Knight, S., Bowdler, I., Ford, H., Zhou, J., (2024). Visual Scoping Review of How Knowledge Graphs and Search Engine Results Page Designs Represent Uncertainty and Disagreement

10.1108/ILS-02-2024-0016

emerald PUBLISHING

Information and Learning Scie

A Visual Scoping Review of How Knowledge Graphs and Search Engine Results Page Designs Represent Uncertainty and Disagreement

Journal:	*Information and Learning Sciences*
Manuscript ID	ILS-02-2024-0016.R1
Manuscript Type:	Article
Keywords:	visual design, materials review, information seeking, search engines, knowledge graphs, systematic review

A Visual Scoping Review of How Knowledge Graphs and Search Engine Results Page Designs Represent Uncertainty and Disagreement

Abstract:
Purpose: Informational conflict and uncertainty are common features across a range of sources, topics, and tasks. Search engines, and their presentation of results via Search Engine Results Pages (SERPs) often underpinned by knowledge graphs (KGs) are commonly used across tasks. Yet, it is not clear how search does, or could, represent the informational conflict that exists across and within returned results. We review KG and SERP designs for representation of uncertainty or disagreement.

Approach: We address our aim through a systematic analysis of material regarding uncertainty and disagreement in knowledge graph (KG) and search engine results page (SERP) contexts. Specifically, we focus on the material representation – user interface design features – that have been developed in the context of uncertainty and disagreement representation for KGs and SERPs.

Findings: Searches identified n = 136 items as relevant, with n = 4 sets of visual materials identified from these for analysis of their design features. Design elements were extracted against sets of design principles, highlighting tensions in the design of such features.

Originality: We conclude by highlighting two key challenges for interface design, and recommending six design principles in representing uncertainty and conflict in SERPs. Given the important role technologies play in mediating information access and learning, addressing the representation of uncertainty and disagreement in the representation of information is crucial.

Keywords: visual design, materials review, information seeking, search engines, knowledge graphs, systematic review

Across a wide range of contexts in professional, civic, and personal decision making, people engage in information seeking that may expose them to conflicting information. Exposure to conflicting information is common in information seeking regarding health and science topics including COVID-19 (Nagler *et al.*, 2020a), and that exposure can lead to confusion, and a distrust of experts beyond the target topic (Nagler, 2014a), suggesting a critical research need in understanding how to help people navigate such information (Knight *et al.*, 2017; Nagler, 2014a; Nagler *et al.*, 2020a). Consider the ways our knowledge about facemasks for COVID-19 shifted, with at least two major bodies changing views at different points, and disagreement grounded in supply chain issues (are there enough?), behavioural concerns (will masks give a sense of artificial safety/produce compliance fatigue?), and transmission uncertainty (is it airborne?), or other health decisions such as vaccination or treatment options (Nagler *et al.*, 2020a). In such cases, we are faced with uncertainty and expert-disagreement. The underlying causes of this uncertainty and disagreement, and its representation to both experts and lay people is a source of significant attention in communication of science (van der Bles *et al.*, 2019; Deroover *et al.*, 2023).

Across a range of topics and contexts, information is frequently accessed via technologies underpinned by knowledge graphs (KGs), including search engines and search engine results pages (SERPs). Knowledge graphs are often used to store descriptions of objects, events, situations or concepts and the relationships between these entities. Knowledge graphs have historically been used by search engines like Google and Bing (as well as question-answering services like WolframAlpha, Apple's Siri and Amazon Alexa) to present facts associated with particular entities in knowledge panels and as answers to user questions (Hogan *et al.*, 2021). Asking who is the President of the United States on Google and Alexa, for example, will surface a series of statements about Joe Biden including where and when he was born, which organisations he founded and who his grandparents are – all of which are powered by knowledge graphs.

However, it is not clear how these interfaces do, or could, support navigation of conflict or uncertainty-oriented information features through their user interface designs. The representation of

conflict and uncertainty are vital for improving decision-making by individuals using these tools because the flag of a statement's potential multiplicity among different groups can be vital for determining the most appropriate actions. Whether a result is in dispute, scientific results are generating conflicting results, or whether details about an event are still being confirmed are all vital to more fully understanding a particular topic. When all statements are presented as if they are similarly certain and uncontested, we risk undermining the ability of individuals to correctly interpret such information.

This paper thus develops a novel systematic approach to analysis of the visual artefacts of these systems, to investigate their representation of uncertainty and disagreement. Our first research question aims to provide insight into the degree of focus on these topics in research to date, in order to identify existing design attention and implications, through analysis of the aims of published works:

> *Research Question 1 What are the aims of research regarding representation of uncertainty or disagreement in Knowledge Graphs (KGs) and Search Engine Results Pages (SERPs)?*

Our second question aims to understand the scope of designs investigated and adopted in tools addressing our focal concern:

> *Research Question 2 How have KG and SERP interfaces been designed to represent uncertainty or disagreement including (a) their general design form; (b) the design heuristics comprising these forms; (c) their implications for cognitive biases in human-information interaction?*

Representation of Uncertainty and Disagreement in SERPs and KGs

Common approaches to tackling the problem of confusion and distrust resulting from exposure to expert-expert conflict focus on technical solutions (such as PageRank or foregrounding of source credibility features) and end-user education (typically targeting knowledge about the nature of science, or credibility assessments). In both cases, the focus is often on informational cues and credibility markers as a function of page rank or in guiding users to high quality sources. However, as COVID-19 has demonstrated, there are many domains in which our knowledge is unstable and evolving, and on which experts have legitimate disagreement that is not well-addressed by strategies that seek to prioritise sources or support evaluation of source-credibility features. These endemic features of the nature of science interact with individual differences (i.e., people's beliefs and attitudes), and are mediated by tools such as SERPs and knowledge graphs.

Moreover, when engaged in search tasks users are capable of distinguishing source quality through their selection of results, including when SERPs are experimentally manipulated to put sources with higher credibility lower in the rank (Salmerón *et al.*, 2013). Some pilot work (Novin and Meyers, 2016) investigated how conflicting results inserted into a SERP at different ranks impacted how students ranked and summarised the information, focusing on how conflicting content was drawn on in how users constructed summaries of the information. However, that preliminary work did not investigate the impact of such insertions on perceptions of conflict among experts (or flow-on impacts of trust in experts from that), nor how these perceptions were resolved. Further work has investigated individual differences in how people integrate information contained in multiple sources, based on how people justify knowledge claims (Ferguson, 2014; Huang, 2020; Kammerer *et al.*, 2015; Knight *et al.*, 2017), however this work has not explicitly probed perceptions of conflict on the basis of SERP features, nor the ways in which these SERP features impact how users synthesise the information.

Existing SERP tools present challenges in navigating uncertainty and disagreement. SERP design and evaluation is precision oriented; that is, SERPs are designed to present the minimal set of information to provide the (definite article) answer to a query. This precision orientation in SERPs may stoke confusion regarding legitimate parts of science such as uncertainty and conflict, and lead to greater acceptance of low quality sources, with SERP designers seeking to provide "seamless single-source answers to complex queries" (Novin and Meyers, 2017, p. 181), with the implication of an "essential truth waiting to be uncovered by an algorithm" (Novin and Meyers, 2017, p. 181). This design is problematic insofar as it ignores the complexity and uncertainty of topics, while also encouraging users to not pursue further research on complex topics; 'slow search' has not been a focus of design (Teevan *et al.*, 2014). The design challenge is significant in that SERPs must both (1) avoid creating a sense of 'false balance', while (2) also presenting change and diversity or multiplicity

in claims in the context of human factors such as confirmation bias and contrast effects (perceptions of divergence in contexts where stimuli are presented alongside each other). There is thus an important gap in understanding how SERP features might be designed for, and exert influence on, user navigation of conflict and uncertainty.

User Interfaces as Mediator and Lens

User interfaces mediate how people access and interact with information, while also providing a lens onto how the expression of information is structured through tools (Hearst, 2009). Modern search interfaces are a reflection of the long recognised challenges of developing user interfaces to meet varied information needs and to address models of human information seeking (Hearst, 2009). A variety of models have been developed to support different types of search, including surface changes such as varying the number of results displayed, as well as backend algorithm changes to promote diversity-aware search results to show a greater range of responses (Verbeke *et al.*, 2009), and more structural changes around the presentation of results. However, there is no clear overview of the variety of interfaces developed, a significant gap given that, in order to support navigation of disagreement and uncertainty, the confluence of interface, information-structure, and user behaviour must be addressed.

Moreover, in seeking to develop a review of user interfaces it is not clear *how* such a review should be conducted. There are well established methods for systematic review of various kinds, with the EQUATOR Network (Enhancing the QUAlity and Transparency Of health Research) hosting many (See, "Ethical issues (consent etc.) | Report Sections | EQUATOR Network", n.d.; Wharton, 2017), 'empirical standards' emerging to fulfil a similar role in computing fields (Ralph *et al.*, 2021) and fields such as education (American Educational Research Association, 2006). However, across reporting standards, guidelines and checklists, there is no clear guidance regarding the review of non-textual material aspects of the research such as user interfaces, which would guide a reviewer in developing a synthesis regarding approaches adopted across studies. Indeed, a recent review of the kinds of reviews used in the Human-Computer-Interaction (HCI) field (Stefanidi *et al.*, 2023) identified a lack of clarity in approaches and their adoption in the field, with a relatively small number (n = 189) of papers adopting a 'review' method. Of these, n = 54 were identified as "Artefact literature review contributions [that] arise from analysing work on artefacts with the goal of classifying them" (Stefanidi *et al.*, 2023, p. 6), however, such reviews would include consideration of features such as design guidelines or other non-visual characteristics of the literature.

Developing a Review Approach for Visual Materials

While no guidance regarding non-textual features exists to our knowledge, recent methods have been developed for the analysis of study *materials* in contrast to outcomes or results, including a method for the systematic review of methods (Gentles *et al.*, 2016) and normative or ethical features in research (Kahrass *et al.*, 2021). In addition, approaches in design research provide inspiration insofar as they provide a theoretical lens for the ways that material artefacts are encoded with features that can be understood to target interaction models and outcomes (see for example, Cobb *et al.*, 2003; Easterday *et al.*, 2016; Hevner *et al.*, 2004; McKenney and Reeves, 2013; Oppl, 2022; The Design-Based Research Collective, 2003; Wilson *et al.*, 2017). User interfaces provide a form of a designed material that helps probe the representation and mediational properties underpinning it, through analysis of the 'claims' made by those artefacts (Carroll and Rosson, 1992; for a critical review, see, McCrickard, 2012). This kind of claim analysis, arising from the field of human-computer-interaction, provides a useful approach for understanding the model of the user and data-structure through analysis of tools and their implied use, providing a probe into designer assumptions (Moran and Carroll, 1996).

Clear articulation of methods for the review of user interfaces and visual design is important because these interfaces may encode assumptions and knowledge that may go unexpressed in explanatory text. There is a significant body of work regarding usability heuristics and the design process, and methods for evidence synthesis regarding targeted theoretical features of interfaces have been conducted. However, these, respectively, (1) focus on usability assessment, but may not articulate an underlying description of the features; (2) may be 'ground up' descriptions of design, without a clear role for

literature and understanding of prior models in the design process; and (3) tend to focus on evaluation at a theory level, while there may be important features in the operationalised design that are not salient to this evaluation.

To conduct a systematic review of uncertainty and conflicting representations in information seeking and decision support tools, we drew together different approaches from methods and design reviews. Our approach draws on Gentles et al.'s (2016) seven best practice principles and strategies for a rigorous systematic methods review, as summarised:

1. Identify a way to delimit a manageable set of literature to address the needs of the review;
2. Consider sources beyond traditional databases (e.g. Google Scholar, or in our case regular Google) as relevant material might be there;
3. Note that metadata (Title, Abstract, Keyword) search is unlikely to be possible via non-traditional databases, and thus purposive sampling may be required;
4. Consider a purposeful strategy to achieve adequate conceptual coverage of relevant material;
5. Consider inductive/deductive approaches to data-abstraction to allow for novel theory to emerge;
6. Key terms may vary in use across materials and thus consider the unit of analysis (e.g., definitions provided in various materials, and how they might be expressed);
7. Develop an appropriate data abstraction approach to ground qualitative analysis in the data in order to provide warrants for abstracted claims.

In addition, a small number of articles in the space of interaction and design provide examples of materials analysis, including Isenberg et al.'s (2013) systematic review of evaluation practice in visualisation research, however approaches to such review are not well defined (Stefanidi *et al.*, 2023, p. 6). To identify relevant material, we combined literature we were previously aware of with initial exploratory searches for review approaches in design identifying three relevant publications (Vial *et al.*, 2022; Wright *et al.*, 2019; Yilmaz *et al.*, 2016).

These were augmented by a search of Scopus (all indexes) using a query intended to capture the [object] [review method] relationship (with the 'within' operator requiring that the items appeared within 10 terms of each other[i]):

> *ALL (("user interface" OR "search interface" OR "search engine results page" OR "visual design" OR "interaction design" OR "visual representation") W/10 ("systematic review" OR "scoping review" OR "evidence synthesis" OR "research synthesis" OR "literature synthesis" OR "rapid review" OR "review of reviews" OR "evidence map" OR "mapping review")) AND (LIMIT-TO (LANGUAGE , "English"))*

This search returned 142 results. The titles and abstracts were reviewed in-browser using the find-in-page function with terms from the search query ('interface' 'search' 'visual' 'interaction'), to identify items where the terms were being used in a way consistent with our aims; i.e., they were applying a [review method] to an [object] or artefact representing a visual interface or interaction feature. Five results were relevant based on this screening. Of these, although relevant to the general problem space, four did not contain methodological information of relevance to developing our approach: (1) the review of HCI reviews (Stefanidi *et al.*, 2023), (2) an overview of visualisation design in clinical decision support (Wright, 2023), (3) a systematic review of interaction designs for group recommender systems (Alvarado *et al.*, 2022), (4) a review of media visual designs in health, which included analysis of visual components, adopting a specific social semiotic framework (Riddell *et al.*, 2022). One piece focused on a systematic analysis of "visual representations for analysing collaborative discourse" (Hu and Chen, 2021), categorising visualisations according to their type (bar, scatter, etc.), interactive properties, and underpinning design principles. Across results, processes for obtaining visualisations, and analysis of any novel features is unclear, and appears to be a gap in the literature.

Across Wright et al.'s (2019) systematic review of patient information displays and Vial et al. (2022) exploratory mapping review of online mental health interventions, features analysed include: the material object, its aims, and visual characteristics; the context of the study and design approach;

and who was involved in development and use of any artefacts. Yilmaz et al. (2016) combined a large database from previous empirical studies to extract concept variations from different design processes and outcomes. From these extracted concepts the authors created 77 design heuristics, which provide strategies to overcome design problems and achieve design goals. For given artefacts or design concepts:

1. Analyse for "functionality, form, and user-interaction features" using a content analysis of "needs, design criteria, functions, and the design solution was performed for each concept" (Yilmaz *et al.*, 2016, p. 100)
2. Identify composite design heuristics within/across the artefacts, e.g. a search bar
3. Iteratively compare design heuristics across artefacts for commonalities in order to generate generalised heuristics (while maintaining the link to the original underlying concept)

These pieces provide some resources for considering the finding, extraction, and analysis of artefacts in published and grey literature domains, although gaps remain regarding appropriate search strategies, identification and inclusion criteria for (visual) artefacts, and the selection of appropriate frameworks for the content analysis of these for reporting in evidence synthesis approaches. Based on this literature, and critical discussion of design principles (Bakker, 2019), we adopt the following to assist in this analysis:

1. **Design elements** are parts of an artefact with a shared form, such as a search box, a results list, etc.; they may be composed of sub-features.
2. **Design principles** are expressions of guidelines or heuristics to support design of a material artefact that addresses the intended user-artefact interaction and outcome, i.e., they succinctly communicate how design decisions can achieve design objectives.
3. **Design heuristics** provide material examples of how design principles can be instantiated into artefacts through combinations of design elements (collections of these may be abstracted and linked to design patterns).

Methods

Search strategy

Overarching approach

The search strategy focused on articles discussing representations of uncertainty, and disagreement in knowledge graph and search engine results page contexts that contained visual representations of methods to foster this navigation to address our research questions.

Eligibility criteria overview: Knowledge graph and SERP models to represent uncertainty and disagreement appear in a number of contexts across both scholarly and non-scholarly venues, and in consumer-oriented tools. As such, searches were inclusive of grey literature, and source-type was not a primary exclusion criterion. That is because our primary focus was design features or characteristics, rather than empirical evaluations such as articles focused on systematic comparison of intervention designs or usability.

Search strategy overview: A set of searches was conducted across scholarly and grey literature with each search developed with respect to the particular affordances (e.g., Boolean search) of the target platform. This strategy is intended to retrieve an adequate range of results to represent the scope of variation and magnitude of available sources, while not being exhaustive in nature.

Quality appraisal: No quality assessment was made of sources retrieved. That is because quality standards are typically applied in the analysis of empirical interventions (including technology designs) and their impacts, while our focus here was on the design implementations developed (whether deployed or not) more broadly.

Venues / sources

Sources were identified through searches of:

1. scholarly indexes (Scopus; Google Scholar),
2. general Google searches, and citation forward/backward chaining particularly to find live prototypes or non-scholarly media (blogs, etc.) discussing them.

Inclusion and exclusion criteria

A set of criteria was developed, excluding

1. **on abstract screening**, items that mention the target concepts (disagreement and uncertainty) and technologies (SERPs and KGs) but without the intersection being the focus of the work; and
2. **on full-text screening**, items that contained no visual design element or representation of any SERP/KG output
3. **on-visual artefact screening,** items for which the visual elements did not represent user search interfaces were excluded

Table 1 summarizes the inclusion and exclusion criteria used in this paper.

Table 1: Inclusion and Exclusion Criteria

Criteria	Inclusion	Exclusion
Literature Type	Journal articles, book chapters, conference proceedings, magazines.	Theses and dissertations
Language	English	Non-English
Timeline	Everything up to October 2022	(Implied: anything published after October 2022)
Subject Area	Materials discussing representation of uncertainty, and disagreement, in knowledge graph and search engine results page, contexts.	Not discussing KG/SERP AND uncertainty/ disagreement
Other		No visual design element Visual design element not related to SERP interface

Searches conducted

In the initial stage of the search process, keywords and databases were identified. Web of Science (WoS) and Association for Computing Machinery (ACM) Full-Text Collection were selected to include a multi-disciplinary perspective and coverage through WoS multi-disciplinary index, and ACM's focused coverage in the field of computing.

Keywords were selected from the project's research questions and project aims and tested in a Google Scholar citation search for papers citing Marchionini's (2006) seminal 'Exploratory search: from finding to understanding', to ensure we were achieving satisfactory conceptual coverage of the field (Gentles *et al.*, 2016). Testing keywords in Google Scholar allowed iterative addition of concepts and keywords, following Gentle et al.'s (2016) advice that you cannot anticipate some concept prior to encountering them in the review process. It further allowed us to consider potentially inconsistent terminology across publications and select keywords that account for this inconsistency (Gentles *et al.*, 2016). Based on this exploration, ten keywords were identified, four relating to terms expressing uncertainty or conflict, and six relating to technology terms.

Following this exploratory search, the keywords were updated to those in Table 2, with terms required from both columns, uncertainty and disagreement, or information evaluation and decision support (column 1), and terms indicating web information seeking (column 2).

Table 2: Updated Search Terms

	Term expressing uncertainty/conflict	**Technology terms**
	Uncertain*, disagree*, conflict*, "decision support", "decision aid", "evaluation of search results"	"knowledge graph", "search engine results page", "SERP", "information seeking behavio*", "online", "web", "digital"

Keywords were combined using Boolean AND/OR terms to form a keyword string, used to search an article's title, abstract and keywords (Full search strings can be found in the Supplementary materials).

Search Results

26 articles that did not appear in the database searches were manually added to the pool of articles. Many of these were non-scholarly sources, such as news articles or prototypes, that focused on new design principles for SERPS. These items were obtained through forward/backward citation chasing including to non-scholarly sources, and targeted searches on non-scholar google.

Duplicate articles across the database were merged using Zotero and the abstracts were then screened according to our inclusion criteria, with 84 records excluded at abstract screening (n = 8 because the full text could not be obtained; n = 1 because language was not English; n = 76 for relevance; note although multiple exclusions may apply to a single item, once classified they were not further coded).

The full texts of 176 remaining articles were scanned for eligibility, with 39 excluded at this stage. This left 132 articles for analysis, with 3 items containing relevant visual elements, to which one additional piece (a new GPT based search tool) was manually added for inclusion in 2023.

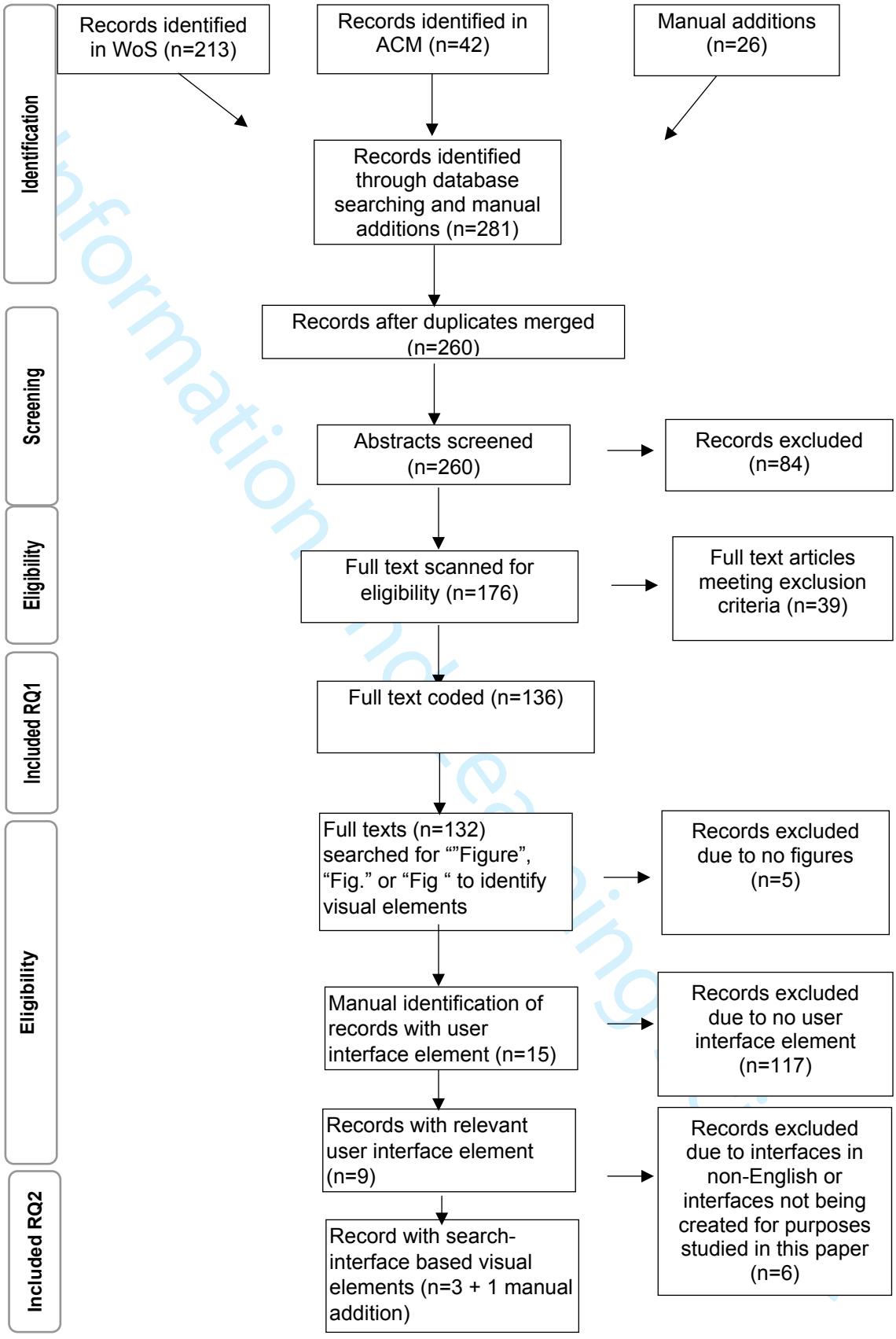

Figure 1 – Adapted PRISMA Flow diagram depicting items retrieved and included for RQ1 and RQ2

Data management and processing

The following procedure was used to search, track, manage, and annotate items, largely within the open source Zotero reference manager (Center for History and New Media, n.d.).

1. **Search:** Searches were conducted in their respective platforms as described above. Zotero was used to track systematic searches, and maintain link of items to their search source.
    a. For each search conducted, where possible all items returned by the search were saved into a parent collection.
    b. Items were screened, and added to collections indicating inclusion/exclusion
    c. Item metadata fields were edited in Zotero, using manual checking where necessary (particularly for 'rights' or copyright status).
    d. Zotero was used to deduplicate items (in Zotero, items may exist in multiple collections, thus they remain in the parent search, and any other searches, and included/excluded collections; deduplication retains collection information for both versions, merging their metadata).
2. **Processing:** Was conducted on items saved in Zotero, using the built-in notes and tagging features, with annotations created for extracted features from the PDFs, and documents tagged with overall topic tags. The annotations were exported to child notes, with each note instantiating a target feature for analysis (e.g., audience, visual-design).
3. **Data extraction:** Data extraction, including selection of features and their export from Zotero for further processing, followed the procedure set out below
4. **Data export:** A combination of Zotero addons and scripts was developed to extract notes from Zotero for further analysis. Notes were exported with embedded images using a markdown export with YAML header containing the note (and parent item) meta-data, including the tags for the particular note. The unit of analysis here is the item/article. The Zotero addon 'Better Notes' was used for this purpose, as it provides a method to export to markdown with image export, to include a YAML header, and to modify the export filename (in this case, to export the parent item key, such that items can be linked to notes).

Data extraction approach

Each article selected was coded in Zotero (steps 2-3 above) to extract data relating to design features. Based on the limited literature regarding evidence synthesis for design objects our initial data extraction included the following, into separate tagged notes:

1. Aim(s) – extracted research questions or aims in papers, and stated purposes or rationales in non-scholarly sources
2. Visual design (object, form, interactions) – extracted visual representations of interfaces and their descriptions
3. Actor(s) / audience(s) – extracted any information regarding who the tool targets/its intended users are
4. Theme(s) or context(s) of the work – extracted components of the articles related to the key words searched and a theoretically grounded set of issues (discussed below)

The user interfaces extracted were divided into two categories: interfaces related to SERPs and interfaces related to decision supports. As these interfaces have different design elements and purposes, such as SERPs providing information to the public, and decision support tools providing information to specialists within a workplace context, we drew on different analysis methods. Those relating to SERPs were analysed based on Novin & Meyers (2017) work as they looked specifically at how the design principles of SERPs can interact with user cognitive biases.

A primary reviewer extracted data, using Zotero tagged notes to excerpt text indicating fields described above. A secondary reviewer assisted to check the extractions captured the intended fields, and agreement regarding inclusion of the items for each question. Following piloting of the approach, no conflicts between reviewers were found.

Our approach drew on directed content analysis, using both theory regarding analysis of design artefacts, and prior research regarding SERP design, to inform the identification of groups of research (Hsieh and Shannon, 2005). The aim of this approach is theory development, in this case around analysis of designs for uncertainty and disagreement in KGs and SERPs, grounded in prior research. As such, our sub-codes (see Table 3) were developed through analysis of the data fields to identify how our target concerns are addressed in the literature, and where one might find design objects (notably, the code: SERP-design-content). These notes and their tags, alongside review of the full article texts were then exported to spreadsheet software, for finalisation of the codes.

Table 3: Data fields extracted and non-mutually exclusive tags applied

Type	Code	Explanation	**Example** (* indicates wildcard stemming)
Information context of work	*Disagreement/ Conflict*	Articles that contain reference to conflicting information or disagreement	disagree*, conflict*,
	Uncertainty and risk	Articles that contain references to uncertainty (or where absent, risk)	Uncertain*
	Multiple-perspectives	Articles that contain reference to multi-perspective answers	multi-perspective, multiple perspective (supplemental searches)
	Decision-support	Articles that contain reference to decision supports or decision aids	"decision support", "decision aid"
	Info-seek-evaluation	Articles that refer to information-seeking behaviour	"evaluation of search results"
Contribution-space of work	*Detect-UDFeatures-in-data-structures*	Focus on methods for identification of features of uncertainty or conflict in data structures	E.g., work identifies confidence of claims, uses clustering to identify claims on same/similar topic, etc.
	Resolve-UDFeatures-in-data-structures	Focus on methods for resolution of features of uncertainty or conflict in data structures	E.g., work uses ranking or clustering to select or merge best candidate claims
	Represent-UDFeatures-in-data-structures	Focus on approaches for representing features of uncertainty or conflict into data structures (e.g., in decision support tools or/and knowledge graph structures) with practical application	E.g., work stores uncertainty or conflict as feature in data (e.g., as a confidence measure) such that multiple candidate outputs could be provided to a user
	Understand-navigation	Focus on understanding how users navigate features of uncertainty or conflict in information	E.g., work investigates how users engage with features of uncertainty or conflict in information
	Design-for-navigation	Focus on understanding interface design for the navigation of uncertainty or conflict in information	E.g., work makes design claims, or includes design focus, regarding navigation of features of uncertainty or conflict in information
	Other	Articles that do not fit into the above categories	

Themes Addressed in Interface Design

To evaluate the visual components of SERP user interfaces, Novin & Meyer's (2017) analysis was drawn on. Novin & Meyers (2017) reviewed the design of SERPs and identified that they can evoke four kinds of cognitive bias in people (though their study focused on post-secondary students specifically): priming, anchoring, framing and availability bias. In this way, SERP designs may interact with the underlying content they serve – and source-biases in this content – and search-users,

and their cognitive biases. While bias occurs both on the user and system side of search engine use, they outlined that changes can be made to SERPs to avoid creating this bias in users. In detail:

1. **Priming** effect refers to when "users are visually drawn to a familiar feature on a SERP" (2017, p. 180). For example, a main priming feature is Google's use of an answer box to provide a single answer to a search. This can be problematic as users are drawn to these familiar sources and often overlook other information. Novin & Meyers (2017) suggest that in response, SERPs should "visually identify conflicting information more explicitly by drawing a user's attention to more perspectives on a topic" (2017, p. 181).
2. **Anchoring** refers to how hierarchical lists assume the importance of the first result, which might emphasise a particular view to users that is not representative of different perspectives on a topic. Users are also biased towards the first result they see, meaning that the first result can affect a user's impression of the significance of the next results.
3. **Framing** refers to how people are influenced by how information is presented, and whether information is presented to make users aware of, and interested in, conflicting views. For example, SERPs are often designed to mistake conflicting perspectives with incorrect or irrelevant results and rank them lower, meaning users do not contextualise different perspectives. Novin & Meyers (2017) argue that SERPs can do more to enable users to understand the relationship between multiple sources.
4. **Availability Bias** refers to how users are more drawn to sources that are easy to access and understand than potentially useful sources that are more complex. A salient challenge when designing SERPs is how to show the usefulness of a source that might undermine assumptions that users have gained from other – perhaps simpler – sources.

Results

RQ1: What are the aims of research regarding representation of uncertainty or disagreement in Knowledge Graphs (KGs) and Search Engine Results Pages (SERPs)?

A set of articles (n = 136) was coded for topic relevance prior to excluding items that did not contain a visual design element. As described above, each of these papers was reviewed with key statements capturing the data fields in Table 5 captured, to inform our understanding of the focus of the works and their design approaches for SERPs and KGs.

Across the published works, as indicated in Table 4, most discussed issues around uncertainty (n = 73), rather than conflict or disagreement in information (n = 39), often with limited discussion. While uncertainty and disagreement were included, they were largely seen as an issue (such as accuracy) to be resolved (n = 64), rather than a feature of the complexity of scientific or other specialised knowledge, and the importance of representing this complexity to users. Investigation of design features (further probed in RQ2) was uncommon (n = 12), with no well-established design models retrieved.

Many items (unique n = 103) addressed issues in the detection (n = 78) or representation (n = 72) of features of interest into data structures, typically in knowledge graphs; in addition, many items treated these features as something to be resolved (n = 1 focused solely on resolution, n = 64 overall). In these cases, the aim was generally to improve accuracy or to cluster claims into single items, rather than to present uncertainty or multiplicity (and conflict within that) as a feature of the underlying information being processed or represented. Across the works, the development of detection methods for incorporating uncertainty or disagreement into knowledge graphs or models incorporating knowledge graphs often treated conflict or uncertainty as a form of error, focusing on these novel techniques as the core contribution (rather than human use of them). Of items that discussed representation of information, articles often had a specific practical application such as targeting medicine contradictions in a hospital setting. There were seven articles that fit into the 'other' category and focused on areas such as crowd sourced knowledge graph generation, affordances of platforms, cognitive systems, using knowledge graphs to find out information on a specific topic and creating algorithms to sort conflicting facts.

N = 24 distinct items addressed design to navigate, or the understanding of navigation of, features of uncertainty and disagreement. We reviewed these to identify design implications

(principles, objectives, or needs), that might be drawn on. Most do not provide clear statements of implications for design:

- Three discuss philosophical issues relating to the nature of search engines and features of learning (Heersmink and Knight, 2018; Knight, 2012; Tavani, 2020)
- Five discuss cognitive process involved in comprehending multiple sources (including on the internet) (Ferguson, 2014; Kammerer *et al.*, 2015; Knight *et al.*, 2017; Kobayashi, 2014; Salmerón *et al.*, 2013), with Salmerón et al., noting that although people do favour the 'top link' in SERP result, this was less likely for controversial topics where the top link did not match relevance expectations, and less likely in bookmarking behaviour.
- Five discuss health information seeking and exposure to conflicting information, making suggestions for public health campaigns and communication (Nagler, 2014b; Nagler *et al.*, 2020b; Rains, 2014; Scarton *et al.*, 2018; Vasantavada *et al.*, 2022)

A further set of eleven discuss issues around SERP design and interaction. Five highlight important areas for study, but without operationalizable implications for design, including: through making a design recommendation to incorporate navigation of uncertainty into platforms, with respect to the particular information environment of those systems (Chowdhury *et al.*, 2014); flagging need to understand interaction between user behaviour and features of ranking, topic, and type of source (Cano-Oron, 2019; Haas and Unkel, 2017); or highlighting concerns regarding user evaluation of misinformation in SERPs (Shah, 2020; Song and Jiang, 2022). Three make concrete suggestions for design, including inclusion in SERPs of information to support interpretation of information cues (although with no detail of how this might be done) (Kattenbeck *et al.*, 2019); suggestion that health SERPs might be dynamic based on query-type and -stage, including use of KG structures to support users in identifying connections between concepts (Chen *et al.*, 2020); and finally, suggesting that KG-based health information could be provided in knowledge panels in searching on a controversial topic (vaccination) to provide unbiased and comprehensible factual information to users (Ludolph *et al.*, 2016). A further three with clear implications for design, are discussed in detail below (Chen *et al.*, 2022; Novin and Meyers, 2016; Schwartz, 2018).

Table 4 – Summary of topics of included article aims and contributions

Type	Code	n
Information context of work	*Disagreement/ Conflict*	39
	Uncertainty and risk	73
	Multiple-perspectives	10
	Decision-support	30
	Info-seek-evaluation	27
Contribution-space of work	*Detect-UDFeatures-in-data-structures*	78
	Resolve-UDFeatures-in-data-structures	64
	Represent-UDFeatures-in-data-structures	72
	Understand-navigation	22
	Design-for-navigation	12
	Other	7

RQ2a: How have user-facing KGs and SERPs been designed to represent uncertainty or disagreement in (a) their general design form (b) the design elements comprising these forms; (c) their implications for cognitive biases in human-information interaction?

Four artefacts were identified as meeting inclusion criteria including relevance of intended use (disagreement and uncertainty), and the availability of a visual artefact or interface. These include three tools that are examples of multi-perspective search interfaces:

1. Recent work (Chen *et al.*, 2022; *Navigating Information Pollution in the Time of COVID-19*, 2020) to represent different perspectives on user input queries via a search input with multi-perspective and multi-source SERP output (see Figure 2);
2. Bing's implementation in some searches of a 'multi-perspective' feature, which displays two 'perspectives' and their key sources as output to a user input query (Bing, 2018; described in, Schwartz, 2018);
3. And emerging generative AI tools, such as the Consensus app which allows users to input queries via a search, generating output from a corpus of expert sources (semantic scholar) to show the different views and give a 'consensus' overview (Consensus, 2023).

In addition to these, we identified that the Contropedia tool (Weltevrede and Borra, 2016) addressed concerns aligned with our aims, and included a visual interface representation. Contropedia aims to represent 'controversy' in the production of the open knowledge encyclopedia Wikipedia (Weltevrede and Borra, 2016). It does this through analysis of editing patterns, including reversions, that indicate the controversial and changing character of knowledge over time, challenging notions of knowledge stability.

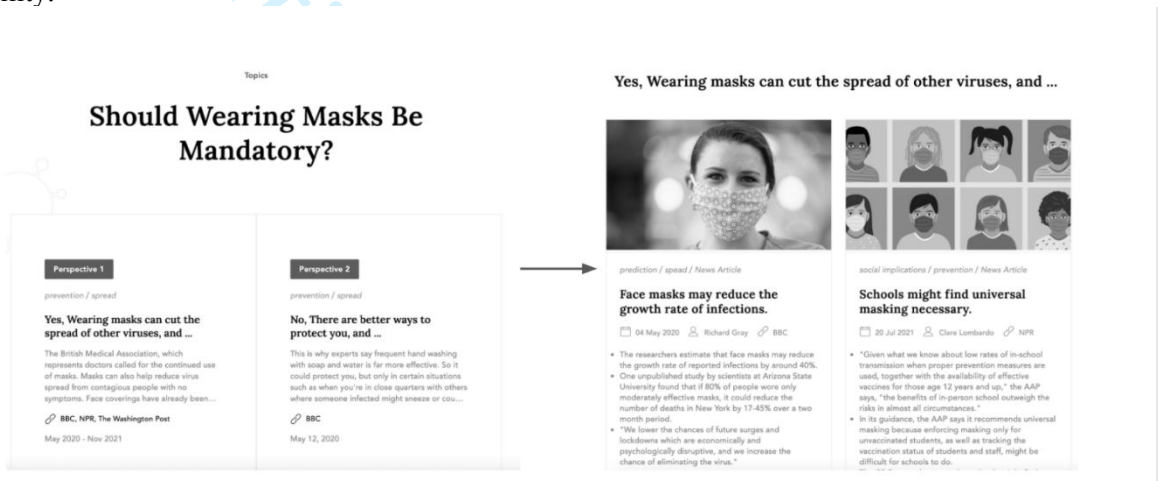

Figure 2: Example interface, from Chen et al., "Clicking on the Perspective 1 on the left page will direct [users] to the page on the right" (2022, p. 298, under a cc-by license; original in colour),

Design Principles and Tensions in Multi-Perspective Search

Drawing on the design approach described above, these interfaces were analysed to understand how they address uncertainty and disagreement, as Table 5 summarises. We sought to distil the *design elements* of each interface, drawing attention to the components of these artefacts particularly where they held a shared form (such as a search text entry box). Each element is mapped to the stems or design objective for four overarching principles (using a shorthand in the table), where the table provides possible ways in which the objective can be addressed:

1. Provide an informative overview of information returned
2. Make different perspectives clear, and make it easy to navigate these perspectives through user interaction
3. Make the connection of perspectives to sources clear, without overwhelming users with redundant information
4. Provide source features

These principles were derived based both on the concerns expressed in Novin and Meyers (2017), alongside our reading of the sources, their key concerns, and their commonalities. In addition, we draw on Novin and Meyers (2017) as a lens onto the ways that each design element relates to the four biases they highlight. In particular, we note that items 3 and 4 provide *framing effects* primary around valuing the representation of *contested information*, and items 1 and 2 provide *framing effects* primarily around valuing the representation of *multiple perspectives*.

Table 5: Overview of Design Elements, Principles, and Biases instantiated in Items

Bias	Principle	Design element	1.	2.	3.	4.
P	Information overview	Includes demarcated box around key feature (search and results)	x	x	x	
P		Use of images to draw attention	x			
P		Design draws immediate attention to multiple perspectives	x	x	x	
Av		Summary of information presented in dot points or short sentences	x		x	
P Av		Knowledge panel Answers in paragraph form		x		
An	Perspective navigation	Displays information/perspectives in non-hierarchical lists	x	x		
An		Juxtaposition of two perspectives side by side	x	x		
An		'vs' is placed in the middle of each answer box to symbolise difference		x		
P		After clicking one perspective it includes the ability to click on another perspective without utilising back function	x			
Av		Summarises perspectives	x	x	x	
P & An F		Use of colour to indicate perspective /contested info - the redder a word is, the more controversial it is (4).			x	x
P		Uses a timeline feature to indicate controversy over time				x
P		Disputed neutrality warning in Wikipedia article				x
P		Includes a 'Consensus Meter' feature with percentages of search results under 'Yes', "Possibly' and 'No', 'Consensus Meter' is colour coded (green, yellow, red)			x	
Fr Av	←cross-cutting→	Makes distinct perspectives visible with sources	x	2 only	^	
Av		Limits redundant perspectives (e.g. prioritises reputable content and reduce repetition)	x	x	x	
P	Source features	Includes disclaimers emphasising the need for further research and not to take 'answers' at face value			x	
Av		Multiple sources for information	x	x	x	
Av		Information has clear source/s	x	x	x	
Av		Ability to choose domain of search (e.g. news articles, medical information)	x			
Av		Includes extra sources details in text buttons, such as type of analysis (case report, meta analysis) as well as if the article is highly cited, or from a rigorous source			x	

1. Chen et al., (2022); see also Cognitive Computation Group (2020); 2. Bing/ (Bing, 2018)/Schwartz (2018); 3. Consensus (app); 4. Weltevrede & Borra, (2016, p. 4)**;** *Colour draws attention to contested parts of article; ^ Consensus meter; possible to filter results by 'yes', 'possibly', 'no'.
P = Priming; F = Framing; An = Anchoring; Av = Availability Bias

Across the items analysed, there are some commonalities in terms of design forms that address the four overarching objectives we identify. The design elements comprising these forms are similar in addressing the 'information overview' objective, through elements that give summaries and draw attention to differing perspectives. There are more differences in elements addressing the second objective, with two items (3 and 4) addressing features of disagreement or lack of consensus more directly, and one (2) indicating this concern through use of a 'v' (versus) suggesting juxtaposing

perspectives. These features have the benefit of priming for visibility of conflict, and may be less likely to provide a hierarchical anchor to users (where perspectives are presented horizontally, not vertically). The items share elements that associate sources to perspectives, with two (1, 3) providing elements that indicate the type of source; this may serve to 'make available' a feature demonstrating the possible usefulness of particular types of sources.

Discussion

Information conflict and uncertainty is commonplace in our current information environment, challenging for information seekers to navigate, yet important to represent and understand given its significance for decision making. Despite this, current common SERPs and interfaces grounded in KGs do not represent these features. Through our analysis we sought to investigate how these informational characteristics have been investigated in the context of user tools and their visual design features. To do this, we identified approaches drawn from multi-disciplinary context to inform the development of a visual scoping review method. We then synthesized the design elements in terms of principles that might support users in navigating uncertainty and disagreement. In order to develop recommendations that expand on these principles, it is first necessary to articulate how design heuristics – instantiated composites of elements that address particular principles – imply tensions that can be observed in the ways that particular design elements relate to or address biases. For example, the use of a 'v' or 'versus' indicator implies contest between ideas, where in many contexts of multiple perspectives, that might lead to either false balance (i.e., placing perspectives as 'equal' when they are not equally supported), or manufactured dissent (i.e., implying disagreement where much of the issue is agreed on). Two overarching tensions exist for interface design:

1. Challenge 1: Interfaces should be understandable to users, avoiding information overload, while simultaneously communicating uncertainty and disagreement at all points that they occur.
2. Challenge 2: Interfaces should aim to make visible the underlying data or model, to help users to understand why they are being shown the information they are; while visual features such as colour, or informational features such as data, may help with this, variation in the underlying data presents challenges around standardisation of language across claims

Table 6 elaborates on these tensions by aligning design elements observed (summarised Table 7) with the biases expressed by Novin and Meyers. In so doing, Table 8 extends their analysis of common search engines (such as Google) through drawing on our review of designs that specifically target uncertainty and disagreement.

Table 6: Tensions in Design Heuristics for Navigating Uncertainty and Disagreement in SERPs, indicating designs that are less or more likely to support user engagement with navigating uncertainty and disagreement

	Less adaptive	**More adaptive**
Priming: representation of perspectives and conflict embedded in design	Summarised overviews may reduce investigation of underlying perspectives. Familiar appearance/genre may be primed for.	Demarked results box around sets of perspectives. Consistent presentation for all perspectives, (e.g. short paragraph or bullet points; images for each result, to minimise risk of bias). Encourages users to engage in deeper reading.
Anchoring: Avoid perception of 'hierarchy' of sources or perspectives	Display of two sources, or 'vs' may imply false balance, manufactured conflict, or false consensus around 'two views'	Perspectives not artificially bifurcated, navigation between perspectives without use e.g. of back button to demonstrate range of perspectives while also indicating

		source features to avoid manufactured conflict.
Framing: Facilitate navigation of groups of sources by perspective and the connections between them	Present sources of different genres, qualities, and perspectives without indicating these features (thus framing as unimportant).	Make clear why results (or sources) appear near each other, and their perspective-relationship and quality; navigation between perspectives and sources, and visual indicators such as use of colour may assist.
Availability: Make navigation of different views or conflict easier	Require browsing of multiple pages, or lower ranked results for identification of uncertainty and conflict	Facilitate navigation between perspectives without 'back', arrange results to facilitate cross-navigation

Following from these tensions, we expand on the four key design principles for representing uncertainty and conflict in SERPs that we identified above:

1. Provide an informative overview of information returned, by bounding sets of perspectives (thus foregrounding their presence) and consistency in summarising of them (e.g., use of bullet points, images, or short paragraphs across them)
2. Make different perspectives clear, and make it easy to navigate these perspectives, by use of horizontal space to display perspectives with perspective sources displayed vertically, it should be possible to move between perspectives without requiring that they are artificially 'distanced' (e.g., through use of 'v', distance on the page, or a vertical hierarchy, or requirements to browse away and use 'back' to gain further insight)
3. Make the connection of perspectives to sources clear, without overwhelming users with redundant information, by providing the sources for perspectives with clear indication of which sources relate to which perspective(s)
4. Provide source features, by providing features that may be used to indicate quality or genre in particular contexts (e.g., date, source, media type, other markers such as 'highly cited' or related to research context, etc.).

Addressing these challenges and recommendations is important for future research. Moreover, further empirical work and evidence syntheses should seek to replicate these results using different methods, and build on them to develop novel systems and their evaluation.

Limitations

This work sought to investigate how SERPs and KGs can be designed to represent uncertainty and disagreement. A key limitation of the approach adopted is that by focusing on existing discussion and literature (including grey literature searches), the data – a relatively small set of results – is limited by what is present, rather than what is possible. This limitation is also expressed in the lack of well-established methods to search for, analyse, and distill or synthesise findings from visual artefacts that meet target criteria. There are likely to be other examples of systems – excluded here – where the issue of uncertainty and disagreement were present, but not a core topic of investigation. Certainly these may provide further examples, although given the importance of uncertainty and disagreement across decision making, further research should incorporate these issues as topics of investigation in their own right.

Conclusions

We should care about how to represent uncertainty and disagreement into our key information seeking tools. This representation matters for helping users to understand information, and navigate it appropriately across decision making contexts, including contexts where it is important users are not provided with a false sense of certainty or consensus. The interface design of such tools plays a key

role in this navigation, and yet there appear to be few methods available to systematically search for, analyse, and understand such interfaces.

Our visual scoping review method is innovative because its orientation to visual artefacts recognises the significance of material objects -SERP and KG interfaces in this case – to people's interaction with and through information. The method builds on established approaches, with this paper demonstrating how these approaches can be adapted to identify relevant material (including figures and their related text), and analyse this material drawing on frameworks in literature that is of relevance to the questions being asked (in this case a directed content analysis drawing on literature including Novin and Meyers, 2016). The approach is beneficial to researchers because the material things worked with in research (and beyond) provide rich insight for addressing questions of design, but they are not a core concern of traditional systematic review models.

Our analysis of the retrieved research indicates that although there is a small amount of research investigating uncertainty and disagreement in SERPs and KGs, design features are not typically the focus. Moreover, a limited number of results provide a visual representation that could inform future interface design or uptake into existing systems. We synthesise using our visual review method, and drawing on prior research regarding bias in SERP interfaces.

As our analysis demonstrates, this is compounded in this research by the relatively small number of results available. Through our analysis we have set out one approach to conducting a systematic visual scoping review, using it to demonstrate the relative lack of focus on SERP and KG design for uncertainty and disagreement. From relevant results identified, we provide a set of design principles that represent key design principles in representing uncertainty and disagreement in SERPs.

References

Alvarado, O., Htun, N.N., Jin, Y. and Verbert, K. (2022), "A Systematic Review of Interaction Design Strategies for Group Recommendation Systems", *Proceedings of the ACM on Human-Computer Interaction*, Vol. 6 No. CSCW2, p. 271:1-271:51, doi: 10.1145/3555161.

American Educational Research Association. (2006), "Standards for reporting on empirical social science research in AERA publications: American Educational Research Association", *Educational Researcher*, Vol. 35 No. 6, pp. 33–40, doi: 10.3102/0013189X035006033.

Bakker, A. (2019), "Design principles in design research: A commentary", in Bikner-Ahsbahs, A. and Peters, M. (Eds.), *Unterrichtsentwicklung macht Schule*, Springer Fachmedien Wiesbaden, Wiesbaden, pp. 177–192, doi: 10.1007/978-3-658-20487-7_10.

Bing. (2018), "Toward a More Intelligent Search: Bing Multi-Perspective Answers", 6 February, available at: https://blogs.bing.com/search-quality-insights/february-2018/Toward-a-More-Intelligent-Search-Bing-Multi-Perspective-Answers/ (accessed 21 July 2023).

van der Bles, A.M., van der Linden, S., Freeman, A.L.J., Mitchell, J., Galvao, A.B., Zaval, L. and Spiegelhalter, D.J. (2019), "Communicating uncertainty about facts, numbers and science", *Royal Society Open Science*, Royal Society, Vol. 6 No. 5, p. 181870, doi: 10.1098/rsos.181870.

Cano-Oron, L. (2019), "Dr. Google, what can you tell me about homeopathy? Comparative study of the top10 websites in the United States, United Kingdom, France, Mexico and Spain", *Profesional De La Informacion*, Ediciones Profesionales Informacion Sl-Epi, Barcelona, Vol. 28 No. 2, p. e280213, doi: 10.3145/epi.2019.mar.13.

Carroll, J.M. and Rosson, M.B. (1992), "Getting Around the Task-artifact Cycle: How to Make Claims and Design by Scenario", *ACM Transactions on Information Systems (TOIS)*, Vol. 10 No. 2, pp. 181–212, doi: 10.1145/146802.146834.

Center for History and New Media. (n.d.). "Zotero Quick Start Guide", available at: http://zotero.org/support/quick_start_guide.

Chen, S., Liu, S., Uyttendaele, X., Zhang, Y., Bruno, W. and Roth, D. (2022), "Design Challenges for a Multi-Perspective Search Engine", *Findings of the Association for Computational Linguistics: NAACL 2022*, presented at the Findings of the Association for Computational Linguistics: NAACL 2022, Association for Computational Linguistics, Seattle, United States, pp. 293–303, doi: 10.18653/v1/2022.findings-naacl.22.

Chen, Y., Zhao, Y. and Wang, Z. (2020), "Understanding online health information consumers' search as a learning process: Search platform sequences, query reformulation and conceptual changes", *Library Hi Tech*, Emerald Publishing Limited, Vol. 38 No. 4, pp. 859–881, doi: 10.1108/LHT-08-2019-0174.

Chowdhury, S., Gibb, F. and Landoni, M. (2014), "A model of uncertainty and its relation to information seeking and retrieval (IS&R)", *Journal of Documentation*, Emerald Group Publishing Limited, Vol. 70 No. 4, pp. 575–604, doi: 10.1108/JD-05-2013-0060.

Cobb, P., Confrey, J., Disessa, A., Lehrer, R., Schauble, L. and others. (2003), "Design experiments in educational research", *Educational Researcher*, Vol. 32 No. 1, pp. 9–13, doi: 10.3102/0013189X032001009.

Consensus. (2023), "Introducing: The Consensus Meter", *Consensus - Evidence-Based Answers, Faster*, 31 January, available at: https://consensus.app/home/blog/introducing-the-consensus-meter/ (accessed 21 July 2023).

Deroover, K., Knight, S., Burke, P.F. and Bucher, T. (2023), "Why do experts disagree? The development of a taxonomy", *Public Understanding of Science*, Vol. 32 No. 2, pp. 224–246, doi: 10.1177/09636625221110029.

Easterday, M., Rees Lewis, D. and Gerber, E. (2016), "The logic of the theoretical and practical products of design research", *Australasian Journal of Educational Technology*, Vol. 32, pp. 125–144, doi: 10.14742/ajet.2464.

"Ethical issues (consent etc.) | Report Sections | EQUATOR Network". (n.d.). , available at: https://www.equator-network.org/?post_type=eq_guidelines&eq_guidelines_study_design=0&eq_guidelines_clinical_specialty=0&eq_guidelines_report_section=ethical-issues-consent-etc&s= (accessed 10 April 2023).

Ferguson, L.E. (2014), "Epistemic Beliefs and Their Relation to Multiple-Text Comprehension: A Norwegian Program of Research", *Scandinavian Journal of Educational Research*, Vol. 59 No. 6, pp. 731–752, doi: 10.1080/00313831.2014.971863.

Gentles, S.J., Charles, C., Nicholas, D.B., Ploeg, J. and McKibbon, K.A. (2016), "Reviewing the research methods literature: principles and strategies illustrated by a systematic overview of sampling in qualitative research", *Systematic Reviews*, Vol. 5 No. 1, p. 172, doi: 10.1186/s13643-016-0343-0.

Haas, A. and Unkel, J. (2017), "Ranking versus reputation: perception and effects of search result credibility", *Behaviour & Information Technology*, Vol. 36 No. 12, pp. 1285–1298, doi: 10.1080/0144929X.2017.1381166.

Hearst, M.A. (2009), *Search User Interfaces*, 1st ed., Cambridge University Press.

Heersmink, R. and Knight, S. (2018), "Distributed learning: Educating and assessing extended cognitive systems", *Philosophical Psychology*, Vol. 31 No. 6, pp. 969–990, doi: 10.1080/09515089.2018.1469122.

Hevner, A.R., March, S.T., Park, J. and Ram, S. (2004), "Design Science in Information Systems Research", *MIS Quarterly*, Management Information Systems Research Center, University of Minnesota, Vol. 28 No. 1, pp. 75–105, doi: 10.2307/25148625.

Hogan, A., Blomqvist, E., Cochez, M., D'amato, C., Melo, G.D., Gutierrez, C., Kirrane, S., *et al.* (2021), "Knowledge Graphs", *ACM Computing Surveys*, Vol. 54 No. 4, p. 71:1-71:37, doi: 10.1145/3447772.

Hsieh, H.-F. and Shannon, S.E. (2005), "Three Approaches to Qualitative Content Analysis", *Qualitative Health Research*, Vol. 15 No. 9, pp. 1277–1288, doi: 10.1177/1049732305276687.

Hu, L. and Chen, G. (2021), "A systematic review of visual representations for analyzing collaborative discourse", *Educational Research Review*, Vol. 34, p. 100403, doi: 10.1016/j.edurev.2021.100403.

Huang, Y.-H. (2020), "Exploring Students' Search Behavior and the Effect of Epistemological Beliefs on Contradictory Issues", in Nah, F.F.-H. and Siau, K. (Eds.), *HCI in Business, Government and Organizations*, Springer International Publishing, Cham, pp. 59–68, doi: 10.1007/978-3-030-50341-3_5.

Isenberg, T., Isenberg, P., Chen, J., Sedlmair, M. and Möller, T. (2013), "A Systematic Review on the Practice of Evaluating Visualization", *IEEE Transactions on Visualization and Computer Graphics*, presented at the IEEE Transactions on Visualization and Computer Graphics, Vol. 19 No. 12, pp. 2818–2827, doi: 10.1109/TVCG.2013.126.

Kahrass, H., Borry, P., Gastmans, C., Ives, J., Graaf, R. van der, Strech, D. and Mertz, M. (2021), "PRISMA-Ethics – Reporting Guideline for Systematic Reviews on Ethics Literature: development, explanations and examples", OSF Preprints, 13 September, doi: 10.31219/osf.io/g5kfb.

Kammerer, Y., Amann, D.G. and Gerjets, P. (2015), "When adults without university education search the Internet for health information: The roles of Internet-specific epistemic beliefs and a source evaluation intervention", *Computers in Human Behavior*, Vol. 48, pp. 297–309, doi: 10.1016/j.chb.2015.01.045.

Kattenbeck, M., Link to external site, this link will open in a new window and Elsweiler, D. (2019), "Understanding credibility judgements for web search snippets", *Aslib Journal of Information Management*, Emerald Group Publishing Limited, Bradford, United Kingdom, Vol. 71 No. 3, pp. 368–391, doi: 10.1108/AJIM-07-2018-0181.

Knight, S. (2012), "Evaluating Google as an Epistemic Tool", *Society of the Query*, 4 December, available at: http://networkcultures.org/wpmu/re-search/2012/12/04/evaluating-google-as-an-epistemic-tool/ (accessed 10 January 2013).

Knight, S., Rienties, B., Littleton, K., Mitsui, M., Tempelaar, D.T. and Shah, C. (2017), "The relationship of (perceived) epistemic cognition to interaction with resources on the internet", *Computers in Human Behavior*, Vol. 73, pp. 507–518, doi: 10.1016/j.chb.2017.04.014.

Kobayashi, K. (2014), "Students' consideration of source information during the reading of multiple texts and its effect on intertextual conflict resolution", *Instructional Science*, Vol. 42 No. 2, pp. 183–205, doi: 10.1007/s11251-013-9276-3.

Ludolph, R., Allam, A. and Schulz, P.J. (2016), "Manipulating Google's Knowledge Graph Box to Counter Biased Information Processing During an Online Search on Vaccination: Application of a Technological Debiasing Strategy", *Journal of Medical Internet Research*, Jmir Publications, Inc, Toronto, Vol. 18 No. 6, p. e137, doi: 10.2196/jmir.5430.

Marchionini, G. (2006), "Exploratory search: from finding to understanding", *Communications of the ACM*, Vol. 49 No. 4, p. 41, doi: 10.1145/1121949.1121979.

McCrickard, D.S. (2012), *Making Claims: The Claim as a Knowledge Design, Capture, and Sharing Tool in HCI*, Morgan & Claypool Publishers.

McKenney, S. and Reeves, T.C. (2013), *Conducting Educational Design Research*, Routledge.

Moran, T.P. and Carroll, J.M. (1996), *Design Rationale: Concepts, Techniques, and Use*, L. Erlbaum Associates Inc.

Nagler, R.H. (2014a), "Adverse Outcomes Associated With Media Exposure to Contradictory Nutrition Messages", *Journal of Health Communication*, Vol. 19 No. 1, pp. 24–40, doi: 10.1080/10810730.2013.798384.

Nagler, R.H. (2014b), "Adverse Outcomes Associated With Media Exposure to Contradictory Nutrition Messages", *Journal of Health Communication*, Vol. 19 No. 1, pp. 24–40, doi: 10.1080/10810730.2013.798384.

Nagler, R.H., Vogel, R.I., Gollust, S.E., Rothman, A.J., Fowler, E.F. and Yzer, M.C. (2020a), "Public perceptions of conflicting information surrounding COVID-19: Results from a nationally representative survey of U.S. adults", *PLOS ONE*, Public Library of Science, Vol. 15 No. 10, p. e0240776, doi: 10.1371/journal.pone.0240776.

Nagler, R.H., Vogel, R.I., Gollust, S.E., Rothman, A.J., Fowler, E.F. and Yzer, M.C. (2020b), "Public perceptions of conflicting information surrounding COVID-19: Results from a nationally

representative survey of U.S. adults", *PLOS ONE*, Public Library of Science, Vol. 15 No. 10, p. e0240776, doi: 10.1371/journal.pone.0240776.

Navigating Information Pollution in the Time of COVID-19: A New Platform. (2020), .

Novin, A. and Meyers, E. (2016), "Controversial Search Engine Results: An Exploratory Study of Information Presentation and Use", *Proceedings of the Annual Conference of CAIS / Actes Du Congrès Annuel de l'ACSI*, doi: 10.29173/cais957.

Novin, A. and Meyers, E. (2017), "Making Sense of Conflicting Science Information: Exploring Bias in the Search Engine Result Page", *Chiir'17: Proceedings of the 2017 Conference Human Information Interaction and Retrieval*, Assoc Computing Machinery, New York, pp. 175–184, doi: 10.1145/3020165.3020185.

Oppl, S. (2022), "On the Ambiguous Nature of Theory in Educational Design-based Research – Reflecting and Structuring from an IS Perspective", *EDeR. Educational Design Research*, Vol. 6 No. 1, doi: 10.15460/eder.6.1.1808.

Rains, S.A. (2014), "Health Information Seeking and the World Wide Web: An Uncertainty Management Perspective", *Journal of Health Communication*, Vol. 19 No. 11, pp. 1296–1307, doi: 10.1080/10810730.2013.872731.

Ralph, P., Ali, N. bin, Baltes, S., Bianculli, D., Diaz, J., Dittrich, Y., Ernst, N., *et al.* (2021), "Empirical Standards for Software Engineering Research", *arXiv:2010.03525 [Cs]*.

Riddell, J., Teal, G., Flowers, P., Boydell, N., Coia, N. and McDaid, L. (2022), "Mass media and communication interventions to increase HIV testing among gay and other men who have sex with men: Social marketing and visual design component analysis", *Health*, SAGE Publications Ltd, Vol. 26 No. 3, pp. 338–360, doi: 10.1177/1363459320954237.

Salmerón, L., Kammerer, Y. and García-Carrión, P. (2013), "Searching the Web for conflicting topics: Page and user factors", *Computers in Human Behavior*, Vol. 29 No. 6, pp. 2161–2171, doi: 10.1016/j.chb.2013.04.034.

Scarton, L.A., Del Fiol, G., Oakley-Girvan, I., Gibson, B., Logan, R. and Workman, T.E. (2018), "Understanding cancer survivors' information needs and information-seeking behaviors for complementary and alternative medicine from short-to long-term survival: a mixed-methods study", *Journal of the Medical Library Association*, Medical Library Assoc, Chicago, Vol. 106 No. 1, pp. 87–97, doi: 10.5195/jmla.2018.200.

Schwartz, B. (2018), "Bing explains how AI-powered intelligent answers can show users two points of view for the same query", *Search Engine Land*, 6 February, available at: https://searchengineland.com/bing-adds-multi-perspective-intelligent-answers-search-results-291369 (accessed 17 October 2022).

Shah, C. (2020), "People can't identify COVID-19 fake news – I Seek, Therefore I Am", available at: https://infoseeking.org/iblog/2020/10/23/people-cant-identify-covid-19-fake-news/ (accessed 23 November 2020).

Song, Q. and Jiang, J. (2022), "How Misinformation Density Affects Health Information Search", *Proceedings of the ACM Web Conference 2022*, presented at the WWW '22: The ACM Web Conference 2022, ACM, Virtual Event, Lyon France, pp. 2668–2677, doi: 10.1145/3485447.3512141.

Stefanidi, E., Bentvelzen, M., Woźniak, P.W., Kosch, T., Woźniak, M.P., Mildner, T., Schneegass, S., *et al.* (2023), "Literature Reviews in HCI: A Review of Reviews", *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, presented at the Conference on Human Factors in Computing Systems - Proceedings, Association for Computing Machinery, New York, NY, USA, pp. 1–24, doi: 10.1145/3544548.3581332.

Tavani, H. (2020), "Search Engines and Ethics", in Zalta, E.N. (Ed.), *The Stanford Encyclopedia of Philosophy*, Fall 2020., Metaphysics Research Lab, Stanford University.

Teevan, J., Collins-Thompson, K., White, R.W. and Dumais, S. (2014), "Slow search", *Communications of the ACM*, Vol. 57 No. 8, pp. 36–38, doi: 10.1145/2633041.

The Design-Based Research Collective. (2003), "Design-Based Research: An Emerging Paradigm for Educational Inquiry", *Educational Researcher*, Vol. 32 No. 1, pp. 5–8, doi: 10.3102/0013189X032001005.

Vasantavada, P.V., Sanderson, R., Ells, L. and Zohoori, F.V. (2022), "Web search engines reveal conflicting information about water fluoridation", *British Dental Journal*, Nature Publishing Group, pp. 1–5, doi: 10.1038/s41415-022-3929-z.

Verbeke, M., Berendt, B. and Nijssen, S. (2009), "Data mining, interactive semantic structuring, and collaboration: A diversity-aware method for sense-making in search", *Proceedings of First International Workshop on Living Web, Collocated with the 8th International Semantic Web Conference (ISWC-2009). CEUR Workshop Proceedings, Washington, DC, USA*, Vol. 515.

Vial, S., Boudhraâ, S. and Dumont, M. (2022), "Human-Centered Design Approaches in Digital Mental Health Interventions: Exploratory Mapping Review", *JMIR Mental Health*, Vol. 9 No. 6, p. e35591, doi: 10.2196/35591.

Weltevrede, E. and Borra, E. (2016), "Platform affordances and data practices: The value of dispute on Wikipedia", *Big Data & Society*, Vol. 3 No. 1, p. 2053951716653418, doi: 10.1177/2053951716653418.

Wharton, T. (2017), "Rigor, Transparency, and Reporting Social Science Research: Why Guidelines Don't Have to Kill Your Story", *Research on Social Work Practice*, Vol. 27 No. 4, pp. 487–493, doi: 10.1177/1049731515622264.

Wilson, E.V., Djamasbi, S., Strong, D. and Ruiz, C. (2017), "Using a Key Informant Focus Group, Formative User Testing, and Theory to Guide Design of a Sleep Health BCSS", presented at the Hawaii International Conference on System Sciences, doi: 10.24251/HICSS.2017.404.

Wright, M.C. (2023), "Information visualization and integration", *Clinical Decision Support and beyond: Progress and Opportunities in Knowledge-Enhanced Health and Healthcare*, pp. 435–463, doi: 10.1016/B978-0-323-91200-6.00032-2.

Wright, M.C., Borbolla, D., Waller, R.G., Del Fiol, G., Reese, T., Nesbitt, P. and Segall, N. (2019), "Critical care information display approaches and design frameworks: A systematic review and meta-analysis", *Journal of Biomedical Informatics*, Vol. 100, p. 100041, doi: 10.1016/j.yjbinx.2019.100041.

Yilmaz, S., Daly, S.R., Seifert, C.M. and Gonzalez, R. (2016), "Evidence-based design heuristics for idea generation", *Design Studies*, Vol. 46, pp. 95–124, doi: 10.1016/j.destud.2016.05.001.

Acknowledgements

[blinded for reiview]

Supplement: Full Search Strategy

All searches were conducted in October 2022.

ACM searches

NOTE: ACM search interface converts a string (the first below) into a query (the second).

Abstract:(("knowledge graph" OR SERP OR "search engine results page*" OR ("information seeking behaviour" AND ("online" OR "digital" OR "web"))) AND (uncertain* OR conflict* OR disagree* OR "decision support" OR "decision aid" OR "evaluation of search results")) OR Title:(("knowledge graph" OR SERP OR "search engine results page*" OR ("information seeking behaviour" AND ("online" OR "digital" OR "web"))) AND (uncertain* OR conflict* OR disagree* OR "decision support" OR "decision aid" OR "evaluation of search results")) OR Keyword:(("knowledge graph" OR SERP OR "search engine results page*" OR ("information seeking behaviour" AND ("online" OR

"digital" OR "web"))) AND (uncertain* OR conflict* OR disagree* OR "decision support" OR "decision aid" OR “evaluation of search results”))

42 Results for: [[[**Abstract**: "knowledge graph"] **OR** [**Abstract**: serp] **OR** [**Abstract**: "search engine results page"] **OR** [[**Abstract**: "information seeking behaviour"] **AND** [[**Abstract**: "online"] **OR** [**Abstract**: "digital"] **OR** [**Abstract**: "web"]]]] **AND** [[**Abstract**: uncertain*] **OR** [**Abstract**: conflict*] **OR** [**Abstract**: disagree*] **OR** [**Abstract**: "decision support"] **OR** [**Abstract**: "decision aid"]]] **OR** [[[**Title**: "knowledge graph"] **OR** [**Title**: serp] **OR** [**Title**: "search engine results page"] **OR** [[**Title**: "information seeking behaviour"] **AND** [[**Title**: "online"] **OR** [**Title**: "digital"] **OR** [**Title**: "web"]]]] **AND** [[**Title**: uncertain*] **OR** [**Title**: conflict*] **OR** [**Title**: disagree*] **OR** [**Title**: "decision support"] **OR** [**Title**: "decision aid"]]] **OR** [[[**Keywords**: "knowledge graph"] **OR** [**Keywords**: serp] **OR** [**Keywords**: "search engine results page"] **OR** [[**Keywords**: "information seeking behaviour"] **AND** [[**Keywords**: "online"] **OR** [**Keywords**: "digital"] **OR** [**Keywords**: "web"]]]] **AND** [[**Keywords**: uncertain*] **OR** [**Keywords**: conflict*] **OR** [**Keywords**: disagree*] **OR** [**Keywords**: "decision support"] **OR** [**Keywords**: "decision aid"]]]

NOTE: A limitation of the ACM query constructor is that it does not have a unified field that represents Title/Abstract/Keyword, thus searches must be manually defined to search these fields. A limitation of the natural approach to using this construction logic is that it requires the intersection of terms to appear within any one of the Title/Abstract/Keyword, but does not capture instances where one term appears in, say, the Title, while another is present in the Abstract. Following our analysis we checked the impact of this, at this later date 47 results were returned by the query above, with 59 results the query below. On screening, none appeared to extend the scope of work conducted, and thus due to resource constraints these results were not further analysed.

Abstract:("knowledge graph" OR SERP OR "search engine results page" OR ("information seeking behaviour" AND ("online" OR "digital" OR "web"))) AND Abstract:(uncertain* OR conflict* OR disagree* OR "evaluation of search results" OR "decision support" OR "decision aid") OR Title:("knowledge graph" OR SERP OR "search engine results page" OR ("information seeking behaviour" AND ("online" OR "digital" OR "web"))) AND Title:(uncertain* OR conflict* OR disagree* OR "evaluation of search results" OR "decision support" OR "decision aid") OR Keyword:("knowledge graph" OR SERP OR "search engine results page" OR ("information seeking behaviour" AND ("online" OR "digital" OR "web"))) AND Keyword:(uncertain* OR conflict* OR disagree* OR "evaluation of search results" OR "decision support" OR "decision aid") OR Abstract:("knowledge graph" OR SERP OR "search engine results page" OR ("information seeking behaviour" AND ("online" OR "digital" OR "web"))) AND Title:(uncertain* OR conflict* OR disagree* OR "evaluation of search results" OR "decision support" OR "decision aid") OR Abstract:("knowledge graph" OR SERP OR "search engine results page" OR ("information seeking behaviour" AND ("online" OR "digital" OR "web"))) AND Keyword:(uncertain* OR conflict* OR disagree* OR "evaluation of search results" OR "decision support" OR "decision aid") OR Title:("knowledge graph" OR SERP OR "search engine results page" OR ("information seeking behaviour" AND ("online" OR "digital" OR "web"))) AND Abstract:(uncertain* OR conflict* OR disagree* OR "evaluation of search results" OR "decision support" OR "decision aid") OR Title:("knowledge graph" OR SERP OR "search engine results page" OR ("information seeking behaviour" AND ("online" OR "digital" OR "web"))) AND Keyword:(uncertain* OR conflict* OR disagree* OR "evaluation of search results" OR "decision support" OR "decision aid") OR Keyword:("knowledge graph" OR SERP OR "search engine results page" OR ("information seeking behaviour" AND ("online" OR "digital" OR "web"))) AND Abstract:(uncertain* OR conflict* OR disagree* OR "evaluation of search results" OR "decision support" OR "decision aid") OR Keyword:("knowledge graph" OR SERP OR "search engine results page" OR ("information seeking behaviour" AND ("online" OR "digital" OR "web"))) AND Title:(uncertain* OR conflict* OR disagree* OR "evaluation of search results" OR "decision support" OR "decision aid")

Web of Science searches

NOTE: Consistent with the queries constructed for the ACM Digital Library, Web of Science Core Collection searches used an 'intra-field' Boolean logic, with terms required within (not across) Title (TI), Abstract (AB), and Keywords (AK).

AB=(("knowledge graph" OR SERP OR "search engine results page*" OR ("information seeking behaviour" AND ("online" OR "digital" OR "web"))) AND (uncertain* OR conflict* OR disagree* OR "decision support" OR "decision aid" OR "evaluation of search results")) OR TI=(("knowledge graph" OR SERP OR "search engine results page*" OR ("information seeking behaviour" AND ("online" OR "digital" OR "web"))) AND (uncertain* OR conflict* OR disagree* OR "decision support" OR "decision aid" OR "evaluation of search results")) OR AK=(("knowledge graph" OR SERP OR "search engine results page*" OR ("information seeking behaviour" AND ("online" OR "digital" OR "web"))) AND (uncertain* OR conflict* OR disagree* OR "decision support" OR "decision aid" OR "evaluation of search results"))

These searches (mirroring those above for ACM) resulted in 216 results. A subsequent check (with limits to the exclude new pieces) indicates 272 items returned for the alternative search strategy using the "Topic" field (TS), or 56 additional items. A review of the titles and abstracts of these items indicates no additional relevant sources.

TS=(("knowledge graph" OR SERP OR "search engine results page*" OR ("information seeking behaviour" AND ("online" OR "digital" OR "web"))) AND (uncertain* OR conflict* OR disagree* OR "decision support" OR "decision aid" OR "evaluation of search results"))

Based on identification of a small number of tools targeting multi-perspective search we conducted supplementary searches (including forward and backward citation chasing), including to check whether using these terms in place of those indicating conflict and uncertainty might return further results, i.e., multi-perspective where our previous terms were not mentioned:
TS=(("multi perspective" OR "multiple perspective") AND ("knowledge graph" OR SERP OR "search engine results page*" OR ("information seeking behaviour" AND ("online" OR "digital" OR "web")))) NOT TS=(("knowledge graph" OR SERP OR "search engine results page*" OR ("information seeking behaviour" AND ("online" OR "digital" OR "web"))) AND (uncertain* OR conflict* OR disagree* OR "decision support" OR "decision aid" OR "evaluation of search results"))

6 results were returned, on screening titles and abstracts non were relevant and thus this strategy was not pursued.

Further searches including grey literature

Known sources were included (e.g., Bing's perspective view), with a review of material citing key works (e.g., Marchionini, 2006), and using the search terms identified above, using both google scholar and google search (excluding scholarly publisher sites). Google searches are limited both by the underlying index (which changes, and is impacted by non-informationally relevant optimization features), and query syntax (e.g., there is no full Boolean syntax, wildcards at the end of terms such as wildcard* - to capture suffixes – should have no impact, because by default non-exact searches include related words, nevertheless this syntax did increase results, etc.). In addition, the number of results flagged as returned by Google (at the top of the SERP, phrased "About 139,000 results (0.40 seconds)"), does not reflect the number displayed, for example for the query generating those 139,000 results, viewing the final page of results provided 91. Between the time of searching and revisions, Google (in our area at least) removed the possibility to paginate pages, this has the effect of making it harder to directly move to the final results, although the single expandable page of results is easier to save as a single file of results. Google searches were not recorded systematically, but included searches for the keywords indicated above, excluding publisher pages, for example:

"diversity aware search" -filetype:pdf -site:researchgate.net -site:semanticscholar -site:springer -site:emerald (returned 91 results, or 2,450 reported by Google)

("search engine" "results page" OR SERP OR "knowledge graph") AND uncertain OR conflicting OR disagreement OR uncertainty -site:sciencedirect.com -site:mdpi.com -site:researchgate.net -site:eprints -site:link.springer.com -site:frontiersin.org -site:hindawi.com -site:dl.acm.org -site:journals.sagepub.com -site:ncbi.nim.nih.gov -site:semantic-web-journal.net -site:direct.mit.edu -site:ec-3.org -site:ncbi.nim.nih.gov -site:*.mitpress.mit.edu -site:ojs.* -site:springerprofessional.de* (returned 264 results, or 550,000 reported by Google).

[i] Without this operator, 27,418 results are returned, or 2,828 when limited to document type 'review' (LIMIT-TO(DOCTYPE, "re")). Using W/20 returns 196. Our rationale for selection of W/10 is that it allows for the terms to span two sentences, intending to capture discussion of the [object] [review method] relationship as a core component of the paper.