Discourse Graphs for Augmented Knowledge Synthesis: What and Why

Joel Chan, University of Maryland College of Information Studies
Email: joelchan@umd.edu | Twitter: @joelchan86

Last updated: August 6, 2021

This document describes the idea of discourse graphs — which builds on a long history of robust information modeling work — and its possible applications for augmenting individual and collective synthesis.

1 Motivation: Effective synthesis is critical but (unnecessarily) hard

To advance science, scientists must *synthesize* what is currently known and unknown about scientific problems. Effective synthesis generates new knowledge, integrating relevant theories, concepts, claims, and evidence into novel conceptual wholes [Strike and Posner, 1983, Blake and Pratt, 2006]. Synthesis may be supported by and manifested in a variety of forms, such as a theory, an effective systematic or integrative literature review, a causal model, a cogent research proposal or problem formulation, or model of a design space, among others. The advanced understanding from synthesis can be a powerful force multiplier for choosing effective studies and operationalizations [van Rooij and Baggio, 2021, McElreath and Smaldino, 2015, Scheel et al., 2020], and may be especially necessary for problems where it is difficult or impossible to construct decisive experimental tests (e.g., the issue of mask efficacy for reducing community transmission [Howard et al., 2020]); indeed, scientific progress may not even be tractable without adequate synthesis (as theory), even with advanced methods and data [Jonas and Kording, 2017]: as Allen Newell famously said, "You can't play twenty questions with nature and win" [Newell, 1973]. To illustrate the power of synthesis for accelerating scientific progress, consider the example of Esther Duflo, who attributed her Nobel-Prize-winning work to the detailed synthesis of problems in developmental economics she obtained from a handbook chapter [Duflo, 2011].

Unfortunately, effective synthesis is rare. Published synthesis outputs, such as review papers and systematic reviews, are scarce, and almost never updated [Shojania et al., 2007, Petrosino, 1999]. Studies of literature reviews in doctoral dissertations [Lovitts, 2007, Holbrook et al., 2004, Boote and Beile, 2005] and even published papers [Alton-Lee, 1998, Alvesson and Sandberg, 2011, van Rooij and Baggio, 2021, McPhetres et al., 2020, Bhurke et al., 2015, Fleming et al., 2013] have found them frequently lacking key aspects of synthesis quality, such as critical engagement with and generative integration of prior work and theory.

There is an important yet relatively neglected reason for this: **the fundamental information models that underlie most scientists' everyday reading and communication practices are not readily amenable to integration, comparison, sharing, and translation across publications, researchers, or domains.** The experience of synthesis work is often described as arduous and effortful [Ervin, 2008, Knight et al., 2019, Granello, 2001], and estimates of the time taken to do synthesis in a rigorous manner, such as in a systematic review, corroborate these subjective experiences [Shojania et al., 2007, Petrosino, 1999, Ervin, 2008], with the labor of transforming the "raw data" of unstructured texts into forms amenable for analysis comprising a major portion of these time costs. One effort to address the difficulty of synthesis is a growing body of work on tools for augmenting systematic review work [O'Connor et al., 2019]. While promising, these efforts are often framed as special-purpose tools disconnected from (and not interoperable with) routine scientific practices [O'Connor et al., 2019].

An interesting line of evidence for the inadequacy of current document-centric information models — or as Qian et al [Qian et al., 2019] call it, "iTunes for papers" — is the desire path of scientists adopting niche tools with different information models [Chan et al., 2020]. For example, there is a subculture of academic researchers who repurpose qualitative data analysis tools like NVivo and Atlas.ti to do literature reviews [Wolfswinkel et al., 2013, Silver, 2020, anujacabraal, 2012]; it is notable that the key affordances of these tools emphasize interacting with different core, more granular information unites — excerpts and themes — than papers. There is also some adoption of niche specialized tools for literature sensemaking, such as LiquidText and Citavi, both of which emphasize the composition of networks of claims that are directly linked to contextualizing excerpts from documents.

2 Discourse Graphs: A promising model for augmenting synthesis

Across scientific disciplines, the past decades have witnessed sweeping efforts to rethink existing formats for scholarly communication, resulting in an array of related approaches — including ontologies, semantically rich data models, and other metadata and linked data standards — to support new modes of knowledge representation, sharing, and transfer [Renear and Palmer, 2009, Kuhn and Dumontier, 2017, de Waard, 2010]. These approaches take a profusion of forms to suit the functional requirements of different research and disciplinary contexts. Of primary interest here is a suite of information models [Ciccarese et al., 2008, Clark et al., 2014, Brush et al., 2016, Shum et al., 2006, **?**, Groth et al., 2010, McCrickard, 2012, de Waard et al., 2009] that share a common underlying model for representing scientific discourse: one that distills traditional forms of publication down into more granular, formalized knowledge *claims*, linked to supporting evidence and *context* through a network or *graph* model. Here, we use the term "**discourse graph**" to refer to this information model, to evoke the core concepts of representing and relating knowledge claims (rather than concepts) as the central unit, and emphasizing linking and relating these claims (rather than categorizing or filing them). Standardizing the representation of scientific claims and evidence in a graph model can support machine reasoning [Kuhn and Dumontier, 2017], but is also widely hypothesized to support human learning across domains and contexts [Shum et al., 2000, Renear and Palmer, 2009, de Waard, 2010, Clark et al., 2014].

To understand why this information model might augment synthesis, consider a researcher who wants to understand what interventions might be most promising for mitigating online harassment. To synthesize a formulation of this complex interdisciplinary problem that can advance the state of the art, she needs material that can help her work through detailed answers to a range of questions. For example, which theories have the most empirical support in this particular setting? Are there conflicting theoretical predictions that might signal fruitful areas of inquiry? What are the key phenomena to keep in mind when designing an intervention (e.g., perceptions of human vs. automated action, procedural considerations, noise in judgments of wrongdoing, scale considerations for spread of harm)? What intervention patterns (whether technical, behavioral, or institutional) have been proposed that are both a) judged on theoretical and circumstantial grounds as likely to be effective in this setting, and b) lacking in direct evidence for efficacy?

The answers to these questions cannot be found simply in the titles of research papers, in groupings of papers by area, or even in citation or authorship networks. The answers lie at lower levels of granularity: the level of theoretical and empirical **claims** or statements made within publications. For example, "viewers in a Twitch chat engaged in less bad behaviors after a user was banned by a moderator for bad behavior" [Seering et al., 2017], and "banning bad actors from a subreddit in 2015 was somewhat effective at mitigating spread of hate speech on other subreddits" [Chandrasekharan et al., 2017] are claims that interrelate in complex ways, both supporting other claims/theories that are in tension with each other. This level of granularity is crucial not just for finding relevant claims to inform the synthesis, but also for constructing more complex arguments and theories, by connecting statements in logical and discursive relationships. Beyond operating at the claim level, our researcher will also need to work through a range of **contextual details**. For example, to judge which studies, findings, or theories are most applicable to her setting, she needs to know key methodological details including the comparability of different studies' interventions, settings, populations, and outcome measures. She might need to reason over the fact that two studies that concluded limited

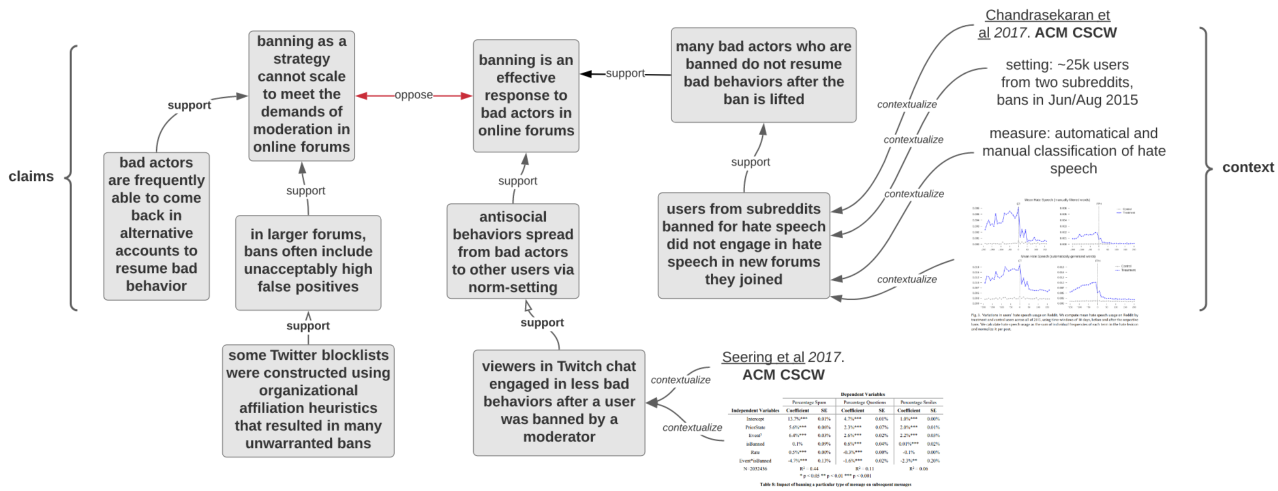

Figure 1: Example discourse graph (with claims and associated context) for theories and findings on effects of bans on bad actors in online forums.

efficacy of bans had ban interventions that were quite short, on a forum with no identity verification. Or she might reason through the fact that a prominent theory of bad faith and discourse was proposed by a philosopher from the early 2000's (before the rise of modern social media). To judge the validity of past findings (e.g., what has been established with sufficient certainty, where the frontier might be), she would need to know, for example, which findings came from which measures (e.g., self-report, behavioral measures), and the extent to which findings have been replicated cross authors from different labs, and across a variety of settings (e.g., year, platform, scale).

2.1 Hypothesized individual benefits: Creative synthesis and exploration

A discourse graph has key **affordances that are hypothesized to enable just these sorts of synthesis operations**. Information is represented primarily at the claim or statement level, and embedded in a graph of relationships with other claims and context. In a discourse graph, claims have many-to-many relationships to support composition of more complex arguments and theories, or "decompression" into component supporting/opposing claims. Contextual entities and information, such as methodological details and metadata, are explicitly included in the discourse graph. This supports direct analysis of claims with their evidentiary context, supporting critical engagement, integration, and even reinterpretation of individual findings. Figure 1 shows how this might be supported in the specific worked example above. Note that discourse graphs need not be represented or manipulated in this visual format; the underlying graph model can be instantiated in a variety of media, such as hypertext notebooks, and also implicitly in various analog implementations that allow for cross-referencing. What is important is the information architecture of representing networks of claims and their context.

Beyond the theoretical match between the kinds of queries scientists need to run over their evidence collection for synthesis, a discourse-centric representation that encodes granular claims instead of document "buckets" could facilitate exploration and conceptual combination. There is theoretical precedent for this in research on expertise and creative problem solving, where breaking complex ideas down into more meaningful, smaller conceptual "chunks" may be necessary for creative recombination into new conceptual wholes [McCrickard et al., 2013, Chase and Simon, 1973, Knoblich et al., 1999, McCaffrey, 2012]. Removing contextual details (though not losing the ability to recover them) may also be necessary and useful for synthesizing ideas and reusing knowledge across boundaries [Star and Griesemer, 1989, McMahan and Evans, 2018]. At the same time, constructive and creative engagement with contextual details, is thought to be necessary for developing novel conceptual wholes from "data", such as in sensemaking [Russell et al., 1993], systematic reviews [Blake and Pratt, 2006], or formal theory development [Marder, 2020, van Rooij and Baggio, 2021, Goldstein, 2018, Gruber and Barrett, 1974]. Further, accurately predicting just which contextual details are necessary to represent directly in an information object is a difficult task [Ackerman and Halverson, 2004, Lutters and Ackerman, 2007] that may be functionally impossible in creative settings. The conjunction of these affordances — having both granular information objects like claims, and the ability to progressively expand their context by traversing/retrieving from the discourse graphs — may help to resolve this tension between granularity and contextualizability. For instance, graph-model and hypertext affordances like hyperlinking or transclusion might enable scientists to hide "extraneous details" (to facilitate compression) without destructively blocking future reusers from obtaining necessary contextual details for reuse [Ackerman and Halverson, 2004, Lutters and Ackerman, 2007].

2.2 Hypothesized collective benefits: Reduced overhead, and enhanced creative reuse

Discourse graphs (or parts thereof) could also significantly reduce the overhead to synthesis through reuse and repurposing over time, across projects, and potentially even across people. For example, imagine collaborators sharing discourse graphs with each other, rather than simple documents full of unstructured notes, to speed up the process of working towards shared mental models and identifying productive areas of divergence; or a lab onboarding new researchers not with long reading lists, but with discourse graph subsets they can build on over time. How much effort could be reduced if this were a reality?

The same affordances of discourse graphs around granularity and contextualizability that are hypothesized to augment individual synthesis should also facilitate exploration and reuse of an evidence collection that was created by someone else, or by oneself in the past. For example, granular representation of scientific ideas at the claim level is a much better theoretical match for the kinds of queries that scientists want to ask of an evidence collection during synthesis [Clark et al., 2014, de Waard, 2010, Hars, 2001, Shum et al., 2000, de Ribaupierre and Falquet, 2017]. These claims may also be more to the level of processing required to be understood and reused by others, compared to raw annotations and marginalia [Marshall and Brush, 2004]. Also, ambiguity around concepts can be a significant barrier to reuse across knowledge boundaries. For example, keyword search is only really useful when there is a stable, shared understanding of ontology [Talja and Maula, 2003]: this condition is almost certainly not present when crossing knowledge boundaries [McMahan and Evans, 2018], and perhaps not even within fields of study with significant ongoing controversy amongst different schools of thought [Hjørland, 2002] In these settings, judging that two things are "the same" is problematic and difficult task; doing so without engagement with context can sometimes introduce *more* destructive ambiguity, not less, a hard-won lesson from the history of Semantic Web [Hayes and Halpin, 2008, Halford et al., 2013], ontology [Gustafson, 2020, Randall et al., 2011] and classification efforts [Bowker and Star, 2000]. A discourse-centric graph that embeds concepts in discourse contexts, traversing through networks of contextual details (such as authors, measures, contexts), and perhaps augmented by formal concepts as hooks, may be a better match for exploring ideas across knowledge boundaries. Further, although in many instances of knowledge reuse, contextual details tend to vary substantially across reuse tasks [Ackerman and Halverson, 2004, Ackerman et al., 2013], there might be sufficient overlap of useful contextual details (e.g., participant information, study context) that remain stable across reuse tasks [Blake and Pratt, 2006].

3 Conclusion: A call for further research

In this document I've described the importance and difficulty of synthesis for scientific progress, diagnosed an information-model barrier to doing synthesis, and described **discourse graphs** as an alternative information model that could augment and accelerate synthesis, both individually and collectively.

I'll close here with a word of caution and a call for further research: despite the significant hypothesized benefits of discourse graphs, we don't yet know very much about whether or how they work in scientific practice. Their efficacy in facilitating synthesis or increasing the speed of research advancement remains uncertain.

There have been attempts to integrate discourse graphs into scientific communities of practice: for example, the ScholOnto model [Shum et al., 2006] was supporting remote PhD mentoring and distributed collaborations; SWAN [Ciccarese et al., 2008] was integrated into the successful Alzforum online research community for Alzheimers research [Clark and Kinoshita, 2007]; and the micropublication model was integrated into the Domeo scientist network [Clark et al., 2014]. However, for a variety of reasons, the impact of these deployments has not been empirically evaluated. This may partly be due to changes in funding and infrastructure for these research software, leading to deprecation of technical infrastructure. Other efforts might not have made it past the experimental prototype stage for similar reasons (lack of funding, incentives). However, the pivot of some information models into more educational or general-purpose applications [Liddo et al., 2012] is suggestive of open problems that might stand in the way of realizing the benefits of discourse graphs in scientific practice more generally.

One potentially promising ongoing effort is the Clinical Interpretation of Variants in Cancer (CIViC) project [Griffith et al., 2017], which has successfully recruited hundreds of volunteer curators to curate evidence on cancer-related mutations, using the Scientific Evidence and Provenance Information Ontology (SEPIO) model [Brush et al., 2016]. There are no empirical evaluations of this project's impact, but a recent qualitative study of clinicians' perceptions of the resource uncovered challenges around mismatches in contextual details that were captured, and the clinically relevant information they needed to use it as a "knowledge base" [Cambrosio et al., 2020]. Another relevant study by Ribaupierre et al [de Ribaupierre and Falquet, 2017] found that a small sample of scientists who tested a prototype tool that allowed them to search over a literature by rhetorical elements (e.g., findings, methods, definitions) self-reported higher signal-to-noise ratio in results when searching for specific findings (e.g., "show all findings of studies that have addressed the issue of gender equality in terms of salary"), compared to using a standard keyword search interface. However, the study did not continue to observe *usage* of literature in a synthesis task.

On the practice side, surveys of tool usage by scientists suggest that document-centric workflows continue to dominate. For example, Bosman et al. [Bosman and Kramer, 2016] reported from a large-scale online survey of approximately 20k researchers worldwide that reference management tools like EndNote, Mendeley, and Zotero were the most frequently mentioned tools for managing and using literature. These large-scale findings are corroborated by more in-depth qualitative investigations of researcher practices, which generally find the predominance of these document-centric tools, as well as mainstream general-purpose software like Microsoft Word and Excel for note-taking [Qian et al., 2020, Willis et al., 2014, Hoang and Schneider, 2018].

Given the high potential of discourse graphs for augmenting synthesis, and the centrality of synthesis for scientific progress, research that directly tests their efficacy and explores how to integrate them into scientific workflows would be very valuable. The time is now ripe for exploring this, because a new generation of consumer-grade software has exploded on the scene under the general rubric of "networked notebooks". These platforms — some particularly popular ones includes RoamResearch[1], Obsidian[2], Logseq[3], Dendron[4], RemNote[5], and AthensResearch[6] — have democratized access to extensible, hypertext notetaking environments, drawing on roots in wiki technology and hacker-culture tools like GNU Emacs' org-mode extension[7]. These platforms have attracted a substantial user base, on the order of tens of thousands of users who have

[1]https://roamresearch.com/

[2]https://obsidian.md/

[3]https://logseq.github.io/

[4]dendron.so

[5]https://www.remnote.io/

[6]https://www.athensresearch.org/

[7]https://orgmode.org/

adopted these tools for knowledge synthesis. Central to the community's interaction is the regular development, sharing, and testing of plugins and extensions for the software, as well as for open-source reference managers like Zotero, editors like LaTeX, and reading and annotation software like Hypothes.is[8], Readwise[9], and Memex[10]. There is also support for sharing graphs with each other through automatic publishing of subsets of notes to a personal website (see, e.g., RoamGarden[11] and Obsidian Publish[12]). This culture and technical infrastructure, paired with the consumer-grade software, is what provides the ideal opportunity to explore and observe discourse graphs in authentic usage across a range of settings. Research to directly test the promise of discourse graphs for synthesis in this favorable setting is outside the scope of this document, but my hope is that this document will frame, support, and spur such work!

[8] https://web.hypothes.is/
[9] https://readwise.io/
[10] https://getmemex.com/
[11] https://roam.garden/
[12] https://obsidian.md/publish

References

[Ackerman et al., 2013] Ackerman, M. S., Dachtera, J., Pipek, V., and Wulf, V. (2013). Sharing Knowledge and Expertise: The CSCW View of Knowledge Management. *Computer Supported Cooperative Work (CSCW)*, 22(4-6):531–573.

[Ackerman and Halverson, 2004] Ackerman, M. S. and Halverson, C. (2004). Organizational Memory as Objects, Processes, and Trajectories: An Examination of Organizational Memory in Use. *Computer Supported Cooperative Work (CSCW)*, 13(2):155–189.

[Alton-Lee, 1998] Alton-Lee, A. (1998). A Troubleshooter's Checklist for Prospective Authors Derived from Reviewers' Critical Feedback. *Teaching and Teacher Education*, 14(8):887–90.

[Alvesson and Sandberg, 2011] Alvesson, M. and Sandberg, J. (2011). Generating research questions through problematization. *Academy of management review*, 36(2):247–271. Publisher: Academy of Management Briarcliff Manor, NY.

[anujacabraal, 2012] anujacabraal (2012). Why use NVivo for your literature review?

[Bhurke et al., 2015] Bhurke, S., Cook, A., Tallant, A., Young, A., Williams, E., and Raftery, J. (2015). Using systematic reviews to inform NIHR HTA trial planning and design: a retrospective cohort. *BMC Medical Research Methodology*, 15(1):108. 00000.

[Blake and Pratt, 2006] Blake, C. and Pratt, W. (2006). Collaborative information synthesis I: A model of information behaviors of scientists in medicine and public health. *Journal of the American Society for Information Science and Technology*, 57(13):1740–1749. 00097 _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/asi.20487.

[Boote and Beile, 2005] Boote, D. N. and Beile, P. (2005). Scholars Before Researchers: On the Centrality of the Dissertation Literature Review in Research Preparation. *Educational Researcher*, 34(6):3–15.

[Bosman and Kramer, 2016] Bosman, J. and Kramer, B. (2016). Innovations in scholarly communication - data of the global 2015-2016 survey. 00014 type: dataset.

[Bowker and Star, 2000] Bowker, G. C. and Star, S. L. (2000). *Sorting Things Out: Classification and Its Consequences*. The MIT Press, Cambridge, Massachusetts London, England, revised edition edition.

[Brush et al., 2016] Brush, M. H., Shefchek, K., and Haendel, M. (2016). SEPIO: A semantic model for the integration and analysis of scientific evidence. *CEUR Workshop Proceedings*, 1747.

[Cambrosio et al., 2020] Cambrosio, A., Campbell, J., Vignola-Gagné, E., Keating, P., Jordan, B. R., and Bourret, P. (2020). 'Overcoming the Bottleneck': Knowledge Architectures for Genomic Data Interpretation in Oncology. In Leonelli, S. and Tempini, N., editors, *Data Journeys in the Sciences*, pages 305–327. Springer International Publishing, Cham. 00000.

[Chan et al., 2020] Chan, J., Qian, X., Fenlon, K., and Lutters, W. G. (2020). Where the rubber meets the road: Identifying integration points for semantic publishing in existing scholarly practice. In *JCDL 2020 Workshop on Conceptual MOdeling*. 00000.

[Chandrasekharan et al., 2017] Chandrasekharan, E., Pavalanathan, U., Srinivasan, A., Glynn, A., Eisenstein, J., and Gilbert, E. (2017). You Can't Stay Here: The Efficacy of Reddit's 2015 Ban Examined Through Hate Speech. *Proceedings of the ACM on Human-Computer Interaction*, 1(CSCW):31:1–31:22. 00175.

[Chase and Simon, 1973] Chase, W. G. and Simon, H. A. (1973). The mind's eye in chess. In Chase, W. G., editor, *Visual Information Processing*, pages 215–281. New York, NY.

[Ciccarese et al., 2008] Ciccarese, P., Wu, E., Wong, G., Ocana, M., Kinoshita, J., Ruttenberg, A., and Clark, T. (2008). The SWAN biomedical discourse ontology. *Journal of Biomedical Informatics*, 41(5):739–751.

[Clark et al., 2014] Clark, T., Ciccarese, P. N., and Goble, C. A. (2014). Micropublications: A semantic model for claims, evidence, arguments and annotations in biomedical communications. *Journal of Biomedical Semantics*, 5:28.

[Clark and Kinoshita, 2007] Clark, T. and Kinoshita, J. (2007). Alzforum and SWAN: the present and future of scientific web communities. *Briefings in Bioinformatics*, 8(3):163–171. 00074.

[de Ribaupierre and Falquet, 2017] de Ribaupierre, H. and Falquet, G. (2017). Extracting discourse elements and annotating scientific documents using the SciAnnotDoc model: A use case in gender documents. *International Journal on Digital Libraries*, pages 1–16.

[de Waard, 2010] de Waard, A. (2010). From Proteins to Fairytales: Directions in Semantic Publishing. *IEEE Intelligent Systems*.

[de Waard et al., 2009] de Waard, A., Shum, S. B., Carusi, A., Park, J., Samwald, M., and Sándor, A. (2009). Hypotheses, Evidence and Relationships: The HypER Approach for Representing Scientific Knowledge Claims. In *Proceedings of the 8th International Semantic Web Conference, Workshop on Semantic Web Applications in Scientific Discourse*, page 12.

[Duflo, 2011] Duflo, E. (2011). Finding the right questions. *Newsletter of the Committee on the Status of Women in the Economics Profession*, pages 4–5. 00000.

[Ervin, 2008] Ervin, A.-M. (2008). Motivating authors to update systematic reviews: practical strategies from a behavioural science perspective. *Paediatric and perinatal epidemiology*, 22(0 1):33–37.

[Fleming et al., 2013] Fleming, P. S., Seehra, J., Polychronopoulou, A., Fedorowicz, Z., and Pandis, N. (2013). Cochrane and non-Cochrane systematic reviews in leading orthodontic journals: a quality paradigm? *European Journal of Orthodontics*, 35(2):244–248. Publisher: Oxford Academic.

[Goldstein, 2018] Goldstein, R. E. (2018). Are theoretical results 'Results'? *eLife*, 7:e40018. 00015 Publisher: eLife Sciences Publications, Ltd.

[Granello, 2001] Granello, D. H. (2001). Promoting Cognitive Complexity in Graduate Written Work: Using Bloom's Taxonomy as a Pedagogical Tool to Improve Literature Reviews. *Counselor Education and Supervision*, 40(4):292–307.

[Griffith et al., 2017] Griffith, M., Spies, N. C., Krysiak, K., McMichael, J. F., Coffman, A. C., Danos, A. M., Ainscough, B. J., Ramirez, C. A., Rieke, D. T., Kujan, L., Barnell, E. K., Wagner, A. H., Skidmore, Z. L., Wollam, A., Liu, C. J., Jones, M. R., Bilski, R. L., Lesurf, R., Feng, Y.-Y., Shah, N. M., Bonakdar, M., Trani, L., Matlock, M., Ramu, A., Campbell, K. M., Spies, G. C., Graubert, A. P., Gangavarapu, K., Eldred, J. M., Larson, D. E., Walker, J. R., Good, B. M., Wu, C., Su, A. I., Dienstmann, R., Margolin, A. A., Tamborero, D., Lopez-Bigas, N., Jones, S. J. M., Bose, R., Spencer, D. H., Wartman, L. D., Wilson, R. K., Mardis, E. R., and Griffith, O. L. (2017). CIViC is a community knowledgebase for expert crowdsourcing the clinical interpretation of variants in cancer. *Nature Genetics*, 49(2):170–174. 00239 Number: 2 Publisher: Nature Publishing Group.

[Groth et al., 2010] Groth, P., Gibson, A., and Velterop, J. (2010). The anatomy of a nanopublication. *Information Services & Use*, 30(1-2):51–56.

[Gruber and Barrett, 1974] Gruber, H. E. and Barrett, P. H. (1974). *Darwin on man: A psychological study of scientific creativity*. Darwin on man: A psychological study of scientific creativity. E. P. Dutton, New York, NY, England. Pages: xxv, 495.

[Gustafson, 2020] Gustafson, J. (2020). What is a Distributed Knowledge Graph? *KFG Notes*. Publisher: PubPub.

[Halford et al., 2013] Halford, S., Pope, C., and Weal, M. (2013). Digital Futures? Sociological Challenges and Opportunities in the Emergent Semantic Web. *Sociology*, 47(1):173–189. 00099 Publisher: SAGE Publications Ltd.

[Hars, 2001] Hars, A. (2001). Designing Scientific Knowledge Infrastructures: The Contribution of Epistemology. *Information Systems Frontiers*, 3(1):63–73.

[Hayes and Halpin, 2008] Hayes, P. J. and Halpin, H. (2008). In defense of ambiguity. *International Journal on Semantic Web and Information Systems (IJSWIS)*, 4(2):1–18. 00077 Publisher: IGI Global.

[Hjørland, 2002] Hjørland, B. (2002). Epistemology and the socio-cognitive perspective in information science. *Journal of the American Society for Information Science and Technology*, 53(4):257–270. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/asi.10042.

[Hoang and Schneider, 2018] Hoang, L. and Schneider, J. (2018). Opportunities for Computer Support for Systematic Reviewing - A Gap Analysis. In Chowdhury, G., McLeod, J., Gillet, V., and Willett, P., editors, *Transforming Digital Worlds*, Lecture Notes in Computer Science, pages 367–377. Springer International Publishing.

[Holbrook et al., 2004] Holbrook, A., Bourke, S., Lovat, T., and Dally, K. (2004). Investigating PhD thesis examination reports. *International Journal of Educational Research*, 41(2):98–120.

[Howard et al., 2020] Howard, J., Huang, A., Li, Z., Tufekci, Z., Zdimal, V., Westhuizen, H.-M. v. d., Delft, A. v., Price, A., Fridman, L., Tang, L.-H., Tang, V., Watson, G. L., Bax, C. E., Shaikh, R., Questier, F., Hernandez, D., Chu, L. F., Ramirez, C. M., and Rimoin, A. W. (2020). Face Masks Against COVID-19: An Evidence Review. 00111 Publisher: Preprints.

[Jonas and Kording, 2017] Jonas, E. and Kording, K. P. (2017). Could a Neuroscientist Understand a Microprocessor? *PLOS Computational Biology*, 13(1):e1005268. 00000 Publisher: Public Library of Science.

[Knight et al., 2019] Knight, I. A., Wilson, M. L., Brailsford, D. F., and Milic-Frayling, N. (2019). Enslaved to the Trapped Data: A Cognitive Work Analysis of Medical Systematic Reviews. In *Proceedings of the 2019 Conference on Human Information Interaction and Retrieval*, CHIIR '19, pages 203–212, New York, NY, USA. ACM. event-place: Glasgow, Scotland UK.

[Knoblich et al., 1999] Knoblich, G., Ohlsson, S., Haider, H., and Rhenius, D. (1999). Constraint relaxation and chunk decomposition in insight problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25(6):1534–1555.

[Kuhn and Dumontier, 2017] Kuhn, T. and Dumontier, M. (2017). Genuine semantic publishing. *Data Science*, 1(1-2):139–154.

[Liddo et al., 2012] Liddo, A. D., Sándor, A., and Shum, S. B. (2012). Contested Collective Intelligence: Rationale, Technologies, and a Human-Machine Annotation Study. *Computer Supported Cooperative Work (CSCW)*, 21(4-5):417–448.

[Lovitts, 2007] Lovitts, B. E. (2007). *Making the Implicit Explicit: Creating Performance Expectations for the Dissertation*. Stylus Publishing, Sterling, Va.

[Lutters and Ackerman, 2007] Lutters, W. G. and Ackerman, M. S. (2007). Beyond Boundary Objects: Collaborative Reuse in Aircraft Technical Support. *Computer Supported Cooperative Work (CSCW)*, 16(3):341–372.

[Marder, 2020] Marder, E. (2020). Theoretical musings. *eLife*, 9:e60703. 00001 Publisher: eLife Sciences Publications, Ltd.

[Marshall and Brush, 2004] Marshall, C. C. and Brush, A. J. B. (2004). Exploring the Relationship Between Personal and Public Annotations. In *Proceedings of the 4th ACM/IEEE-CS Joint Conference on Digital Libraries*, JCDL '04, pages 349–357, New York, NY, USA. ACM.

[McCaffrey, 2012] McCaffrey, T. (2012). Innovation Relies on the Obscure: A Key to Overcoming the Classic Problem of Functional Fixedness. *Psychological Science*, 23(3):215–218.

[McCrickard, 2012] McCrickard, D. S. (2012). Making Claims: Knowledge Design, Capture, and Sharing in HCI. *Synthesis Lectures on Human-Centered Informatics*, 5(3):1–125.

[McCrickard et al., 2013] McCrickard, D. S., Wahid, S., Branham, S. M., and Harrison, S. (2013). Achieving Both Creativity and Rationale: Reuse in Design with Images and Claims. In Carroll, J. M., editor, *Creativity and Rationale*, number 20 in Human–Computer Interaction Series, pages 105–119. Springer London.

[McElreath and Smaldino, 2015] McElreath, R. and Smaldino, P. E. (2015). Replication, Communication, and the Population Dynamics of Scientific Discovery. *PLOS ONE*, 10(8):e0136088. 00107 Publisher: Public Library of Science.

[McMahan and Evans, 2018] McMahan, P. and Evans, J. (2018). Ambiguity and Engagement. *American Journal of Sociology*, 124(3):860–912. 00010.

[McPhetres et al., 2020] McPhetres, J., Albayrak-Aydemir, N., Mendes, A. B., Chow, E. C., Gonzalez-Marquez, P., Loukras, E., Maus, A., O'Mahony, A., Pomareda, C., Primbs, M., Sackman, S., Smithson, C., and Volodko, K. (2020). A decade of theory as reflected in Psychological Science (2009-2019). Technical report, PsyArXiv. 00000 type: article.

[Newell, 1973] Newell, A. (1973). You can't play 20 questions with nature and win: Projective comments on the papers of this symposium. Technical report.

[O'Connor et al., 2019] O'Connor, A. M., Tsafnat, G., Gilbert, S. B., Thayer, K. A., Shemilt, I., Thomas, J., Glasziou, P., and Wolfe, M. S. (2019). Still moving toward automation of the systematic review process: a summary of discussions at the third meeting of the International Collaboration for Automation of Systematic Reviews (ICASR). *Systematic Reviews*, 8(1):57.

[Petrosino, 1999] Petrosino, A. (1999). Lead authors of cochrane reviews: Survey results. *Report to the Campbell Collaboration. Cambridge, MA: University of Pennsylvania*.

[Qian et al., 2019] Qian, X., Erhart, M. J., Kittur, A., Lutters, W. G., and Chan, J. (2019). Beyond iTunes for Papers: Redefining the Unit of Interaction in Literature Review Tools. In *Conference Companion Publication of the 2019 on Computer Supported Cooperative Work and Social Computing*, CSCW '19, pages 341–346, Austin, TX, USA. Association for Computing Machinery.

[Qian et al., 2020] Qian, X., Fenlon, K., Lutters, W. G., and Chan, J. (2020). Opening Up the Black Box of Scholarly Synthesis: Intermediate Products, Processes, and Tools. In *Proceedings of ASIST 2020*.

[Randall et al., 2011] Randall, D., Procter, R., Lin, Y., Poschen, M., Sharrock, W., and Stevens, R. (2011). Distributed ontology building as practical work. *International Journal of Human-Computer Studies*, 69(4):220–233.

[Renear and Palmer, 2009] Renear, A. H. and Palmer, C. L. (2009). Strategic reading, ontologies, and the future of scientific publishing. *Science*, 325(5942):828.

[Russell et al., 1993] Russell, D. M., Stefik, M. J., Pirolli, P., and Card, S. K. (1993). The Cost Structure of Sensemaking. In *Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems*, CHI '93, pages 269–276, New York, NY, USA. ACM.

[Scheel et al., 2020] Scheel, A. M., Tiokhin, L., Isager, P. M., and Lakens, D. (2020). Why Hypothesis Testers Should Spend Less Time Testing Hypotheses. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, page 1745691620966795. 00008.

[Seering et al., 2017] Seering, J., Kraut, R., and Dabbish, L. (2017). Shaping Pro and Anti-Social Behavior on Twitch Through Moderation and Example-Setting. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, CSCW '17, pages 111–125, New York, NY, USA. ACM.

[Shojania et al., 2007] Shojania, K. G., Sampson, M., Ansari, M. T., Ji, J., Doucette, S., and Moher, D. (2007). How Quickly Do Systematic Reviews Go Out of Date? A Survival Analysis. *Annals of Internal Medicine*, 147(4):224.

[Shum et al., 2000] Shum, S. B., Motta, E., and Domingue, J. (2000). ScholOnto: An ontology-based digital library server for research documents and discourse. *International Journal on Digital Libraries*, 3(3):237–248.

[Shum et al., 2006] Shum, S. J. B., Uren, V., Li, G., Sereno, B., and Mancini, C. (2006). Modeling naturalistic argumentation in research literatures: Representation and interaction design issues. *International Journal of Intelligent Systems*, 22(1):17–47.

[Silver, 2020] Silver, C. (2020). The crux of literature reviewing: structuring critical appraisals and using CAQDAS-packages. Library Catalog: www.qdaservices.co.uk.

[Star and Griesemer, 1989] Star, S. L. and Griesemer, J. R. (1989). Institutional Ecology, 'Translations' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. *Social Studies of Science*, 19(3):387–420.

[Strike and Posner, 1983] Strike, K. and Posner, G. (1983). Types of synthesis and their criteria.

[Talja and Maula, 2003] Talja, S. and Maula, H. (2003). Reasons for the use and non-use of electronic journals and databases: A domain analytic study in four scholarly disciplines. *Journal of Documentation*, 59(6):673–691. Publisher: MCB UP Ltd.

[van Rooij and Baggio, 2021] van Rooij, I. and Baggio, G. (2021). Theory Before the Test: How to Build High-Verisimilitude Explanatory Theories in Psychological Science. *Perspectives on Psychological Science*, page 1745691620970604. 00029 Publisher: SAGE Publications Inc.

[Willis et al., 2014] Willis, M., Sharma, S., Snyder, J., Brown, M., Østerlund, C., and Sawyer, S. (2014). Documents and Distributed Scientific Collaboration. In *Proceedings of the Companion Publication of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing*, CSCW Companion '14, pages 257–260, New York, NY, USA. ACM.

[Wolfswinkel et al., 2013] Wolfswinkel, J. F., Furtmueller, E., and Wilderom, C. P. M. (2013). Using grounded theory as a method for rigorously reviewing literature. *European Journal of Information Systems*, 22(1):45–55. Publisher: Taylor & Francis.