the provider needs to extract information from documents and cite or embed it in their messages. Such an embedding allows the seeker to get to the information they are seeking more efficiently. For the provider, document-related meta-information (cf. Section 2.1) is thus subsumed in domain-knowledge-related meta-information for conversational search. Other concepts that target domain knowledge encompass relationships between documents or entities mentioned in the documents, including collection-related meta-information (cf. Section 2.3). For conversational search, however, the same concepts that target the provider's domain knowledge can as well be applied to the seeker. Dialog 3 shows the beginning of an imaginary search for which the messages illustrate plausible uses of domain knowledge.

Dialog 3. Imaginary text-based dialog on "DIY music server". Underlining indicates links.

Turn	**Role**	**Message**
1	Seeker	I want to build myself a small web server to play my music collection. It should be available from my WiFi only. Can you help me?
2	Provider	Most people suggest to use a Raspberry Pi for that. I found this how-to for programming a music server. But you need to know how to program a Raspberry Pi. The book "Programming the Raspberry Pi: Getting Started with Python" is the best-rated one for learning to program a Raspberry Pi. And if you tell me your router name, I can look for a how-to for setting up the Raspberry Pi in your WiFi.
3	Seeker	Oh, I have already read that book!
4	Provider	Noted! But in this case you might be more interested in this how-to titled "DIY Raspberry Pi 4 Music Server in 30 Minutes," as it targets more advanced users.

[Example 10] The first use of domain knowledge for the provider is to constrain the search results. Domain knowledge can help to rule out specific results, to reason about result utility, and thus rank them and give direct suggestions. The first sentence of Turn 2 of the Dialog 3 shows such reasoning. Since μ_1 and μ_2 fulfill the same role here, they are modeled identically in this regard—though other concepts differentiate between the seeker's restrictions and the provider's suggestions.

Example 10. Utilization of domain knowledge to constrain results in Turns 1 and 2 of Dialog 3.

Information		Formal model	
σ	"build myself a small web server to play my music collection"'	(1)	σ *sets* topic
μ_1	"[the server] should be available from my WiFi only' Target: Conversation Concept: Facets [42] Intent: Constrain results	(2)	μ_1 *elaborates* (1)
μ_2	"use a Raspberry Pi for [the server]" Target: Domain knowledge Concept: Fields [42] Intent: Constrain results	(3)	μ_2 *elaborates* (1)
μ_3	"[use a Raspberry Pi as] most people suggest" Target: Domain knowledge Concept: Metadata [32] Intent: Justify constraint	(4)	μ_3 *justifies* μ_2

[Example 11] The second use of domain knowledge for the provider is to provide information to the seeker. For conversational search, this can extend far beyond listing results. Even today, commercial search engines enrich their results pages with various other result elements (or "information items") like definitions, translations, encyclopedic knowledge, graphs, or widgets. A conversational search system will heavily employ such result formats and extend it with the capability to reason about the utility of complementary information items, which Radlinski and Craswell call the property of set retrieval [42]. Turn 2 of Dialog 3 continues with a seeming solution to the initial problem. But the provider then uses its domain knowledge to point out new problems that arise from this solution: knowing how to program a Raspberry Pi and how to configure the WiFi. The provider thus shifts the conversation first to one and then to the other new problem.

Example 11. Utilization of domain knowledge in Turn 2 of Dialog 3. Continuation of Example 10.

Information		Formal model	
μ_4	"this how-to [. . .] for programming a music server" Target: Domain knowledge Concept: Document information elements [70] Intent: Provide answer	(5)	μ_4 *closes* (1)
μ_5	"You need to know how to program a Raspberry Pi [for the task]" Target: Domain knowledge Concept: Set retrieval [42] Intent: Describe missing part	(6)	μ_5 *shows gap in* μ_4
μ_6	"The book [enables] learning to program a Raspberry Pi" Target, concept, intent: same as μ_4	(7)	μ_6 *closes* (6)
μ_7	"[you need] a how-to for [. . .] your WiFi" Target, concept, intent: same as μ_5	(8)	μ_7 *shows gap in* μ_4
μ_8	"If you tell me your router name, I can look for a how-to" Target: Conversation Concept: Facets [42] Intent: Request specification	(9) (10)	μ_8 *requests* knowledge (9) *may close* (8)

[Example 12] However, also the seeker might want to give information about their domain knowledge to the provider, to inform the provider about what they already know. They might give single pieces of information, similar to the concept of atomic "information nuggets" in documents [40]. Or, like in Dialog 3, they could state that they know an entire document.

Example 12. Revealment of the seeker's knowledge in Turns 3 and 4 of Dialog 3. Continuation of Example 11.

Information		**Formal model**	
μ_9	"I have already read that book" Target: Domain knowledge Concept: Seeker knowledge Intent: Constrain results	(10) (11)	μ_9 *sets* (seeker) knowledge (10) *elaborates* (1)
μ_{10}	"this how-to [. . .] targets more advanced users" Target: Domain knowledge Concept: User criteria [70] Intent: Improve answer	(12)	μ_{10} *improves on* μ_4

To work effectively and efficiently with a conversational search system, the seeker needs to know the provider's capabilities. For instance, Turn 1 of Dialog 3 shows a strong faith in the provider's capabilities. Thus the provider should put the effort in both advertising and demonstrating capabilities, a process called "system revealment" [42]. For an effective revealment, on the other hand, the provider must keep track of what has been revealed so far and when. They can achieve such bookkeeping through modeling the conversation with meta-information. For example, μ_{10} reveals the capability of judging the seeker's experience level for a topic.

3.4 Information-Need-Related Meta-Information

One of the central promises of conversational search systems is to assist the seeker with problems they can not formalize, yet [42]. Taylor [56] hypothesizes that information needs occur at four levels: (1) the visceral, unexpressed need, which leads to (2) a conscious need, which, however, has to be turned into (3) a formalized need to provide a question that can be answered. As Taylor notes, knowledge can play an essential role in determining the seeker's question. For various reasons, the seeker might not communicate a formalized description at Level 3 to the provider, but (4) a description that has been compromised to fit the provider's capabilities, for example, a query. Traditional search engines work with the query at Level 4 they get, but, in conversations, especially Levels 2 and 3 are relevant: Taylor specifically notes conversations as a way to get from a description at Level 2 as "an ambiguous and rambling statement" to a "properly qualified and rational statement" at Level 3. Cases where the seeker needs assistance to reach a Level 3 statement are not uncommon, occurring even in known-item search [8]. The provider can detect the current stage in this transition within the conversation from the seeker's messages and react accordingly. Such detection has already been successful in text messages [49]. However, one might even detect a "vague sort of dissatisfaction" at Level 1 through meta-information like hesitation or body language.

In conversational search, the provider can gain a deeper understanding of what the seeker does and does not say as the conversation progresses and through explicit requests. The information need is an anomalous state of knowledge [15], where the seeker, "faced with a problem, recognizes that their state of knowledge is inadequate for resolving that problem, and decides that obtaining information about the problem area and its circumstances is an appropriate means towards its resolution." The provider should engage in user revealment [42], i.e., to help the seeker express (or discover) the actual information need and possibly also long-term preferences. In terms of

Table 3. Model of the information search process according to the uncertainty principle as per Kuhlthau [31].

Tasks	Initiation	Selection	Exploration	Formulation	Collection	Presentation
Feelings *(affective)*	Uncertainty	Optimism	Confusion, frustration, doubt	Clarity	Sense of direction, confidence	Satisfaction or disappointment
Thoughts *(cognitive)*	Vague ——————> Focused ——— Increased interest ———>					
Actions *(physical)*	Seeking relevant information, exploring ——————> Seeking pertinent information, documenting					

meta-information, one may model such circumstances through what is missing. To understand what is missing, however, the provider has to use their domain knowledge.

[Example 13] In Dialog 3, the provider uses the domain knowledge "web server *requires* server hardware" to identify what the seeker needs.

Example 13. Identification of an anomalous state of knowledge in Turns 1 and 2 of Dialog 3.

Information		**Formal model**	
σ	"to play my music collection"'	(1)	σ *sets* topic
μ_1	"build myself a small web server" Target: Information need Concept: User revealment, facets [42] Intent: Describe missing part	(2)	μ_1 *requests* facet
μ_2	"use a Raspberry Pi for [the server]" Target: Domain knowledge Concept: Facets [42] Intent: Offer specification	(3) (4)	μ_2 *offers* facet *for* (2) (3) *closes* (2)

3.5 Seeker-Related Meta-Information

Although the "user context" has been the focus of much empirical work in information retrieval (cf. Section 2.6), as far as we know, the principle of uncertainty [31] is the only theoretical work that is concerned with the seeker beyond the information need. Table 3 illustrates the process of information search according to this principle. The initial state of uncertainty in the seeker corresponds to the anomalous state of knowledge [15] discussed earlier. But it is considered under an affective point of view in the principle of uncertainty. The principle considers the process of seeking and finding information from an affective (feelings of the seeker), cognitive (thoughts of the seeker), and physical (actions of the seeker) perspective. During the process, the principle assumes that the seeker progresses from the initiation of the search to the result ("presentation") through four intermediate states. Concerning meta-information, the principle thus provides meta-information on top of the user context described already for non-conversational search (Section 2.6). For example, if some meta-information sets the *feelings* to "clarity," and both *thoughts* and *actions* have a fitting value, then the principle of uncertainty provides that the *task* is "formulation."

3.6 Provider-Related Meta-Information

In conversational search, it becomes necessary to model the provider's actions just like the seeker's, as one can only interpret the seeker's actions in the whole conversation context. Especially in cases where the seeker asks about the reasons for the provider's actions, such meta-information becomes necessary to allow the provider to explain their actions. At the moment, though, not much theoretical work on the provider exists. Indeed, research on retrieval models comes closest to a theory of the provider for traditional information retrieval, but most retrieval models are empirical in nature [54]. Still, one can treat their scores as meta-information on the results. The only theoretical framework that we are aware of is the conversation action space, which categorizes both the provider's result answers and the seeker's response options [42]. The provider can answer with either regular items or so-called partial items, the latter of which are either described by a field (e.g., a product's price), a field with value (e.g., one specific price or price range), or a cluster (e.g., "low-budget products"). Response options that the provider can allow are a rating, a preference, a lack of preference, a critique of the (partial) items in which the seeker says why the items are not suited, or unstructured text. Like retrieval models, the conversation action space framework uses a score for all (partial) items that is updated on each turn.

Moreover, some providers may adopt a personality. For such providers, explicit persona traits [73] (e.g., "I enjoy trash movies.") can be meta-information to messages—both of the seeker and the provider—to model the influence of the seeker's personality on the conversation. A different approach is to measure and control the manifestation of a provider's personality in single messages, for example, their expressed warmth and competence [33].

3.7 Discussion: Intents of Meta-Information

As Sections 2 and 3 illustrate, one can see most, if not all, analyses related to information retrieval and conversations as meta-information that target the different entities of information-seeking conversations. For humans, this meta-information layer is natural to conversations, and the grand vision of conversational search thus requires systems to understand and reason about meta-information. As a first step towards a formalization of meta-information, this work presents several examples of meta-information uses, each alongside an RDF-like formal model.

As our examples show, one can use meta-information with various intents in mind. Table 4 compiles an overview of the intents discussed in the examples throughout the paper. As the table highlights, one can express the same intent in various concepts, and meta-information of the same concept can express various intents. Naturally, more intents can be envisioned, and the intents can be analyzed using other concepts. For example, the conversation state concepts (context spaces, conversational roles, QRFA) can express several of the listed intents. But even in the few examples given, a large variety of ways to constrain search results in a conversation becomes apparent. Moreover, related to this observation, meta-information can be used to elaborate and justify constraints and evaluate the results retrieved. Other intents are not directly related to the search results but provide signals that can be combined with additional information to influence the results. For instance, "locate seeker" is irrelevant to the retrieval until a ranking by distance is desired or expected, at which point it becomes of critical importance. Though our discussion of meta-information has been mostly theoretical, note that many of the examples—including all from the next section—are taken from datasets that simulate interactions with conversational search systems as accurately as possible. Such simulations are necessary as hardly any system exists which implements complex multi-turn conversational search capabilities.

Table 4. Intended effects of meta-information, the concepts employed, alongside the examples given.

Intent	**Concepts (example)**
Answer request	Elements (Example 19)
Clarify answer	Examples (Example 19)
Clarify request	Examples (Example 19)
Constrain duration	User context (Example 1), word choice (Example 2)
Constrain results	Context spaces (Example 8), context terms (Example 17), conversational implicature (Example 9), conversational roles (Example 7), document features (Example 18), facets (Example 3, 6, 10, 17), fields (Example 10), memory (Example 6), QRFA state (Example 7), query strategies for extra-topical preferences (Example 14), seeker knowledge (Example 12), user criteria (Example 20)
Describe missing part	Elements (Example 18), facets (Example 13), set retrieval (Example 11), user revealment (Example 13)
Describe result	Fields (Example 15, 16)
Elaborate on constraint	Facets (Example 3)
Emphasize topic change	Word choice (Example 2)
Evaluate result	Anomalous state of knowledge (Example 20), fields (Example 15), user criteria (Example 20), metadata (Example 17)
Improve answer	User criteria (Example 12)
Justify constraint	Metadata (Example 10)
Locate seeker	User context (Example 5)
Offer result	Conversational roles (Example 20)
Offer specification	Context spaces (Example 8), conversational roles (Example 7), facets (Example 13), QRFA state (Example 7), user revealment (Example 17, 18)
Provide answer	Context spaces (Example 8), conversational roles (Example 7), document information elements (Example 11), QRFA state (Example 7)
Rank results	Extra-topical dimensions (Example 5)
Request detail	Document features (Example 18), facets (Example 15)
Request information	Context spaces (Example 8), conversational roles (Example 7), QRFA state (Example 7)
Request specification	Conversational implicature (Example 9), facets (Example 11)
Select result	Facets (Example 15, 16)
Show affective state	Affective/physiological/behavioral features (Example 2)

4 DATASETS FOR CONVERSATIONAL SEARCH WITH META-INFORMATION

To investigate possible uses of meta-information in real-life conversational search, we analyze both the Spoken Conversational Search dataset (SCS; [62–64]) and the Microsoft Information-Seeking Conversations dataset (MISC; [59]). From an original list of 27 conversational search datasets, only these two datasets fulfill our minimum requirements of closeness to human conversations and inclusion of meta-information: (1) more than two turns per dialog, (2) human-human or human-wizard information-seeking dialogs with minimal restrictions on each participant's actions and utterances, and (3) suitably complex search tasks to necessitate the inspection of more than one search result. After discussing the two datasets (Section 4.1), we provide insights from both a quantitative (Section 4.2) and qualitative analysis (Section 4.3) of the datasets concerning meta-information. Based on this analysis, Section 4.4 derives additional dataset properties that would allow for more in-depth studies of the use of meta-information in conversational search.

4.1 Dataset Overview

In both datasets, one human seeker engages in a voice-only conversation with one human provider, who accesses a search engine on the seeker's behalf to fulfill an information search task. The tasks in both datasets are designed very similarly. The seekers received the task at the start of a conversation, which was limited to ten minutes. The tasks encompass both directed and exploratory search tasks with detailed backstories. But while for SCS, the seekers were instructed to stop whenever they were satisfied [65], the participants in MISC had to give answers to the questions in their task and were interviewed about their perceived success. Next to pre-task and post-task questionnaires, MISC also contains audio and video recordings of the participants as well as data on prosody and facial expressions. For MISC, several seekers read their task to the provider, which led to an implausible situation of the provider being fully aware of the seeker's task from the start. On the contrary, to read the task to the provider was explicitly disallowed for SCS. The SCS dataset contains 39 conversations from 13 pairs of participants, and the MISC dataset contains 110 conversations from 22 pairs of participants.

As per manual inspection, the conversations in both datasets seem adequate for analyzing the use of meta-information in conversational search. Even though seekers formulated their queries in some cases as if they were typing them into a traditional search engine, the conversations that develop around these queries do not seem forced but natural. Specifically, the datasets contain query formulations, result set presentations, result selections, and the presentation of relevant information within the selected results, as one would expect in conversational search.

As a critical shortcoming of MISC, its transcription contains many erroneous words due to the employed automatic speech recognition, which prevented us from performing a quantitative analysis of the dataset. Section 4.3 still reports on a qualitative analysis of some examples that we manually transcribed or took from other work that uses MISC [58]. If the transcription issues can be resolved in future work, MISC would present an excellent dataset to study the use of meta-information, given the questionnaires and recordings provided along with it.

4.2 Quantitative Analysis of the SCS Dataset

For a rough overview of the prevalence of meta-information in conversational search, we manually annotated each turn of the SCS dataset for meta-information that targets documents, results, or the collection. We restricted ourselves to these types of meta-information as these are well-defined in the literature (cf. Section 2) and as their use is probably not affected by the setup in which the dataset was collected. Each of the 1044 dialog turns in the dataset was annotated with the types of meta-information—based on the ones discussed in Section 2—it mentions. A turn in the dataset

Table 5. Number of turns with explicit mentions of the respective meta-information in the SCS dataset by speaker role and task complexity. Percentages are relative to the overall number of turns in the same row.

Role	Complexity	Turns	Turns with meta-information								
			Author	Count	Date	Format	Genre	Origin	Site	Source	Any
Seeker	Remember	126	0 0.0%	0 0.0%	0 0.0%	3 2.4%	0 0.0%	0 0.0%	1 0.8%	0 0.0%	4 3.2%
	Understand	178	0 0.0%	0 0.0%	2 1.1%	6 3.4%	4 2.2%	1 0.6%	9 5.1%	2 1.1%	19 10.7%
	Analyze	224	4 1.8%	0 0.0%	9 4.0%	7 3.1%	7 3.1%	0 0.0%	9 4.0%	3 1.3%	32 14.3%
	Total	528	4 0.8%	0 0.0%	11 2.1%	16 3.0%	11 2.1%	1 0.2%	19 3.6%	5 0.9%	55 10.4%
Provider	Remember	122	0 0.0%	3 2.5%	0 0.0%	7 5.7%	3 2.5%	3 2.5%	10 8.2%	8 6.6%	20 16.4%
	Understand	174	0 0.0%	3 1.7%	1 0.6%	15 8.6%	14 8.0%	3 1.7%	20 11.5%	10 5.7%	43 24.7%
	Analyze	220	3 1.4%	2 0.9%	26 11.8%	13 5.9%	19 8.6%	1 0.5%	25 11.4%	24 10.9%	61 27.7%
	Total	516	3 0.6%	8 1.6%	27 5.2%	35 6.8%	36 7.0%	7 1.4%	55 10.7%	42 8.1%	124 24.0%

corresponds to a series of uninterrupted utterances by one participant. Specifically, we detected the following types of meta-information: the document's *author*, the *count* of results on the results page, the document's publication *date*, information *format* (e.g., list, tabular comparison, image, video), *genre* (e.g., news article, forum thread, blog post), country or place of *origin*, *site*, and the information *source*. To ensure annotation quality, we reviewed all annotations independently and resolved two dozen potential errors in a discussion. We also reviewed a sample of 100 turns where we found no mention of meta-information in the first pass, and only two instances contained previously undiscovered meta-information. In total, we found mentions of meta-information in 32 (82%) of the dialogs. Our annotations are available online as Webis SCSmeta 2021.[9] The aforementioned errors in the transcripts of MISC prevented a similar analysis of this dataset.

Table 5 provides statistics about the meta-information used by the seeker and the provider for the entire dataset, as well as dependent on task complexity. For the whole dataset, we observe a higher amount of meta-information communication by providers. A likely reason for this is their interaction with the search engine, making the available meta-information directly apparent to them. On the other hand, seekers have to rely, to some degree, on providers revealing which meta-information is available, or resort to making assumptions in this regard.

For tasks of higher complexity, both seekers and providers mention more meta-information. As the measure of complexity, we use the dataset's task categorization that is based on the cognitive dimensions the task requires as per Anderson et al. [4]: remember, understand, and analyze. Remember tasks are the least complex and require seekers to identify and retrieve relevant knowledge but require no analysis. Understand tasks require seekers to build connections between knowledge items. Analyze tasks are the most complex and require seekers to "break material into its constituent parts and determine how the parts relate to one another and to an overall structure or purpose." As the table shows, the increasing complexity of tasks is reflected in an increase in both the total number of dialog turns and the percentage of turns that mention meta-information. However, different types of meta-information are affected differently. Especially the result date seems to be important for analyze tasks, mentioned in every ninth turn of the provider. Overall, 24% of turns mention meta-information, which shows its widespread and intuitive usage by seekers.

[9] The dataset is available at https://webis.de/data.html#webis-scsmeta-21 and https://doi.org/10.5281/zenodo.4108195.

4.3 Qualitative Analysis of the SCS and MISC datasets

This section presents a detailed analysis of meta-information in five excerpts from the SCS and MISC datasets to understand further how humans use meta-information in information-seeking conversations. We chose these excerpts to shed light on four crucial cases of meta-information usage that are also challenging for conversational search systems: result constraints, document description, user revealment, and the handling of unclear situations.

4.3.1 Meta-information to constrain the results. So far, this paper mainly discussed the use of meta-information to constrain results directly, and the SCS and MISC datasets indeed contain several such cases. Dialog 4 shows an example for constraining results that is different from the examples discussed in that the seeker suggests changing the entire collection.

Dialog 4. SCS dialog excerpt on "health benefits of marine vegetation as food or drugs."

Turn	Role	Message
7	Seeker	Do you have, uhm… No, I want you to put them all in the same… in the same expression. And do you have access to Google Scholar?
8	Provider	Google Scholar… Yes.
9	Seeker	So search on Google Scholar for "health" and "algae" and "seaweed" and "kelp"

[Example 14] Dialog 4 illustrates how seekers ask for searching specific collections. In this case, the seeker identifies the collection by a service they know instead of, e.g., asking for a search within scientific papers. Their knowledge of the service makes it easier for the seeker to describe their need and for the provider to fulfill the request. Had the seeker not known the alternative engine, would they have asked? We believe this to be less likely, since web search engines only marginally advertise switching to specialized search engines, especially when this would lead the user to a competitor's product, and since such options are invisible in a conversation. It is thus necessary for the provider to reveal their capability to do such specialized searches.

Example 14. Request to use specific collection in Dialog 4.

Information		Formal model	
σ	"Search […] for 'health' and 'algae' and 'seaweed' and 'kelp'"	(1)	σ *sets* topic
μ_1	"Search on Google Scholar" Target: Query Concept: Query strategies for extra-topical preferences [7] Intent: Constrain results	(2)	μ_1 *sets* collection

4.3.2 Meta-information for document description. The dialogues of the dataset often show the use of meta-information to describe results (cf. Table 5). Such use seems especially useful for voice-based search where the seeker can not skim results. But all search systems can benefit from an abstractive description of the results, as this allows the seekers to stay engaged with the conversation for longer instead of forcing them to go back-and-forth between the conversation and the result list.

Dialog 5. SCS dialog excerpt on "airport security" with the following backstory for the seeker: "Every time you go through the security screening at an airport, you wonder whether it is making any difference. Find out how effective the many new measures (beyond just standard screening) at airports actually are, both for scrutinizing of passengers and their checked and carry-on baggage."

Turn	Role	Message
1	Seeker	Can you type in, uhm, "effective"... "effectiveness of new security measures at airports"?
2	Provider	Australia, or is it just airports?
3	Seeker	Put, uhm, "international"... "international airports".
4	Provider	OK, the first [result] that comes up is "Airport Security Measures Aren't Good Enough, Here's a Fix", uhm, it doesn't say who it's by... looks like... seems to be an article... 2014. So it is not that recent.
5	Seeker	Does it... is it from a newspaper, or is it from a...
6	Provider	It is theconversation.com, so, uhm, so it may be, or maybe a blog.
7	Seeker	OK.
8	Provider	"TSA guidelines for passengers on new security...", that's East Texas Airpark; there's "Airport Security", Wikipedia; there's "The Debate over Airport Security", uhm, that seems to be from an organisation called the CFR, uhm...
9	Seeker	Can you just look at the CFR website?
10	Provider	This is a 2010, so...
11	Seeker	OK, try and... yeah, OK... forget that one.
⋮		
16	Provider	The next one is the impact of nine eleven, so that's from 2007. So a lot of these tend to be quite old, 2013, 2010. Are you wanting anything newer?
17	Seeker	Ah, uhm, "new security measures after Brussels bombing"
18	Provider	OK. Effectiveness of new security measures at international airports... just add it?
19	Seeker	Yes, yes.
20	Provider	Okay, so in "Brussels Airport Bombing Brings New Security Measures in the US", that is 2016, so that's much better. April 9, uhm, we've got March 2016, "Brussels Attacks: How Airport Bombings will Change..."—it is probably the "Air"-what? "Security"? That's news.com, Washington Post was the first, uhm, article...
21	Seeker	Can you just look at the news.com?
22	Provider	Yep, that's the second one, that's "The Brussels Attacks: How Airport Bombings will Change Air-[inaudible segment]" [long pause] OK. How attacks at Zaventem... Zaventem Airport in Brussels will change air travel, uhm. It is just a picture of the... I think he is the mastermind, uhm [...]

[Example 15] Turns 4 to 7 of Dialog 5 illustrate how a provider can extensively use meta-information to describe just one result: its title, the lack of author information, the type of the result ("seems to be an article") and its publication year, along with the interpretation of the article being "not that recent." The seeker then requests more meta-information ("is it from a newspaper") and seemingly dismisses the result based on that, without information on the article's actual content beyond the title. Note that, though the provider might have taken some information from the documents rather than the results page, they present all information like a result to the seeker.

Example 15. Document-targeted meta-information in Turns 4 to 7 of Dialog 5.

Information		**Formal model**	
σ	"The first [result] that comes up"	(1)	σ *sets* document d_1
μ_1	"Airport Security Measures Aren't Good Enough, Here's a Fix" Target: Result Concept: Fields [42] Intent: Describe result	(2)	μ_1 *sets* d_1.title
μ_2	"It doesn't say who it's by" Target, concept, intent: same as μ_1	(3)	μ_2 *sets* d_1.author *to* unknown
μ_3	"Seems to be an article" Target, concept, intent: same as μ_1	(4)	μ_3 *sets* d_1.genre
μ_4	"2014" Target, concept, intent: same as μ_1	(5)	μ_4 *sets* d_1.date
μ_5	"[2014] is not that recent" Target, concept: same as μ_1 Intent: Evaluate result	(6)	μ_5 *evaluates* d_1 *by* date
μ_6	"Is it from a newspaper[?]" Target: Document Concept: Facets [42] Intent: Request detail	(7)	μ_6 *requests* d_1.site
μ_7	"It is theconversation.com, so [...] maybe a blog" Target, concept, intent: same as μ_1	(8) (9)	μ_7 *sets* d_1.site (8) *closes* (7)

4.3.3 Meta-information for user revealment. One central promise of the conversational search paradigm is that it allows to assist seekers even if they can not formalize their problem, yet [42], and, as the following examples show, meta-information can be an essential tool both for picking up clues from the seeker and for assisting them.

[Example 16] Turns 8 and 9 of Dialog 5 illustrate how implicit clues can reveal parts of the seeker's information need to the provider. In Turn 8, the provider adapts their result presentation to reflect the site's importance in the seeker's decision to dismiss the previous result: they name just the title and the site for the next results. Furthermore, note how the seeker then uses the site to pick a result in Turn 9, which can be another hint at the seeker's perceived importance of the site. We want to highlight that the seeker did not need to formulate this criterion and may have had just vague thoughts in this direction (cf. the principle of uncertainty in Section 3.5).

Example 16. Result selection by meta-information in Turns 8 and 9 of Dialog 5.

Information		**Formal model**	
σ	"There's 'The Debate over Airport Security'"	(1)	σ *sets* document d_3
μ_1	"Seems to be from an organisation called the CFR" Target: Result Concept: Fields [42] Intent: Describe result	(2)	μ_1 *sets* d_3.site
μ_2	"Can you just look at the CFR website?" Target: Document Concept: Facets [42] Intent: Select result	(3)	μ_2 *selects* d_3 *by* μ_1

[Example 17] Turns 16 to 19 of Dialog 5 illustrate user revealment through clarifications. The provider, perhaps reminded of the dismissal of a result based on the publication date in Turn 11, presumably recognizes that most or all results would be dismissed on the same grounds and offers to extend the query with a corresponding facet. But the way the seeker sets the facet in Turn 17 is ambiguous, so the provider has to confirm in Turn 18 that the facet should be added to the query and not replace previous facets. Such clarifications show the intricacies of natural language search and thereby the need for user revealment. On the other hand, the provider does not clarify—maybe because they do not even notice—the ambiguity of whether the seeker meant "after" to imply a temporal or causal relationship. But the conversation carries on successfully even though this ambiguity remains. Thus, not all ambiguities have to be resolved, and it will be a challenge for conversational search systems to decide for or against a resolution.

Example 17. User revealment in Dialog 5 Turns 16 and 17.

Information		Formal model	
σ	"The next one is [...]"	(1)	σ *sets* result
μ_1	"So a lot of these tend to be quite old, 2013, 2010" Target: Collection Concept: Metadata [32] Intent: Evaluate result set	(2)	μ_1 *evaluates* results
μ_2	"Are you wanting anything newer?" Target: Conversation Concept: User revealment [42] Intent: Offer specification	(3) (4)	μ_2 *offers* facet (3) *closes* (2)
μ_3	"New security measures after Brussels bombing" Target: Conversation Concept: Facets [42] Intent: Constrain results	(5) (6)	μ_2 *sets* topic (5) *closes* (3)

[Example 18] Dialog 6 shows an example where a question, initially directed at the current document, becomes part of the query. The seeker asks specifically for numbers in Turn 53. The provider notices in Turn 54 that there are no other numbers in the document, so they ask whether to extend the search for statistics towards other documents, which the seeker acknowledges.

Dialog 6. SCS dialog excerpt on "airport security."

Turn	Role	Message
52	Provider	About infrastructure. And this talks about general domestic screening procedures and general measures. So, the Australian passengers carry-on baggage, uhm, is very expensive, requiring screening for all flights. Could lead to some communities losing air services. And it talks about highest security risk, uhm…
53	Seeker	Does it give you any numbers saying?
54	Provider	Yeah, so 96% of Australian domestic passengers in Australia depart from screened airports, uhm… It doesn't give you any other statistics or numbers. It then talks about cockpit doors, training of police, uhm, reconciliation of passengers with their bags. I can go down and see, if… Is it statistics that you are looking for?
55	Seeker	Yeah, some measure whether it's maybe effective or not. Can you just, then… If there is no statistics there, like, that are really visible…
56	Provider	No, no there isn't.
57	Seeker	Can you just, uhm, type in "how many people get caught at airport security checks"?

Example 18. Requests for statistics in Dialog 6.

Information		Formal model	
σ_1	"It talks about highest security risk"	(1)	σ_1 *sets* content
μ_1	"Does it give you any numbers saying?" Target: Document Concept: Document features [61] Intent: Request detail	(2)	μ_1 *requests* statistics *for* (1)
σ_2	"96% [...] depart from screened airports"	(3)	σ_2 *closes* (2)
μ_2	"It doesn't give you any other statistics or numbers" Target: Document Concept: Elements [71] Intent: Describe missing part	(4)	μ_2 *opposes* (3)
μ_3	"Is it statistics that you are looking for?" Target: Conversation Concept: User revealment [42] Intent: Offer specification	(5)	μ_3 *offers* document feature
μ_4	"Yeah, some measure whether it's maybe effective or not." Target: Conversation Concept: Document features [61] Intent: Constrain results	(6) (7)	μ_4 *sets* document feature (6) *closes* (5)

4.3.4 Meta-information to handle unclear situations. Situations that require user revealment are naturally less certain and less clear—for both the seeker and the provider. We found two methods in the dialogs of the dataset that humans employ to handle such situations: examples and explanations.

[Example 19] Dialog 7 illustrates how complex situations can be explained through meta-information and examples. The seeker needs to find several results for one topic: selfless or heroic acts. They stress the importance of having several results in Turn X+3 by giving examples of what they think are plausible result counts. However, in Turn X+4, the provider responds that there are rather too many results, indicating that one retrieved document contains even 22 instances of such acts on its own. As expected for conversational search (cf. Section 3), the human provider does not see the document as a single result here but sees every item within the document as a result of its own. To clarify this situation, the provider adapts to the seeker in providing examples.

Dialog 7. MISC dialog excerpt on "selfless or heroic acts" as per our own transcription.

Turn	Role	Message
X+1	Seeker	Can you tell me, like, on an, uhh, like, how many actually come up? We can use one but, like...
X+2	Provider	Well...
X+3	Seeker	Is it ten? Is there fifteen? Is there one? Is there three? You know? Just like, uhm...
X+4	Provider	Uhh, I... it's pages, so uhm... you know, it's kind of random acts of kindness [unintelligible]. Here's one that comes up: "Ten Heartwarming Acts of Kindness You Didn't Hear About", uhm... "22 Acts of Kindness that will Restore..."—Let's see what this one is—"that will Restore your Faith in Humanity", so that one might, uh, uhm... you know, I'm sure there's gonna be a lot of different articles because this is kind of a big thing, these days, but, like, this one was, uh, here's one that says that, um, a pre-paid vending machine treat, um, you know, that's the pay-it-forward act that a lot of people do.

Example 19. Usage of examples for clarification in Dialog 7.

Information		**Formal model**	
σ	"How many actually come up?"	(1)	σ *requests* result count
μ_1	"Is it ten? Is there fifteen? [. . .]" Target: Conversation Concept: Examples Intent: Clarify request	(2)	μ_1 *elaborates* (1)
μ_2	"It's pages" Target: Conversation Concept: Elements [71] Intent: Answer request	(3)	μ_2 *closes* (1)
μ_3	"Here's one that comes up: [. . .]" Target: Conversation Concept: Examples Intent: Clarify answer	(4)	μ_3 *elaborates* (3)

[Example 20] Dialog 8 illustrates how a provider can explain their actions to the seeker. Even though the seeker did not ask for reputable sources, the provider seems to think that the topic demands such and tells the seeker that this is what they want to present. The seeker does not interrupt the provider, which the provider probably takes as silent agreement, as they continue to put a high emphasis on this criterion. As the provider needs some time to find a fitting result, they decide to provide the seeker with the best they found so far (from "American Family Physician"), but with the clear statement that it does not fulfill the criterion, so they continue to search. Again, the provider gives the seeker the choice to interrupt them or to agree silently. The provider then uses explanations for the information items they found within a document, just like they did for the documents they found: stating what they found, but also their reservations. Though the provider cannot make sense of what they found, a seeker with the relevant domain knowledge—what it means to relax the blood vessels—might be. Explanations thus seem to be a valuable tool for collaboration in complex search tasks.

Dialog 8. MISC dialog excerpt on "migraine treatments." Transcript taken from [58].

Turn	**Role**	**Message**
X+1	Seeker	I need to research beta-blockers and calcium channel blockers. . . Um, I guess, as to their applicability to migraines. . . and their effectiveness to migraines. And then, after that, explore other options, if I don't want to take medicines. I guess, I'd just look for beta-blockers.
X+2	Provider	(LONG PAUSE) Yeah. . . I just got beta-blockers, migraine prevention, here. . . I'm trying to find a vaguely reputable site to go with. . . (LONG PAUSE) I found something called "American Family Physician" that I have never heard of. I want to go back to "WebMD"—that can kinda be sketchy but should give some sources. . . (LONG PAUSE) In general, it says "beta-blockers work to relax the blood vessels" and it is not clear how they work to prevent migraines. . . It says, "beta-blockers have been shown to prevent migraines".

Example 20. Use of explanations in Dialog 8 Turn X+2.

Information			Formal model
σ_1	"I just got beta-blockers, migraine prevention"	(1)	σ_1 *confirms* topic
μ_1	"I'm trying to find a vaguely reputable site to go with" Target: Conversation Concept: User criteria [70] Intent: Constrain results	(2)	μ_1 *constrains* reputation
μ_2	"Something called 'American Family Physician'" Target: Conversation Concept: Conversational roles (COR) [53] Intent: Offer result	(3) (4)	μ_2 *sets* document d_1 μ_2 *sets* dialog act *to* offer
μ_3	"That I have never heard of" Target, concept: same as μ_1 Intent: Evaluate result	(5)	μ_3 *evaluates* d_1 *by* (provider) familiarity
μ_4	"I want to go back to 'WebMD'" Target, concept, intent: same as μ_2	(6) (7)	μ_4 *sets* document d_2 μ_4 *sets* dialog act *to* withdraw offer, offer
μ_5	"[WebMD] can kinda be sketchy but should give some sources" Target, concept, intent: same as μ_3	(8)	μ_5 *evaluates* d_2 *by* (provider) experience
σ_2	"Beta-blockers work to relax the blood vessels"	(9)	σ_2 *closes* topic
μ_6	"It is not clear how they work to prevent migraines" Target: Information need Concept: Anomalous state of knowledge [15] Intent: Evaluate result	(10) (11)	μ_6 *evaluates* (9) *by* logical relation μ_6 *evaluates* (12) *by* evidence
σ_3	"Beta-blockers have been shown to prevent migraines"	(12)	σ_3 *closes* topic

4.4 Requirements for Future Meta-Information Datasets

Altogether, our dataset analyses show that existing corpora are still largely insufficient to study the use of meta-information in conversational search extensively. The fact that only one dataset of relatively limited size met our minimum requirements somewhat calls into question our findings' reliability. The presented findings and suggested insights thus need to be verified in future studies on larger and more complete datasets. We acknowledge that the creation of such datasets is a challenge in its own right. However, it seems necessary for further investigations in this field.

As meta-information can emerge from various sources at various stages in the search process, it is crucial to collect the data in a setting that contains as many of such natural sources of meta-information as possible. The common practice of providing seekers with pre-made search tasks is somewhat at odds with studying the seeker's use of meta-information to describe their information need to the provider, as the "need" is artificial and has to be described to the seeker in the first place. In an attempt to alleviate this problem, Arguello et al. [7] force the participants of their study to internalize the task (by not showing the task description and the search interface at the same time) and to spent cognitive effort on interpreting it (by telling the seekers about one aspect of inappropriate results, thereby demanding them to invert the aspect for their query). Though promising, without a point of comparison with natural behavior, it is unclear to which degree the attempt succeeded in alleviating the problem. However, albeit less helpful overall, it may be more feasible to create more diverse datasets that give insight into a few specific aspects of the conversation only. For example, to allow for better comparisons between seekers and tasks, a study could employ human providers that received specific training and instructions. Similarly, a long-term study of conversational search that investigates the collection and tracking of a seeker's preferences, opinions, and interests, is still missing. Moreover, human conversations employ several channels, and the collection of, for example, the participant's tone of voice, facial expression, or eye movements requires a complicated setup, significant preprocessing effort, and consent from participants. To this end, it would be helpful to improve the transcripts of the MISC dataset, which currently hinders the exploitation of the plethora of raw data that comes with it (cf. Section 4.1).

For holistic analyses, the participants' observations usually need to be complemented by their explicit participant feedback, but existing questionnaires do not usually focus on meta-information. The development of questions on the use and perceived utility of meta-information is challenging. Standardization of questionnaires can alleviate the load on researchers, having them avoid pitfalls, and lead to increased comparability of studies. This paper's categorizations of meta-information by target, concept, and intent may inspire in this regard. In any case, the categorizations illustrate the difficulty of collecting statements on some phenomenon that is as diverse as meta-information.

5 APPLICATIONS OF META-INFORMATION TO CONVERSATIONAL SEARCH

The theoretical analysis, the observations from existing datasets, and the hypothetical examples presented so far illustrate that the proper use and understanding of meta-information will be critical for conversational search systems in order to meet the high expectations for this paradigm shift from the research community, industry, and end users. Applications of meta-information in this context are manifold. This section illustrates several applications, including ones we already alluded to, and ones we have not mentioned yet, to provide an overview. This overview is not complete, though; there is still much room for heretofore unconsidered applications. Section 5.1 looks at research into visually impaired seekers' information-seeking behavior, which has recently produced a working prototype of a search interface that incorporates meta-information. This analysis highlights lessons learned from the design of natural-language search interfaces. Section 5.2 then looks into other challenges of incorporating meta-information into conversational search systems, touching upon newly emerging requirements and some of the approaches taken in this regard. Finally, Section 5.3 discusses further uses of meta-information in conversational search by envisioning how various instances of meta-information can benefit the provider's conversation management.

5.1 Case Study: Information-Seeking for Visually Impaired Users

Conversational search is often associated with voice search as both are perceived to lead to more natural interactions with a search engine. Intuitively, it thus makes sense to investigate the state of the art for visually impaired seekers, who have much experience with voice search from their daily lives, to draw conclusions for conversational search in general. Typically, these seekers navigate the Internet with the aid of screen readers. A recent study by Upadhyay [66] examines and contrasts the web search interaction behavior of sighted and non-sighted seekers. Upadhyay states that the sequential nature of non-visual web access using screen readers "may result in increased costs of navigation and cognitive overload from excessive auditory information," which causes screen reader users to adapt their browsing strategies. Visually impaired seekers cannot use so-called proximal cues, such as links, layout, color, and keywords that form an information scent [36, 41] for them to follow. These cues can be seen as pieces of meta-information akin to the various meta-information concepts related to documents and results (see Sections 2.1 and 2.2). But not all cues are equally relevant [55]. The study by Upadhyay [66] therefore identifies which cues make up an information scent for non-sighted seekers. After the seeker sends their voice query to a regular search engine, they investigate the search engine results page. The seeker listens to the result headings, snippets, and sometimes URLs. The seeker might then reformulate the query based on keywords from the headings or snippets using "previously acquired knowledge of the content structure to explore more scent." [66] Such reformulations show the importance of utilizing meta-information about the set of retrieved results to minimize the risk of examining a document extensively, only to find that it is irrelevant: it takes much more time to assess a whole document's relevance via audio than via visual inspection. To cope with this problem, seekers first listen to the headings to "pick up scent," [66] exploring more details for a result on the results page only if it seems relevant. As a relevance judgment from the headings alone is not trivial, non-sighted seekers revisit results pages more often than sighted seekers. The authors thus suggest providing meta-information along with the headings to give some scent.

Vtyurina et al. [68] attempt to tackle these issues for visually impaired seekers combining voice assistant and screen reader technology. Their system, VERSE, aims to offer both the convenience of voice assistants and the fine-grained control and deep engagement with documents facilitated by screen readers. The system first provides concise answers like voice assistants but allows the seeker to explore the documents from which the answer has been generated. To facilitate such exploration, the system first presents meta-information on the documents. For example, the provider can state that it "found two entities, nine web pages, eight related search queries, ten videos, [and] ten Wikipedia articles." Such a statement informs the seeker of the sources and suggests how the seeker can continue their exploration. If the seeker selects one kind of source, the system again provides meta-information to allow the seeker to judge the depth and scope of a document [68], for example, by stating that it "has 16 sections and 3127 words." These capabilities were well-received by participants of a user study. The participants compared the web search interaction favorably to screen readers, with one participant stating that "this gives you much more structure." Another participant appreciated that "different forms of data were being pulled together," noting that, as opposed to the "stream of responses" obtained from Google, VERSE "gathers the relevant stuff and groups it in different ways," referring to the overview of the obtained results [68].

This case study implies that the use of meta-information in answers from the provider alone can already have an enormous positive impact on the usability of voice-based conversational search systems. One example is a more sophisticated and structured presentation of both result sets and individual results. VERSE shows that incorporating meta-information is essential in this regard. However, such improvements do not solely apply to voice interaction. Expanded capabilities of a conversational search system in result overview, navigation, and presentation can equally be utilized through a visual interface. A possible multi-modal scenario is taken into account in VERSE [68]: users can send selected results to their smartphone, where they often have more sophisticated screen reading software. However, also seekers who are not visually impaired may start a search with a simple query through voice-based interaction and then choose to move to a device with a screen if their search scope widens, if they require more detail, or if they need to view images or videos. Search success in these situations depends on seekers' ability to pick up "information scent" [41, 66] early on, which is greatly facilitated by a structured overview of results that incorporates meta-information.

5.2 Challenges for the Design and Implementation of Conversational Search Systems

At this point, several challenges remain to create a conversational search system that utilizes meta-information to its full effect. These challenges can be broadly categorized into those of meta-information acquisition, interpretation, and presentation. Each of these challenges demands in-depth investigations. For illustration, the following paragraphs present exemplary challenges for each category.

The first challenge is to acquire the meta-information reliably. As the examples presented throughout this article show, meta-information can be transmitted through a great variety of carriers. However, especially as meta-information is often not stated explicitly, it is a challenge to judge the reliability of gathered meta-information—a judgment which is likely made based on other meta-information. For example, though conversational systems and traditional web search engines already employ meta-information from past events (e.g., [37]), it is still a challenge to differentiate between short-term and long-term preferences as well as between generic and task-specific ones. Moreover, the acquisition of meta-information of an affective, physiological, or behavioral nature requires additional sensors and complex processing. However, commercial organizations are already moving in this direction, with the next generation of Echo smart speakers, at the time of writing, being equipped with cameras and developer kit support for basic movement detection.[10] Finally, not every seeker may want all their signals to be stored, processed, or interpreted. For privacy concerns, it must thus be communicated clearly to the seekers which signals are used, how they are used, and how to disable their use.

The second challenge is to interpret and utilize the acquired meta-information, which includes extending it by reasoning. Knowledge graphs seem to present themselves here as the foundational data-structure. Knowledge graphs and reasoning algorithms have been extensively studied over the past decade.[11] At the same time, publicly available knowledge graphs, like Wikidata,[12] have grown to a large size. As these knowledge graphs already contain domain knowledge, enriching them with meta-information on the respective conversation and employing the well-studied reasoning algorithms seems like a straightforward approach. However, the great variety of meta-information and their limited reliability poses a challenge to this approach that needs to be overcome to utilize meta-information to its greatest effect.

The third challenge is the presentation of meta-information within a conversation. In conversational search, both the seeker and the provider should be able to refine the search intuitively and quickly. To first present a result list that is directly described by meta-information, but may turn out to contain completely irrelevant results only, may be preferable over a mixed bag of results: in the best case, the description would allow the seeker to pinpoint what is wrong immediately, interrupt the conversation, and to refine the conversation's topic accordingly, even without looking at—or hearing—the first results. But which meta-information should be chosen for the description and presented to the seeker? First approaches exist to select the meta-information to present. They either use topic modeling and seeker feedback [30] or query log mining [72]. With an increasing multi-modality of systems,[13] it is also essential to consider which modalities are available or desirable at each moment of a conversation. Furthermore, not all modalities are equally suited for every piece of meta-information. For example, to disambiguate the topic, the system could show an image representing the topic it assumes the seeker refers to (e.g., a photo of the apple company logo). But this modality is not available on all devices. In other situations, seekers may say, "I can't see you [the provider] from here," expecting a shift towards the audio-only search. In this example, the seeker used meta-information for a temporary effect on the conversation. But for how long exactly? Another challenge in this category is thus for the system to keep the seeker consistently updated on the meta-information it employs—e.g., that it continues to assume the seeker can't see it—, so that the seeker can better comprehend and update the system.

[10] https://www.amazon.science/blog/the-science-behind-echo-show-10

[11] E.g., in the workshop series on graph structures for knowledge representation and reasoning, https://graphkr.github.io/

[12] At the time of writing, Wikidata contains more than 90,000,000 items; https://www.wikidata.org/

[13] The aforementioned Echo smart speaker can also shake itself: the first step for a gesture modality.

5.3 Use Cases for Meta-Information in Conversation Management

Though the incorporation of meta-information into conversational search systems thus comes with several challenges, it also provides several opportunities to enrich the conversations. Some of these opportunities and use cases, for which meta-information plays a key role, are described in the following paragraphs. Some use cases have already been alluded to in the analyses and examples in this paper, but are discussed in more detail here (see Table 4 for an overview of the examples).

5.3.1 Adapting to a change in information. The seeker may retract some information that the provider has already related to other information, or the domain knowledge might change due to new scientific results or events in the world. For instance, the seeker may retract a preference by stating, "I don't like red cars anymore," which may require the provider to re-evaluate the obtained results or retrieve new ones. If the provider organized the knowledge as meta-information, they can trace this retraction's influence through their knowledge and make changes as needed. Moreover the provider might want to clarify the retraction: do they still like cars?

5.3.2 Adapting to analogous knowledge. A seeker may reveal analogous knowledge to the provider, which may give the provider clues about the seeker's expertise and how they can present concepts more effectively with relation to this existing knowledge. For instance, consider a seeker familiar with the C++ programming language who wishes to learn Java. They are more likely to quickly grasp the new information when the provider presents it by comparing C++'s and Java's concepts.

5.3.3 Adapting to the level of expertise. The provider often has to select an answer based on what they assume the seeker to understand (cf. Example 12). For instance, a seeker's message may describe the content of the provider's information as "too technical," which should cause the provider to adjust the information on the seeker's expertise and adapt the retrieval and presentation strategies accordingly. Conversely, the provider may make similar judgments about the seeker's messages, for instance, acknowledging higher domain expertise when the seeker submits complex queries. Moreover, the provider may want to slowly teach the seeker more ways of interaction that they support. Thus, the provider must remember and infer which capabilities the seeker already knows, or which the seeker might have forgotten.

5.3.4 Adapting to different seekers. Conversations are not limited to two participants, and some conversational search systems might want to allow several seekers to collaborate. Such collaboration might be especially feasible for voice-based search, where the seekers can effectively share the search interface without blocking each other's view. However, such a setting requires the provider to identify the seeker through their messages—most probably using vocal features. Such identification allows the provider to associate the seeker's message with the respective user model. This problem occurs, to a limited extent, also when different seekers use the conversational search system sequentially. However, a sequential use allows for other recognition techniques like facial recognition or separate authentication credentials. For instance, consider a situation in which two persons plan to have a night out and visit a few bars in a foreign area. One of the persons is an avid beer drinker, while the other enjoys cocktails. In conversation with a search system, the beer drinker asks, "What bars around here could you recommend me?" and the system responds with a few options. Then, the cocktail drinker asks, "And which ones would I like?" prompting the system to identify them by their voice and to adapt its results and presentation to their preferences.

5.3.5 Adapting to time constraints. Seekers may enter a conversation with constraints for its duration, which influences the seeker's and the provider's entire interaction [20]. A seeker with little time is likely to prefer concise information about a topic, wishing to spend less time on an overview of the results, navigating them, or refining their questions substantially (cf. Example 1).

The provider should adapt to this, both in terms of the retrieval strategy and result presentation. With little time, seekers are likely to look for summaries about topics, which a provider could either generate or focus on for retrieval. Suppose a seeker actively wishes to engage in an extended exchange with the provider, bringing lots of time to the search process. In that case, the provider may equally adapt to this, encouraging deeper interaction with the results to deliver more information. For example, the provider may suggest additional related documents in such a situation.

5.3.6 Determining and resolving conflicts in the information. As meta-information requires interpretation, misinterpretations will happen and become apparent when different (meta-)information conflicts with each other. For example, consider a seeker responds to some request by saying "yes, fine" in an annoyed tone (cf. Example 2). The words seem to be positive, but the tone implies the opposite. If the provider incorporates further meta-information, for example, the seeker's facial expression or whether previous turns in the conversation were successful, the provider can identify that the seeker may be short of seeing the conversation as failed and that clarifications are needed.

5.3.7 Determining conversation context. Meta-information on the conversation state can help conversational search systems to keep track of the current topic. Though conversational search systems are already able to resolve straightforward cases of coreferences (e.g., "Who is Barack Obama? When did *he* leave office?"), they cannot resolve coreferences across conversation context shifts (e.g., if one search in a conversation is interrupted by the question for an explanation, and the conversation returns to the search after the explanation is given and accepted). The concept of context spaces (cf. Example 8) seems especially suited in this regard.

5.3.8 Determining mutual understanding. The seeker and the provider may misunderstand each other. In the extreme case, this may lead to a double illusion of transparency, where each party believes the other to understand the information exchanged, yet one has not. Consider an alternate version to Dialog 3, where the provider directly refers the seeker to the how-to mentioned in Turn 4 without conveying the meta-information that it requires some expertise. The seeker would have no indication to question their ability to follow the how-to, and then possibly be overwhelmed by its technical jargon. In the dialog given, however, the provider assumes a teaching role and prevents this misunderstanding by actively encouraging user revealment.

5.3.9 Explaining statements. The seeker might be confused or intrigued by the provider's messages and thus ask the provider to explain that message. The provider can then reveal the (chain of) meta-information that led to the statement. For example, consider a seeker who wants to buy a smartphone. A provider can use information about the seeker's preferences in their user model to arrive at a recommendation more quickly, requiring less facet elicitation from the seeker. Additionally, the provider may elicit the context in which the seeker's need became obvious to gain further insight into the seeker's requirements. If a seeker is then curious about why the provider recommends that particular smartphone, the provider can explain their choice by stating the meta-information they used to arrive at that particular product. For measuring the quality of such recommendations, see Tintarev and Masthoff [60] and Balog and Radlinski [9].

These use cases illustrate the possibilities of meta-information for conversational search, which go far beyond what seekers employ today in user studies. Therefore, they instead show a possible direction for conversational search systems. There is a plethora of information to consider, and the various ways in which it can become meta-information allow for a considerable expansion of both the interaction space and naturalness of interaction with conversational search systems. Therefore, it is crucial to examine these relations in more detail to allow the paradigm of conversational search to meet the high expectations placed on it.

6 CONCLUSION

Information-seeking conversations involve the exchange of meta-information to complete the seeker's state of knowledge collaboratively. The manifold of relevant meta-information ranges from explicit constraints to implicit conversational cues. While humans request, send, receive, and interpret meta-information without effort as a matter of course, computational retrieval systems will require sophisticated algorithms to keep up. Meta-information provides an information-centric view on a conversation and may complement frameworks that emphasize an action-centric view, such as CORS [53], QRFA [67], or the conversation action space [42].

The outset of our research is Floridi's [23] recognized definition of meta-information, which we further refine for the information retrieval context: (1) Meta-information is identifiable as such only by its relation to other information, (2) this relation is dynamic, and (3) the meaning of meta-information depends on both the referred information and the recipient. Our literature survey discusses 36 approaches to exploit meta-information, which have been introduced for classic information retrieval (p. 9, Table 1), as well as for conversational search and conversation theories in general (p. 15, Table 2). We categorize these approaches by their primary target, which include "classic" IR targets, such as the "result" and the "document," as well as conversational search targets, such as the "message" and "domain knowledge." Our theoretical considerations are underpinned by quantitative and qualitative analyses of two conversational search datasets, SCS [62–64] and MISC [59]. In this regard, we illustrate more than twenty different intents for meta-information usage in conversational search (p. 25, Table 4), for which we also provide a formal model. To examine the practical usage of meta-information in conversational search systems, we discuss an existing search system for the visually impaired, identify main challenges for conversational search systems in the acquisition, interpretation, and presentation of meta-information, and highlight nine use cases for meta-information to enrich information-seeking conversations. These use cases include opportunities for the provider system to adapt, determine context, and explain its actions.

"Classical" IR systems always have, and very successfully, been exploiting meta-information of various types, even from such noisy signals as dwell time on a result page. With our article, we express the vision to extend this effective exploitation to the field of conversational search, which in turn leads to the question of whether the information retrieval toolbox already contains the necessary means, and, if not, which of the methods can be adapted, and which have to be invented. Besides studying individual pieces of meta-information in the context of conversational search, specific next steps include creating annotated datasets at scale, schematizing our formal model, as well as inviting the wider IR community to participate, e.g., in a TREC-style evaluation.

ACKNOWLEDGMENTS

We sincerely thank the anonymous reviewers for their valuable comments and helpful suggestions.

This work has been partially funded by the Thüringer Ministerium für Wirtschaft, Wissenschaft und Digitale Gesellschaft under grant 5575/10-5 (MetaReal).

REFERENCES

[1] Yamen Ajjour, Henning Wachsmuth, Dora Kiesel, Patrick Riehmann, Fan Fan, Giuliano Castiglia, Rosemary Adejoh, Bernd Fröhlich, and Benno Stein. 2018. Visualization of the Topic Space of Argument Search Results in args.me. In *2018 Conference on Empirical Methods in Natural Language Processing (EMNLP 2018) - System Demonstrations*, Eduardo Blanco and Wei Lu (Eds.). Association for Computational Linguistics, 60–65. http://aclweb.org/anthology/D18-2011

[2] James F. Allen, Curry I. Guinn, and Eric Horvtz. 1999. Mixed-initiative interaction. *IEEE Intelligent Systems and their Applications* 14, 5 (1999), 14–23. https://doi.org/10.1109/5254.796083

[3] Raviteja Anantha, Svitlana Vakulenko, Zhucheng Tu, Shayne Longpre, Stephen Pulman, and Srinivas Chappidi. 2020. Open-Domain Question Answering Goes Conversational via Question Rewriting. *CoRR* abs/2010.04898 (2020). arXiv:2010.04898 https://arxiv.org/abs/2010.04898

[4] Lorin W. Anderson, David R. Krathwohl, Peter W. Airasian, Kathleen A. Cruikshank, Richard E. Mayer, Paul R. Pintrich, James Raths, and Merlin C. Wittrock. 2001. *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives.* Longman.

[5] Ioannis Arapakis, Joemon M. Jose, and Philip D. Gray. 2008. Affective feedback: An investigation into the role of emotions in the information seeking process. In *ACM SIGIR 2008 - 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Proceedings.* ACM Press, New York, New York, USA, 891. https://doi.org/10.1145/1390334.1390566

[6] Ioannis Arapakis, Yashar Moshfeghi, Hideo Joho, Reede Ren, David Hannah, and Joemon M. Jose. 2009. Enriching user profiling with affective features for the improvement of a multimodal recommender system. *CIVR 2009 - Proceedings of the ACM International Conference on Image and Video Retrieval* (2009), 216–223. https://doi.org/10.1145/1646396.1646433

[7] Jaime Arguello, Bogeum Choi, and Robert Capra. 2018. Factors influencing users' information requests: Medium, target, and extra-topical dimension. *ACM Transactions on Information Systems (TOIS)* 36, 4 (2018), 1–37. https://doi.org/10.1145/3209624

[8] Jaime Arguello, Adam Ferguson, Emery Fine, Bhaskar Mitra, Hamed Zamani, and Fernando Diaz. 2021. Tip of the Tongue Known-Item Retrieval: A Case Study in Movie Identification. In *Conference on Human Information Interaction and Retrieval (CHIIR '21)*, Falk Scholer, Paul Thomas, David Elsweiler, Hideo Joho, Noriko Kando, and Catherine Smith (Eds.). ACM, 5–14. https://doi.org/10.1145/3406522.3446021

[9] Krisztian Balog and Filip Radlinski. 2020. Measuring Recommendation Explanation Quality: The Conflicting Goals of Explanations. In *43rd International Conference on Research and Development in Information Retrieval (SIGIR 2020)*, Jimmy Huang, Yi Chang, Xueqi Cheng, Jaap Kamps, Vanessa Murdock, Ji-Rong Wen, and Yiqun Liu (Eds.). ACM, 329–338. https://doi.org/10.1145/3397271.3401032

[10] Carol L. Barry. 1998. Document representations and clues to document relevance. *Journal of the American Society for Information Science* 49, 14 (1998), 1293–1303. https://doi.org/10.1002/(sici)1097-4571(1998)49:14<1293::aid-asi7>3.3.co;2-5

[11] Carol L. Barry and Linda Schamber. 1998. Users' criteria for relevance evaluation: A cross-situational comparison. *Information Processing & Management* 34, 2-3 (mar 1998), 219–236. https://doi.org/10.1016/S0306-4573(97)00078-2

[12] Carol L. Barry. 1994. User-defined relevance criteria: An exploratory study. *Journal of the American Society for Information Science* 45, 3 (apr 1994), 149–159. https://doi.org/10.1002/(SICI)1097-4571(199404)45:3<149::AID-ASI5>3.0.CO;2-J

[13] Nicholas J. Belkin, Colleen Cool, Adelheit Stein, and Ulrich Thiel. 1995. Cases, scripts, and information-seeking strategies: On the design of interactive information retrieval systems. *Expert systems with applications* 9, 3 (1995), 379–395. https://doi.org/10.1016/0957-4174(95)00011-W

[14] Nicholas J. Belkin, Pier G. Marchetti, and Colleen Cool. 1993. BRAQUE: Design of an interface to support user interaction in information retrieval. *Information Processing and Management* 29, 3 (1993), 325–344. https://doi.org/10.1016/0306-4573(93)90059-M

[15] Nicholas J. Belkin, Robert N. Oddy, and Helen M. Brooks. 1982. Ask for Information Retrieval: Part I. Background and Theory. *Journal of Documentation* 38, 2 (1982), 61–71. https://doi.org/10.1108/eb026722

[16] Andrei Z. Broder. 2002. A taxonomy of web search. *SIGIR Forum* 36, 2 (2002), 3–10. https://doi.org/10.1145/792550.792552

[17] B. Barla Cambazoglu, Leila Tavakoli, Falk Scholer, Mark Sanderson, and W. Bruce Croft. 2021. An Intent Taxonomy for Questions Asked in Web Search. In *Proceedings of the 2021 Conference on Human Information Interaction and Retrieval (CHIIR '21)*, Falk Scholer, Paul Thomas, David Elsweiler, Hideo Joho, Noriko Kando, and Catherine Smith (Eds.). ACM, 85–94. https://doi.org/10.1145/3406522.3446027

[18] Duen Horng Chau, Brad A. Myers, and Andrew Faulring. 2008. What to do when search fails: finding information by association. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI '08)*, Mary Czerwinski, Arnold M. Lund, and Desney S. Tan (Eds.). ACM, 999–1008. https://doi.org/10.1145/1357054.1357208

[19] Colleen Cool, Nicholas Belkin, Paul Kantor, and Ophir Frieder. 1993. Characteristics of texts affecting relevance judgments. *Proceedings - National Online Meeting* 14 (1993), 77–83.

[20] Anita Crescenzi, Robert G. Capra, Bogeum Choi, and Yuan Li. 2021. Adaptation in Information Search and Decision-Making under Time Constraints. In *Conference on Human Information Interaction and Retrieval (CHIIR '21)*, Falk Scholer, Paul Thomas, David Elsweiler, Hideo Joho, Noriko Kando, and Catherine Smith (Eds.). ACM, 95–105. https://doi.org/10.1145/3406522.3446030

[21] J. Shane Culpepper, Fernando Diaz, and Mark D. Smucker. 2018. Research Frontiers in Information Retrieval: Report from the Third Strategic Workshop on Information Retrieval in Lorne (SWIRL 2018). *SIGIR Forum* 52, 1 (2018), 34–90. https://doi.org/10.1145/3274784.3274788

[22] David Elsweiler, Stefan Mandl, and Brian Kirkegaard Lunn. 2010. Understanding casual-leisure information needs: a diary study in the context of television viewing. In *Information Interaction in Context Symposium*, Nicholas J. Belkin and Diane Kelly (Eds.). ACM, 25–34. https://doi.org/10.1145/1840784.1840790

[23] Luciano Floridi. 2019. Semantic Conceptions of Information. In *The Stanford Encyclopedia of Philosophy* (winter 2019 ed.), Edward N. Zalta (Ed.). Metaphysics Research Lab, Stanford University.
[24] B. J. Fogg and Hsiang Tseng. 1999. The Elements of Computer Credibility. In *Proceeding of the Conference on Human Factors in Computing Systems (CHI 1999)*, Marian G. Williams and Mark W. Altom (Eds.). ACM, 80–87. https://doi.org/10.1145/302979.303001
[25] Alexandru-Lucian Gînsca, Adrian Popescu, and Mihai Lupu. 2015. Credibility in Information Retrieval. *Foundations and Trends in Information Retrieval* 9, 5 (2015), 355–475. https://doi.org/10.1561/1500000046
[26] Herbert Paul Grice. 1989. *Studies in the Way of Words.* Harvard University Press.
[27] Matthias Hagen, Martin Potthast, Benno Stein, and Christof Bräutigam. 2011. Query Segmentation Revisited. In *20th International Conference on World Wide Web (WWW 2011)*, Sadagopan Srinivasan, Krithi Ramamritham, Arun Kumar, M. P. Ravindra, Elisa Bertino, and Ravi Kumar (Eds.). ACM, 97–106. https://doi.org/10.1145/1963405.1963423
[28] Makoto P Kato, Takehiro Yamamoto, Hiroaki Ohshima, and Katsumi Tanaka. 2014. Cognitive search intents hidden behind queries: A user study on query formulations. In *WWW 2014 Companion - Proceedings of the 23rd International Conference on World Wide Web.* ACM Press, New York, New York, USA, 313–314. https://doi.org/10.1145/2567948.2577279
[29] Anders Kofod-Petersen and Agnar Aamodt. 2003. Case-based Situation Assessment in a Mobile Context-Aware System. In *Artificial Intelligence in Mobile Systems.* 41–49.
[30] Weize Kong and James Allan. 2014. Extending Faceted Search to the General Web. In *Proceedings of the 23rd ACM International Conference on Information and Knowledge Management.* Association for Computing Machinery, New York, NY, USA, 839–848. https://doi.org/10.1145/2661829.2661964
[31] Carol Collier Kuhlthau. 1993. A Principle of Uncertainty for Information seeking. *Journal of Documentation* 49, 4 (1993), 339–355. https://doi.org/10.1108/eb026918
[32] John A. Kunze and Thomas Baker. 2007. *The Dublin Core Metadata Element Set.* RFC 5013. RFC Editor.
[33] Irene Lopatovska, Elena Korshakova, Diedre Brown, Yiqiao Li, Jie Min, Amber Pasiak, and Kaige Zheng. 2021. User Perceptions of an Intelligent Personal Assistant's Personality: The Role of Interaction Context. In *Conference on Human Information Interaction and Retrieval (CHIIR '21)*, Falk Scholer, Paul Thomas, David Elsweiler, Hideo Joho, Noriko Kando, and Catherine Smith (Eds.). ACM, 15–25. https://doi.org/10.1145/3406522.3446018
[34] Kelly L. Maglaughlin and Diane H. Sonnenwald. 2002. User perspectives on relevance criteria: A comparison among relevant, partially relevant, and not-relevant judgments. *Journal of the American Society for Information Science and Technology* 53, 5 (2002), 327–342. https://doi.org/10.1002/asi.10049
[35] Christos Makris, Yannis Plegas, and Sofia Stamou. 2012. Web query disambiguation using PageRank. *Journal of the American Society for Information Science and Technology* 63, 8 (2012), 1581–1592.
[36] David Maxwell and Leif Azzopardi. 2018. Information Scent, Searching and Stopping - Modelling SERP Level Stopping Behaviour. In *40th European Conference on IR Research (ECIR 2018) (Lecture Notes in Computer Science, Vol. 10772)*, Gabriella Pasi, Benjamin Piwowarski, Leif Azzopardi, and Allan Hanbury (Eds.). Springer, 210–222. https://doi.org/10.1007/978-3-319-76941-7_16
[37] Seungwhan Moon, Pararth Shah, Anuj Kumar, and Rajen Subba. 2019. Memory Graph Networks for Explainable Memory-grounded Question Answering. In *Proceedings of the 23rd Conference on Computational Natural Language Learning (CoNLL).* Association for Computational Linguistics, Hong Kong, China, 728–736. https://doi.org/10.18653/v1/K19-1068
[38] Yashar Moshfeghi and Joemon M. Jose. 2013. An effective implicit relevance feedback technique using affective, physiological and behavioural features. In *Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval - SIGIR '13.* ACM Press, New York, New York, USA, 133. https://doi.org/10.1145/2484028.2484074
[39] Taemin Kim Park. 1993. The Nature of Relevance in Information Retrieval: An Empirical Study. *The Library Quarterly* 63, 3 (jul 1993), 318–351. https://doi.org/10.1086/602592
[40] Virgiliu Pavlu, Shahzad Rajput, Peter B. Golbus, and Javed A. Aslam. 2012. IR system evaluation using nugget-based test collections. In *Proceedings of the Fifth International Conference on Web Search and Web Data Mining, WSDM*, Eytan Adar, Jaime Teevan, Eugene Agichtein, and Yoelle Maarek (Eds.). ACM, 393–402. https://doi.org/10.1145/2124295.2124343
[41] Peter Pirolli and Stuart Card. 1999. Information foraging. *Psychological Review* 106, 4 (1999), 643–675. https://doi.org/10.1037/0033-295X.106.4.643
[42] Filip Radlinski and Nick Craswell. 2017. A Theoretical Framework for Conversational Search. In *Proceedings of the 2017 Conference on Human Information Interaction and Retrieval.* 117–126. https://doi.org/10.1145/3020165.3020183
[43] Rachel Reichman. 1985. *Getting Computers to Talk Like You and Me: Discourse Context, Focus, and Semantics:(an ATN Model).* MIT press.
[44] Soo Young Rieh. 2010. Credibility and cognitive authority of information. In *Encyclopedia of Library and Information Sciences* (3rd ed.). Taylor and Francis, 1337–1344. https://doi.org/10.1081/E-ELIS3-120044103

[45] J. J. Rocchio and Gerard Salton. 1965. Information search optimization and interactive retrieval techniques. In *Proceedings of the 1965 fall joint computer conference, (AFIPS 1965)*, Robert W. Rector (Ed.). ACM, 293–305. https://doi.org/10.1145/1463891.1463926
[46] Daniel M. Russell. 2012. Using context terms to focus your search. SearchReSearch Blog. http://searchresearch1.blogspot.com/2012/07/using-context-terms-to-focus-your.html
[47] Daniel M. Russell. 2019. *The Joy of Search: A Google Insider's Guide to Going Beyond the Basics.* MIT Press.
[48] Ian Ruthven. 2011. Information Retrieval in Context. In *Advanced Topics in Information Retrieval*, Massimo Melucci and Ricardo Baeza-Yates (Eds.). The Information Retrieval Series, Vol. 33. Springer, 187–207. https://doi.org/10.1007/978-3-642-20946-8_8
[49] Ian Ruthven. 2019. The language of information need: Differentiating conscious and formalized information needs. *Inf. Process. Manag.* 56, 1 (2019), 77–90. https://doi.org/10.1016/j.ipm.2018.09.005
[50] Ian Ruthven. 2019. Making Meaning: A Focus for Information Interactions Research. In *Proceedings of the Conference on Human Information Interaction and Retrieval, (CHIIR '19)*, Leif Azzopardi, Martin Halvey, Ian Ruthven, Hideo Joho, Vanessa Murdock, and Pernilla Qvarfordt (Eds.). ACM, 163–171. https://doi.org/10.1145/3295750.3298938
[51] Tefko Saracevic. 2016. *The Notion of Relevance in Information Science: Everybody knows what relevance is. But, what is it really?* Morgan & Claypool Publishers. https://doi.org/10.2200/S00723ED1V01Y201607ICR050
[52] Claude E. Shannon and Warren Weaver. 1949. *The Mathematical Theory of Communication.* University of Illinois Press, Urbana, IL.
[53] Stefan Sitter and Adelheit Stein. 1992. Modeling the illocutionary aspects of information-seeking dialogues. *Information Processing & Management* 28, 2 (1992), 165–180. https://doi.org/10.1016/0306-4573(92)90044-Z
[54] Benno Stein, Tim Gollub, and Maik Anderka. 2017. Retrieval Models. In *Encyclopedia of Social Network Analysis and Mining (ESNAM)*, Reda Alhajj and Jon G. Rokne (Eds.). Springer, Berlin Heidelberg New York, 1–7. https://doi.org/10.1007/978-1-4614-7163-9_117-1
[55] S. Shyam Sundar, Silvia Knobloch-Westerwick, and Matthias R. Hastall. 2007. News cues: Information scent and cognitive heuristics. *Journal of the American Society for Information Science and Technology* 58, 3 (2007), 366–378. https://doi.org/10.1002/asi.20511
[56] Robert S. Taylor. 1962. The process of asking questions. *American Documentation* 13, 4 (oct 1962), 391–396. https://doi.org/10.1002/asi.5090130405
[57] Jaime Teevan, Eytan Adar, Rosie Jones, and Michael A. S. Potts. 2007. Information re-retrieval: repeat queries in Yahoo's logs. In *SIGIR 2007: Proceedings of the 30th Annual International ACM*, Wessel Kraaij, Arjen P. de Vries, Charles L. A. Clarke, Norbert Fuhr, and Noriko Kando (Eds.). ACM, 151–158. https://doi.org/10.1145/1277741.1277770
[58] Paul Thomas, Mary Czerwinski, Daniel McDuff, Nick Craswell, and Gloria Mark. 2018. Style and Alignment in Information-Seeking Conversation. In *Proceedings of the Conference on Human Information Interaction and Retrieval (CHIIR '18)*, Vol. 2018-March. ACM Press, New York, New York, USA, 42–51. https://doi.org/10.1145/3176349.3176388
[59] Paul Thomas, Daniel McDuff, Mary Czerwinski, and Nick Craswell. 2017. MISC: A data set of information-seeking conversations. In *SIGIR 1st International Workshop on Conversational Approaches to Information Retrieval (CAIR'17).*
[60] Nava Tintarev and Judith Masthoff. 2015. Explaining Recommendations: Design and Evaluation. In *Recommender Systems Handbook*, Francesco Ricci, Lior Rokach, and Bracha Shapira (Eds.). Springer, 353–382. https://doi.org/10.1007/978-1-4899-7637-6_10
[61] Anastasios Tombros, Ian Ruthven, and Joemon M. Jose. 2005. How users assess Web pages for information seeking. *Journal of the American Society for Information Science and Technology* 56, 4 (feb 2005), 327–344. https://doi.org/10.1002/asi.20106
[62] Johanne R. Trippas, Damiano Spina, Lawrence Cavedon, Hideo Joho, and Mark Sanderson. 2018. Informing the Design of Spoken Conversational Search: Perspective Paper. In *Proceedings of the Conference on Human Information Interaction and Retrieval (CHIIR '18)* (New Brunswick, NJ, USA). ACM, New York, NY, USA, 32–41. https://doi.org/10.1145/3176349.3176387
[63] Johanne R. Trippas, Damiano Spina, Lawrence Cavedon, and Mark Sanderson. 2017. A conversational search transcription protocol and analysis. In *Proceedings of SIGIR 1st International Workshop on Conversational Approaches to Information Retrieval (CAIR'17).*
[64] Johanne R. Trippas, Damiano Spina, Lawrence Cavedon, and Mark Sanderson. 2017. How Do People Interact in Conversational Speech-Only Search Tasks: A Preliminary Analysis. In *Proceedings of the Conference on Human Information Interaction and Retrieval (CHIIR '17)* (Oslo, Norway). ACM, New York, NY, USA, 325–328. https://doi.org/10.1145/3020165.3022144
[65] Johanne R. Trippas and Paul Thomas. 2019. Data Sets for Spoken Conversational Search. In *BIIRRR@ CHIIR*. 14–18.
[66] Pooja Upadhyay. 2020. Comparing Non-Visual and Visual Information Foraging on The Web. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems Extended Abstracts* (Honolulu, HI, USA) *(CHI '20).* Association for Computing Machinery, New York, NY, USA, 1–8. https://doi.org/10.1145/3334480.3383025

[67] Svitlana Vakulenko, Kate Revoredo, Claudio Di Ciccio, and Maarten de Rijke. 2019. QRFA: A data-driven model of information-seeking dialogues. In *European Conference on Information Retrieval*. Springer, 541–557. https://doi.org/10.1007/978-3-030-15712-8_35

[68] Alexandra Vtyurina, Adam Fourney, Meredith Ringel Morris, Leah Findlater, and Ryen W. White. 2019. Verse: Bridging screen readers and voice assistants for enhanced eyes-free web search. *ASSETS 2019 - 21st International ACM SIGACCESS Conference on Computers and Accessibility* (2019), 414–426. https://doi.org/10.1145/3308561.3353773

[69] Henning Wachsmuth, Martin Potthast, Khalid Al-Khatib, Yamen Ajjour, Jana Puschmann, Jiani Qu, Jonas Dorsch, Viorel Morari, Janek Bevendorff, and Benno Stein. 2017. Building an Argument Search Engine for the Web. In *4th Workshop on Argument Mining (ArgMining 2017) at EMNLP*, Kevin Ashley, Claire Cardie, Nancy Green, Iryna Gurevych, Ivan Habernal, Diane Litman, Georgios Petasis, Chris Reed, Noam Slonim, and Vern Walker (Eds.). Association for Computational Linguistics, 49–59. https://www.aclweb.org/anthology/W17-5106

[70] Peiling Wang and Dagobert Soergel. 1998. A cognitive model of document use during a research project. Study I. Document selection. *Journal of the American Society for Information Science* 49, 2 (1998), 115–133. https://doi.org/10.1002/(SICI)1097-4571(1998)49:2<115::AID-ASI3>3.0.CO;2-1

[71] Iris Xie and Edward Benoit. 2013. Search result list evaluation versus document evaluation: similarities and differences. *Journal of Documentation* 69, 1 (jan 2013), 49–80. https://doi.org/10.1108/00220411311295324

[72] Hamed Zamani, Bhaskar Mitra, Everest Chen, Gord Lueck, Fernando Diaz, Paul N. Bennett, Nick Craswell, and Susan T. Dumais. 2020. Analyzing and Learning from User Interactions for Search Clarification. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*. Association for Computing Machinery, New York, NY, USA, 1181–1190. https://doi.org/10.1145/3397271.3401160

[73] Saizheng Zhang, Emily Dinan, Jack Urbanek, Arthur Szlam, Douwe Kiela, and Jason Weston. 2018. Personalizing Dialogue Agents: I have a dog, do you have pets too?. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, ACL 2018*, Iryna Gurevych and Yusuke Miyao (Eds.). Association for Computational Linguistics, 2204–2213. https://doi.org/10.18653/v1/P18-1205

[74] Yinglong Zhang and Robert Capra. 2019. Understanding How People use Search to Support their Everyday Creative Tasks. In *Proceedings of the Conference on Human Information Interaction and Retrieval (CHIIR '19)*, Leif Azzopardi, Martin Halvey, Ian Ruthven, Hideo Joho, Vanessa Murdock, and Pernilla Qvarfordt (Eds.). ACM, 153–162. https://doi.org/10.1145/3295750.3298936