Where the rubber meets the road: Identifying integration points for semantic publishing in existing scholarly practice

JOEL CHAN, XIN QIAN, KATRINA FENLON, and WAYNE LUTTERS, University of Maryland

Semantic publishing has significant potential to transform scholarly work. While much progress has been made on conceptual models and technical infrastructure, authorship remains an open problem. Here, we explore whether and how semantic publishing labor might be integrated into the existing practices of scholars reading and synthesizing the literature. From a series of studies of scholars, we observe rich practices across a variety of workflows and tool ecologies that overlap with key aspects of semantic publishing: 1) creating granular knowledge artifacts (Compression), extracting and specifying provenance and contextual details (Contextualizability), and 3) specifying semantically typed entities and relations between knowledge units (Composability). We discuss implications of these findings for developing sustainable, scholar-powered models of semantic publishing.

CCS Concepts: • **Information systems** → **Document representation**; • **Human-centered computing** → **Interaction paradigms**.

Additional Key Words and Phrases: synthesis, semantic publishing, infrastructure, practice

ACM Reference Format:
Joel Chan, Xin Qian, Katrina Fenlon, and Wayne Lutters. 2020. Where the rubber meets the road: Identifying integration points for semantic publishing in existing scholarly practice. In *JCDL 2020 Workshop on Conceptual Modeling, August 01, 2020, Wuhan, China.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/1122445.1122456

1 INTRODUCTION

Semantic publishing holds significant promise for transforming scholarly knowledge work [4, 17, 23, 28]. In particular, there is exciting potential for formal semantics to *augment* — rather than replace — scholars' ability to synthesize knowledge from the literature. We resonate with Renear and colleagues' [23] articulation of a vision of scholarly communication infrastructures that are substantially enriched by "computational access to causal and ontological relationships", and an "increasingly rich layer of indexing, linking, and annotation information" (p. 832).

Much progress has been made towards this vision. We now have a robust ecology of conceptual models, formal standards, and technical infrastructure for semantic publishing of granular, semantically interlinked claims and concepts — e.g., as annotations [8], micropublications [9], nanopublications [13], or webs of arguments [26] — along with detailed provenance information, such as evidence [5], uncertainty [11], and connections to project-specific contextual metadata via Research Objects [3].

Yet, adoption of semantic publishing, particularly outside biomedical domains, remains low, and is often concentrated amongst a relatively small set of authors. For example, while Kuhn et al [18]

JCDL '20, August 01, 2020, Wuhan, China

ACM ISBN 978-1-4503-XXXX-X/18/06…$15.00
https://doi.org/10.1145/1122445.1122456

report around 10 million nanopublications published by 2018, almost all were within bioinformatics, and contributions were overwhelmingly from a small (N=41) set of authors. Discussion of the roots of this slow adoption is ongoing, but we agree with an emerging consensus that the barriers are not strictly technical (in the sense of models / standards) [17, 23]: rather, the key barrier to the vision of semantic publishing is an *authoring bottleneck*: **what is the labor of semantic publishing, and who will do it?**

Currently, the bulk of semantic publishing seems to be done by specialized ontology engineers, crowdsourcing, or volunteer curators. This specialized curator model is powerful when funding or significant volunteer labor is available. However, we are more compelled by a broader vision of **scholar-powered semantic publishing** [16, 23] — which aims to integrate the labor of semantic publishing into different points of the research lifecycle, such as manuscript preparation [14, 19] or peer review [6], where these activities are not a separate task, but part of their routine scholarly practices. We situate our efforts within this broader scholar-powered approach because of our interest in developing *sustainable infrastructures* for semantic publishing and synthesis.

In this paper, investigate the following question: **where (if at all) are there integration points between semantic publishing and existing practices of reading and synthesizing the scholarly literature**? Identifying these integration points could help us see where we might be able to "graft" semantic authoring tools and interfaces into scholarly practices to leverage the rich semantic work that is already happening. We might also improve sustainability by improving, rather than disrupting, existing scholarly practices, better aligning collective good with individual benefits. These strategies of integration hold significant promise for building more sustainable knowledge infrastructures [12].

2 METHODS

2.1 Data sources

To address these questions, we draw observations from three complementary data sources. First, we draw on results from two empirical studies of scholars' synthesis work. The Protocol Study includes guided tours [30] of ten scholars' workflows and tool setups, along with fine-grained think-aloud protocols of the same scholars' synthesis work, including both initial processing of sources (45 minutes) and later reuse (45 minutes), observed in-person and with a video recording from a head-mounted camera. The Interview Study involves in-depth semi-structured contextual interviews with 10 PhD students reflecting on their process and setup for a recent synthesis effort (more details available in [22]). Finally, we draw on initial observations from the first author's preparation for an extended auto-ethnographic and participatory observation study of knowledge management / hacker communities. While no formal data has yet been collected (the IRB for this study is in process), here we report observations from publicly available sources from the community, such as Youtube videos and blog posts.

2.2 Our lens on the labor of semantic publishing

For this work, we define semantic publishing labor as any set of practices that directly or indirectly produce artifacts — such as notes, annotations, or documents — that could serve as resources for semantic publicationsMore specifically, we look for instances where scholars create artifacts that satisfy one or more of three primary categories of semantic properties:

(1) Compression: afford reasoning about more atomic / granular units of knowledge, such as claims or concepts.
(2) Contextualizability: afford reasoning about context. For example, if the publication is about a particular claim, it should include information that a scholar would need to appropriately

(re)interpret the claim, such as authorship, provenance, uncertainty, evidence, or relationships with other claims or concepts.

(3) Composability: afford composition of atomic units of knowledge, such as concepts, into more complex representations, such as arguments, maps, timelines, and causal models. This is typically enabled by encoding semantic publications with some kind of formal representation (e.g., typed entities and relations), since this enables computational support for retrieval, aggregation, and reasoning about collections of knowledge units.

We expect these categories of semantic properties to anyone familiar with the Semantic Web and semantic publishing. For example, Compression resembles the core concerns of standards like micropublications [9] and nanopublications [13], which were developed in part to enable reasoning over more granular units of knowledge. Similarly, the property of Contextualizability resembles the goal of representing evidence [5], uncertainty [11], and provenance of publications [13]. Finally, the idea of formal semantics and Composability connect well to the core vision of enabling machine assisted-reasoning and higher-level synthesis in the Semantic Web [4, 17].

In addition to drawing on the literature on semantic publishing, we have also synthesized these ideas from fields relevant to understanding the scholarly practice of synthesis, such as sensemaking [32], knowledge reuse [1], and creative problem formulation [21].

3 FINDINGS

We organize our findings by sketching out a series of personas aggregated from and grounded in the data from our three studies. We present them not as formal contributions in their own right (we expect robust discussion and disagreement about their coherence, sufficiency, and differences); instead, we present them as thematic clusters that emerged from our observations, which help organize and make sense of many disparate examples of semantic-publishing-adjacent practices. These personas are also useful as conceptual handles that can frame downstream design/development work towards integrating semantic authoring into scholarly practice.

Here, we discuss three personas from our data: 1) **Virtuosos**, who optimize purpose-built mainstream tools for synthesis, 2) **Explorers**, who adopt niche synthesis tools or appropriate other tools for synthesis work, and 3) **Hackers**, who create new tools and practices to support synthesis work.

3.1 The Virtuosos

The most common persona observed in our data was the **virtuoso**. She focuses a lot less on the tools she uses, and more on the work she is doing. She is unlikely to shop around a lot for new tools. Instead, she focuses on configuring and optimizing the tools she has to get the job done. Overall, the virtuoso's toolset tends to consist of familiar "mainstream" tools like Google Scholar, word processors like Word and Google Docs, reference managers like Zotero and Mendeley, printed papers, post-its, and highlighters, and some note-taking apps like OneNote. Yet, virtuosos use these tools to support sophisticated practices that resemble many core aspects of semantic publishing work as we have defined it.

For example, the virtuoso often employs sophisticated approaches to annotation. They use highlights and notes to identify and extract key ideas in source texts with (Compression). Often these annotations are color-coded to identify types of building blocks (Composability). They also mark up key contextual details that might help with downstream interpretation and synthesis (Contextualizability). Consider the concrete example in Figure 1, where a participant in our Protocol Study used a blue highlight to identify a key idea that "slavery cannot be represented", and a green highlight to explicitly mark the author of the quote. These color-codes are used to

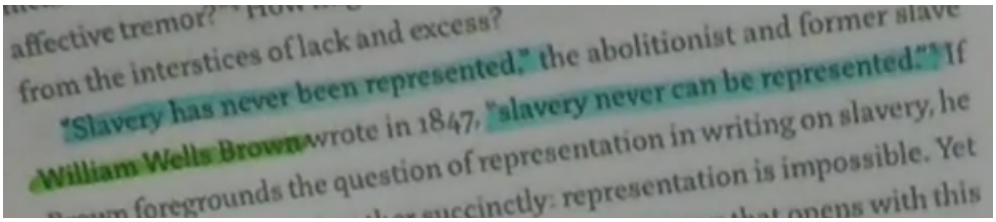

Fig. 1. Example annotation with Compression and Contextualizability, using color coding.

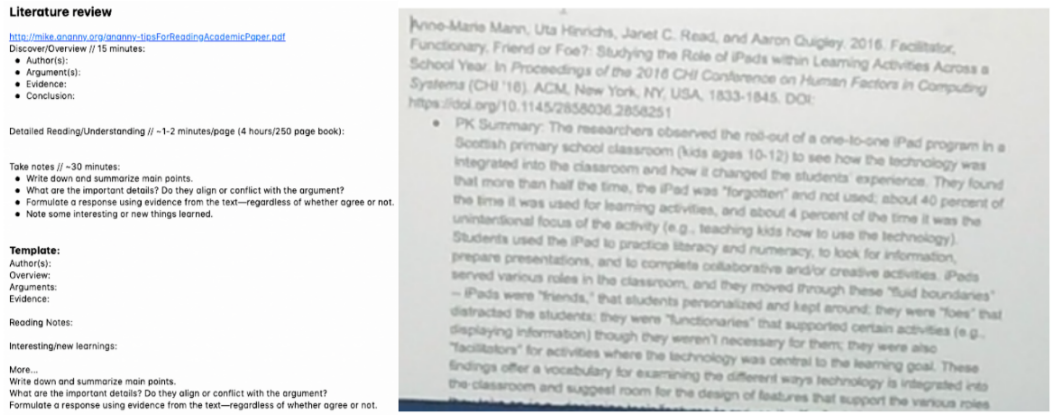

Fig. 2. Examples of structured summaries that include features of Compression, Contextualizability, and Composability.

support sensemaking in the moment, and also aid recall of the key ideas and contextual details when revisiting sources.

Virtuosos are also likely to employ sophisticated approaches to summarizing key sources. These summaries often include details about key claims from the source (Compression), along with contextual details that aid interpretation (Contextualizability). The claims/contributions are often linked together into an argument with typed relations like alignment, contradiction, or support (Composability). Consider the template from a participant in the interview study (Figure 2, left), which is meant to ensure that he captures the key arguments and evidence (with contextual details) for important papers/books. Other participants we've observed create summaries that are less obviously structured (see Figure 2, right), but include many similar elements, such as *key ideas*, *contextual details* like authors and methods details, and some *semantics in relations between ideas*, expressed in natural language. Often these more sophisticated practices are motivated by the desire to share knowledge with collaborators and advisors.

3.2 The Explorers

We have also encountered a relatively small number of **explorers** (about 10-15% of our participants), who are defined by their adoption or appropriation of more exotic tools. These tools provide affordances that are either missing or difficult to achieve in more mainstream tools. In our conversations with these participants, we often hear themes about a desire to "stay grounded in the sources" (Contextualizability), and to better support "linking and reuse of ideas across papers" (Compression and Composability).

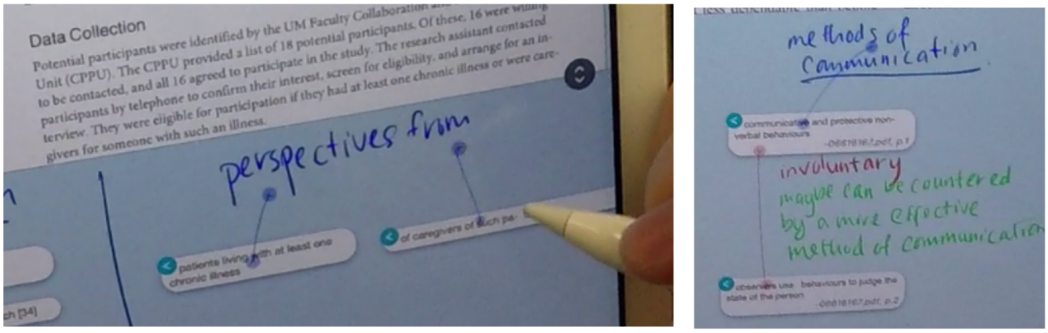

Fig. 3. Example excerpts and notes on LiquidText canvas, with hooks to context of excerpts, as well as semanticaly typed relations between excerpts and notes.

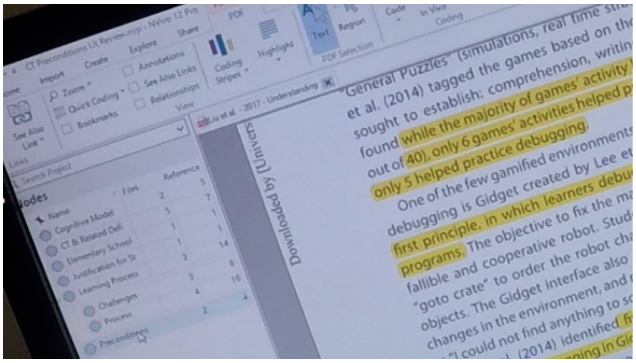

Fig. 4. Screenshot from NVivo interface in use by one of our participants, showcasing "coding" of excerpts from a research paper into semantically structured hierarchies.

For example, a participant from our Protocol Study uses the LiquidText[1] [29] iPad app for reading and synthesizing, which was purpose-built to support active reading of scholarly papersIn the app, she is able to extract segments from PDFs onto a canvas (Compression), where she can relate them in "typed" ways on a canvas (Composability), such as identifying different sources of perspectives on an issue by drawing "perspectives from..." connections between two text segments describing participant populations (Fig. 3, left panel). Each of these excerpts have rich software-enabled affordances for Contextualizability: for example, each segment can include signals of the name of the PDF and the page number from which the excerpt came (Fig. 3, right panel). Additionally, by tapping on the arrow icon, the user can follow a live "link" back to the specific location in the PDF from which the excerpt came. Notably, these affordances come "for free" from the app, without any extra manual effort by the user.

Going a bit further than adopting niche tools that were purpose-built for synthesis work, some scholars appropriate tools meant for other purposes, such as qualitative data analysis software (QDAS). For example, in our Protocol Study, one participant uses NVivo [2] after struggling with OneNote and other PDF readers. Using the program, he can "code" excerpts from papers, and place them in a code tree for sensemaking (see Fig. 4). This allows him to achieve Compression

[1]https://www.liquidtext.net/
[2]https://www.qsrinternational.com/NVivo-qualitative-data-analysis-software/home

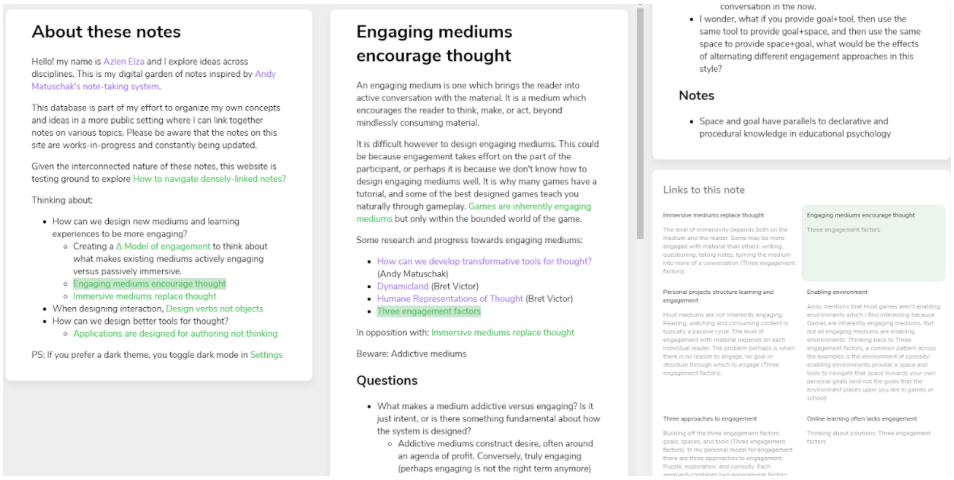

Fig. 5. Example of a networked notebook. Here, each "pane" is a note. Notice the atomic titles, in the middle pane, the linkages to other notes (green links), and "bi-directional links to the note on the right ("links to this note" pane). These notebooks also include links to sources (purple links).

more explicitly, creating excerpts that are manipulable by themselves, but also make sure they are typed and have particular semantically meaningful relationships to other compressed segments (Composability). Similar to LiquidText, highlighting in this way also gives him substantial ability to achieve Contextualizability, since each coded excerpt retains direct links that the user can follow back to the source itself. Each excerpt can also be viewed in narrow (i.e., only highlighted claim) and broad (claim plus some surrounding text) context settings. Importantly, as with LiquidText, these affordances are enabled with almost no manual effort other than selecting the relevant PDF segments. While still not mainstream, the use of QDAS for literature reviews is common enough that there are tutorials [2] and even conceptual papers [31] discussing this approach.

Another example of Explorers' niche tools is a growing ecosystem of "networked notebooks", which are a particularly interesting category with deep intellectual roots in hypertext [7, 20]. A prominent example is RoamResearch[3]; others include TiddlyWiki[4] and Obsidian.md[5]. Scholars who use these tools create and maintain relatively atomic notes on concepts or some kind of focused claim (Compression). These notes are densely linked to each other (Composability), typically bi-directionally: every time a link is made from one source note to a target note, both the source and target notes record the link (see Fig. 5). In this way, links between notes are more accessible, since links can be followed from either source or target notes. This, together with other affordances like autocompletion of links during text editing, enables easier tending to connections between notes (Composability). The links also enable users to compress quite complex ideas into a single statement (e.g., "knowledge is contextual") while retaining links to the less compressed ideas that "unpack different aspects and subtleties of the more complex idea. In this way, tending to the notes and links also enhances the Contextualizability of each entry.

Finally, since notes that collect bi-directional links can be as small as a single concept, the act of deliberately linking notes partially accomplishes the work of developing folksonomies (Composability). A key affordance in these networked notebooks is that it is quite easy to rename

[3]https://roamresearch.com/
[4]https://tiddlywiki.com/
[5]https://obsidian.md/

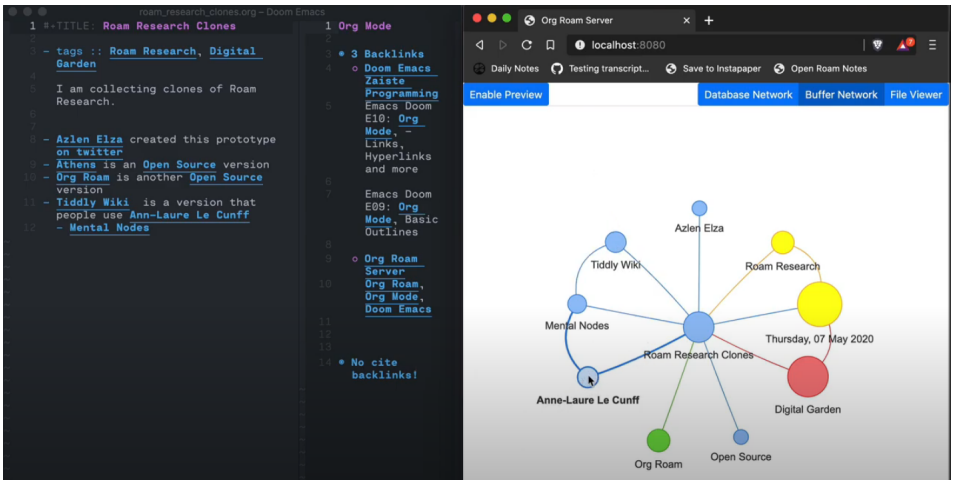

Fig. 6. Screenshot from org-roam interface, showcasing key features of atomic notes and bi-directional links that support Compression, Contextualizability and Composability.

note titles, with changes automatically propagating throughout the database. This enables more agile and evolving folksonomies.

3.3 The Hackers

Finally, the first-author's initial participant observation adds the **Hacker** persona: scholars who have homebrewed their own synthesis system. While Hackers' tooling can look far more exotic than those of the Virtuosos or even Explorers, in practice, their workflows and practices are quite similar to those of the Explorers in terms of the level at which they do work to support Compression, Contextualizability and Composability. A good example of a Hacker system is the plain-text, command-line-based org-mode[6] [27], which has been extended by the package org-roam[7] to support affordances of bi-directional linking and easily contextualizable links to source documents and excerpts (see Fig. 7).

The Hacker persona is worth calling out separately from the Explorers because innovations from Hackers can percolate back into the more "commercial" tools, as plugins/extensions that Explorers or even Virtuosos can adopt. For instance, many Explorers use mechanisms for extracting annotations from PDFs to their note-taking systems; one important example of this is the zotfile extension, which enables automatic extraction of annotations, with pointers back to the specific PDF location from which the extraction came (not unlike LiquidText and NVivo), significantly enhancing Contextualizability.

4 DISCUSSION

In this paper, we sought to explore whether and how scholars might be engaged in work practices that overlap with semantic publishing labor, defined as producing artifacts with features of Compression, Contextualizability, and/or Composability. Our preliminary findings across our three studies suggest that a surprising amount of semantic publishing labor is already happening:

[6]https://orgmode.org/
[7]https://org-roam.readthedocs.io/en/master/

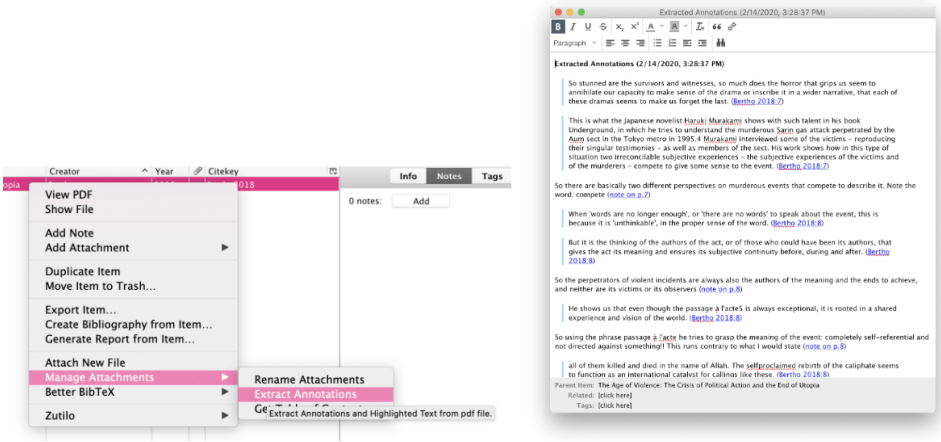

Fig. 7. The zotfile extension on the popular open-source Zotero reference manager enables stronger CONTEXTUALIZABILITY for PDF annotations.

we have presented cases of real scholars shaping their tools, practices, and workflows to enable these semantic properties.

The fact that these rich practices were intrinsically motivated rather than externally imposed suggests that there is significant opportunity to align the more distant benefits of semantic publishing — often framed as for the good of the impersonal "field" — with more immediate and personal incentives, such as thinking through a problem better or enhancing collaborations with others. Integrating semantic publishing labor with these intrinsic incentives could be a powerful path to sustainability. One open challenge and opportunity to realize this path is that the practices of semantic work, especially of virtuosos, are often experienced as arduous. Consequently, these practices are often abandoned or attenuated, especially once scholars move past the beginning stage and are subject to increased external pressures.

Another key question is how to bridge the often informal semantics in scholars' practices with the formality of most semantic publishing standards. For example, the notes and folksonomies, even of Explorers and Hackers, tend to be quite contextual, which is a liability if the goal of semantic publishing is seen as creating a field- or even science-wide Semantic Web. But what if the goal were changed to enabling better interfaces, but not necessarily via standardization? Could alternative conceptual modeling commitments — such as a "federated wiki" approach [10] — that respect the deeply contextual nature of knowledge [15, 24] open up more interesting and feasible ways to integrate semantic publishing work with the rich semantic practices we have described? How much value might we get from more "local" Semantic Webs? For example, perhaps social contexts of scholarly work like advising, collaborations, literature repositories for labs, might be good "first stops" for semantic publishing, to provide immediate value for interoperability and other benefits well-suited for semantic We think these issues deserve deeper discussion from a conceptual modeling standpoint.

Finally, we need to understand how much of this "organic" semantic publishing labor is actually happening. What is the scale at which scientists are producing annotations and notes? How much currently "wasted effort" is potentially available to power semantic publishing? While the absolute numbers of Hackers might be small, the Explorers group may cover more than we think, especially beginning students who might be more open to trying different tools and workflows. As tools mature in this space and become more mainstream, we might see adoption by Virtuosos. The

first author has personally witnessed adoption of some of the more niche tools (e.g., MarginNote, RoamResearch) by at least three of the students in the interview study, and the data for the Protocol Study was not explicitly sampled for Explorers, yet includes both NVivo and LiquidText. In this light, we could take inspiration from diffusion of innovation theory [25] to focus first on integrating semantic publishing into a smaller number of enthusiastic adopters to provide visibility and potentially trialability/testability, then expanding to wider adoption.

REFERENCES

[1] Mark S. Ackerman, Juri Dachtera, Volkmar Pipek, and Volker Wulf. 2013. Sharing Knowledge and Expertise: The CSCW View of Knowledge Management. *Computer Supported Cooperative Work (CSCW)* 22, 4-6 (Aug. 2013), 531–573. https://doi.org/10.1007/s10606-013-9192-8

[2] anujacabraal. 2012. Why use NVivo for your literature review? https://anujacabraal.wordpress.com/2012/08/01/why-use-nvivo-for-your-literature-review/

[3] Sean Bechhofer, Iain Buchan, David De Roure, Paolo Missier, John Ainsworth, Jiten Bhagat, Philip Couch, Don Cruickshank, Mark Delderfield, Ian Dunlop, Matthew Gamble, Danius Michaelides, Stuart Owen, David Newman, Shoaib Sufi, and Carole Goble. 2013. Why linked data is not enough for scientists. *Future Generation Computer Systems* 29, 2 (Feb. 2013), 599–611. https://doi.org/10.1016/j.future.2011.08.004

[4] Tim Berners-Lee and James Hendler. 2001. Publishing on the semantic web. *Nature* 410, 6832 (April 2001), 1023–1024. https://doi.org/10.1038/35074206 Number: 6832 Publisher: Nature Publishing Group.

[5] Matthew H Brush, Kent Shefchek, and Melissa Haendel. 2016. SEPIO: A Semantic Model for the Integration and Analysis of Scientific Evidence. In *CEUR Workshop Proceedings*. 6.

[6] Cristina-Iulia Bucur, Tobias Kuhn, and Davide Ceolin. 2019. Peer Reviewing Revisited: Assessing Research with Interlinked Semantic Comments. In *Proceedings of the 10th International Conference on Knowledge Capture*. ACM, Marina Del Rey CA USA, 179–187. https://doi.org/10.1145/3360901.3364434

[7] Vannevar Bush. 1945. As we may think. *Atlantic Monthly* (1945).

[8] Paolo Ciccarese, Marco Ocana, Leyla Jael Garcia Castro, Sudeshna Das, and Tim Clark. 2011. An open annotation ontology for science on web 3.0. *Journal of Biomedical Semantics* 2, 2 (May 2011), S4. https://doi.org/10.1186/2041-1480-2-S2-S4

[9] Tim Clark, Paolo N. Ciccarese, and Carole A. Goble. 2014. Micropublications: a semantic model for claims, evidence, arguments and annotations in biomedical communications. *Journal of Biomedical Semantics* 5 (July 2014), 28. https://doi.org/10.1186/2041-1480-5-28

[10] Alan Davoust, Alexander Craig, Babak Esfandiari, and Vincent Kazmierski. 2015. P2Pedia: a peer-to-peer wiki for decentralized collaboration. *Concurrency and Computation: Practice and Experience* 27, 11 (2015), 2778–2795. https://doi.org/10.1002/cpe.3420 _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/cpe.3420.

[11] Anita De Waard and Jodi Schneider. 2012. Formalising uncertainty: an ontology of reasoning, certainty and attribution (ORCA). In *Proceedings of the Joint 2012 International Conference on Semantic Technologies Applied to Biomedical Informatics and Individualized Medicine - Volume 930 (SATBI+SWIM'12)*. CEUR-WS.org, Boston, 10–17.

[12] Paul N Edwards, Geoffrey C Bowker, Steven J Jackson, and Robin Williams. 2009. Introduction: an agenda for infrastructure studies. *Journal of the Association for Information Systems* 10, 5 (2009), 6.

[13] Paul Groth, Andrew Gibson, and Jan Velterop. 2010. The anatomy of a nanopublication. *Information Services & Use* 30, 1-2 (Jan. 2010), 51–56. https://doi.org/10.3233/ISU-2010-0613

[14] Tudor Groza, Knud Möller, Siegfried Handschuh, Diana Trif, and Stefan Decker. 2007. SALT: Weaving the Claim Web. In *The Semantic Web (Lecture Notes in Computer Science)*, Karl Aberer, Key-Sun Choi, Natasha Noy, Dean Allemang, Kyung-Il Lee, Lyndon Nixon, Jennifer Golbeck, Peter Mika, Diana Maynard, Riichiro Mizoguchi, Guus Schreiber, and Philippe Cudré-Mauroux (Eds.). Springer Berlin Heidelberg, 197–210.

[15] Susan Halford, Catherine Pope, and Mark Weal. 2013. Digital Futures? Sociological Challenges and Opportunities in the Emergent Semantic Web. *Sociology* 47, 1 (Feb. 2013), 173–189. https://doi.org/10.1177/0038038512453798 Publisher: SAGE Publications Ltd.

[16] Tobias Kuhn, Paolo Emilio Barbano, Mate Levente Nagy, and Michael Krauthammer. 2013. Broadening the Scope of Nanopublications. In *The Semantic Web: Semantics and Big Data (Lecture Notes in Computer Science)*, Philipp Cimiano, Oscar Corcho, Valentina Presutti, Laura Hollink, and Sebastian Rudolph (Eds.). Springer Berlin Heidelberg, 487–501.

[17] Tobias Kuhn and Michel Dumontier. 2017. Genuine semantic publishing. *Data Science* 1, 1-2 (Jan. 2017), 139–154. https://doi.org/10.3233/DS-170010

[18] Tobias Kuhn, Albert Meroño-Peñuela, Alexander Malic, Jorrit H. Poelen, Allen H. Hurlbert, Emilio Centeno Ortiz, Laura I. Furlong, Núria Queralt-Rosinach, Christine Chichester, Juan M. Banda, Egon Willighagen, Friederike Ehrhart,

Chris Evelo, Tareq B. Malas, and Michel Dumontier. 2018. Nanopublications: A Growing Resource of Provenance-Centric Scientific Linked Data. *arXiv:1809.06532 [cs]* (Sept. 2018). http://arxiv.org/abs/1809.06532 arXiv: 1809.06532.

[19] Barend Mons. 2005. Which gene did you mean? *BMC Bioinformatics* 6, 1 (June 2005), 142. https://doi.org/10.1186/1471-2105-6-142

[20] T. H. Nelson. 1965. Complex information processing: a file structure for the complex, the changing and the indeterminate. In *Proceedings of the 1965 20th national conference (ACM '65)*. Association for Computing Machinery, Cleveland, Ohio, USA, 84–100. https://doi.org/10.1145/800197.806036

[21] Stellan Ohlsson. 1992. Information-processing explanations of insight and related phenomena. In *Advances in the psychology of thinking*, M. T. Keane and K. J. Gilhooly (Eds.). Vol. 1. Harvester Wheatsheaf, Hertfordshire, UK, 1–44.

[22] Xin Qian, Katrina Fenlon, Wayne G. Lutters, and Joel Chan. 2020. Opening Up the Black Box of Scholarly Synthesis: Intermediate Products, Processes, and Tools.. In *Proceedings of ASIST 2020*.

[23] Allen H. Renear and Carole L. Palmer. 2009. Strategic Reading, Ontologies, and the Future of Scientific Publishing. *Science* 325, 5942 (Aug. 2009), 828–832. https://doi.org/10.1126/science.1157784

[24] David Ribes, Andrew S Hoffman, Steven C Slota, and Geoffrey C Bowker. 2019. The logic of domains. *Social Studies of Science* 49, 3 (June 2019), 281–309. https://doi.org/10.1177/0306312719849709 Publisher: SAGE Publications Ltd.

[25] Everett M. Rogers. 2003. *Diffusion of Innovations, 5th Edition* (5th edition ed.). Free Press, New York.

[26] Jodi Schneider, Tudor Groza, and Alexandre Passant. 2013. A review of argumentation for the Social Semantic Web. *Semantic Web* 4, 2 (Jan. 2013), 159–218. https://doi.org/10.3233/SW-2012-0073 Publisher: IOS Press.

[27] Eric Schulte, Dan Davison, Thomas Dye, and Carsten Dominik. 2012. A Multi-Language Computing Environment for Literate Programming and Reproducible Research. *Journal of Statistical Software* 46, 1 (Jan. 2012), 1–24. https://doi.org/10.18637/jss.v046.i03 Number: 1.

[28] David Shotton. 2009. Semantic publishing: the coming revolution in scientific journal publishing. *Learned Publishing* 22, 2 (2009), 85–94. https://doi.org/10.1087/2009202 _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1087/2009202.

[29] Craig S. Tashman and W. Keith Edwards. 2011. LiquidText: A Flexible, Multitouch Environment to Support Active Reading. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. ACM, New York, NY, USA, 3285–3294. https://doi.org/10.1145/1978942.1979430

[30] Leslie Thomson. 2015. The guided tour technique in information science: Explained and illustrated. *Proceedings of the Association for Information Science and Technology* 52, 1 (2015), 1–5. https://doi.org/10.1002/pra2.2015.1450520100135

[31] Joost F Wolfswinkel, Elfi Furtmueller, and Celeste P M Wilderom. 2013. Using grounded theory as a method for rigorously reviewing literature. *European Journal of Information Systems* 22, 1 (Jan. 2013), 45–55. https://doi.org/10.1057/ejis.2011.51 Publisher: Taylor & Francis.

[32] Zhang Pengyi and Soergel Dagobert. 2014. Towards a comprehensive model of the cognitive process and mechanisms of individual sensemaking. *Journal of the Association for Information Science and Technology* 65, 9 (Aug. 2014), 1733–1756. https://doi.org/10.1002/asi.23125