Episemantics: Aboutness as Aroundness †

Elliott Hauser* and Joseph T. Tennis**

*University of North Carolina at Chapel Hill, School of Information and Library Science,
100 Manning Hall, Chapel Hill, NC 27599-3360, <eah13@email.unc.edu>
**University of Washington, The Information School, Box 352840,
Mary Gates Hall, Ste. 370, Seattle, WA 98195-2840, <jtennis@uw.edu>

Elliott Hauser, MSIS, is a doctoral candidate and at the University of North Carolina at Chapel Hill, where he is a member of the Royster Society of Fellows. His research interests include information organization, data science education, the ethics of algorithms, and sociotechnical methods of systems analysis. His dissertation research on the phenomenon of certainty in information systems was recently selected for the Litwin Books Dissertation Research Award. Elliott has a background in programming, co-founded the online coding education platform Trinket, and teaches programming and information studies topics at the graduate and undergraduate levels.

Joseph T. Tennis is Associate Professor and Associate Dean of Faculty Affairs at the University of Washington Information School, Adjunct Associate Professor in linguistics, and a member of the textual studies, computational linguistics, and museology faculty advisory groups at the University of Washington. He served as President of the International Society for Knowledge Organization from 2014-2018. He is on the *Library Quarterly* and *Knowledge Organization* editorial boards and served as a core member of the InterPARES Trust research team from 2005- 2019. Tennis works in classification theory, metadata versioning, ethics of knowledge organization work, descriptive informatics, and authenticity.

Hauser, Elliott and Joseph T. Tennis. 2019. "Episemantics: Aboutness as Aroundness." *Knowledge Organization* 46 (8): 590-595. 16 references. DOI:10.5771/0943-7444-2019-8-590.

Abstract: Aboutness ranks amongst our field's greatest bugbears. What is a work about? How can this be known? This mirrors debates within the philosophy of language, where the concept of representation has similarly evaded satisfactory definition. This paper proposes that we abandon the strong sense of the word aboutness, which seems to promise some inherent relationship between work and subject, or, in philosophical terms, between word and world. Instead, we seek an etymological reset to the older sense of aboutness as "in the vicinity, nearby; in some place or various places nearby; all over a surface." To distinguish this sense in the context of information studies, we introduce the term episemantics. The authors have each independently applied this term in slightly different contexts and scales (Hauser 2018a; Tennis 2016), and this article presents a unified definition of the term and guidelines for applying it at the scale of both words and works. The resulting weak concept of aboutness is pragmatic, in Star's sense of a focus on consequences over antecedents, while reserving space for the critique and improvement of aboutness determinations within various contexts and research programs. The paper finishes with a discussion of the implication of the concept of episemantics and methodological possibilities it offers for knowledge organization research and practice. We draw inspiration from Melvil Dewey's use of physical aroundness in his first classification system and ask how aroundness might be more effectively operationalized in digital environments.

Received: 29 September 2019; Revised: 28 October 2019; Accepted: 31 October 2019

Keywords: meaning, aboutness, classification, subject, episemantics

† Presented at NASKO 2019: Knowledge Organization: Community and Computation, sponsored by ISKO-Canada/United States at Drexel University, Philadelphia, Pennsylvania, USA, June 13-14, 2019. We are grateful to an anonymous reviewer for encouraging us to explore the literal aroundness central to early library classification schemes.

1.0 Introduction

This paper discusses and synthesizes two conceptions of the term episematics developed independently by the authors in prior work. Both conceptions deny that meaning is an inherent property of language, but take distinct approaches in relating this idea to the field of KO, and information studies more broadly. Tennis (2016) proposes episemantics as a potential new field of study, analogous to epigenetics, just recently made possible due to the advent of new technologies and research methods. Hauser (2018a) asks what it might mean to remove aboutness as a core com-

https://doi.org/10.5771/0943-7444-2019-8-590 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

ponent of our understanding of information at all. After discussing both proposals, we present a synthesis of each that connects Tennis's methodological proposal with Hauser's theoretical approach via a shared pragmatism, in Star's sense of "consequences, not antecedents." The result is discussed in relation to classification theory and particularly in light of Melvil Dewey's pragmatic approach to his first classification system. Finally, we consider what this might mean for organization practices in digital environments.

2.0 Tennis's episemantics: Epigenetics for KO

> The idea of episemantics is to account for meaning as it changes over time outside of the scheme, and relate that to the scheme. Instead of reifying the subject in the context of the scheme alone, and linking those subjects to a body of documents, episemantics would establish models for the investigation of particular relationships. These models would be networks of meaning that show how relationships between terms are established.
>
> Tennis, "Methodological Challenges in Scheme Versioning and Subject Ontogeny Research," 578

Tennis employs an analogy to epigenetics, the study of the effects and behavior of genetic material within living organisms, as opposed to limiting the scope of study to "a" genetic sequence. Epigenetic research has determined that the activation and inhibition of specific genes often occurs in response to environmental or organismal factors in what must be regarded as emergent properties not detectable from a mere sequence of nucleotides. Just as rapid and inexpensive sequencing techniques allowed the relative rates of expression of genes to be contemplated as a subject of research, thereby enabling a new field, Tennis envisions digital methods providing new epistemic access to phenomena of deep importance to subject ontogeny research.

The challenge that Tennis's proposed epistemantics addresses is the "location" of meaning in indexing languages in relation to literary warrant. Most indexing languages rely on their structure and the intellect of the indexer to triangulate the meaning in indexing terms. Further, meaning can be inferred from the range of materials that are indexed with that term. What has heretofore been lacking is the link to the literature except in the rare cases of citations to literature in thesauri (Soergel 1974) and *Library of Congress Subject Headings* (e.g., Library of Congress 2019). However, there are no explicit links between these sparse citations and wider network of literature.

Elsewhere, Tennis has presented on the circumstantial evidence relating term appearance in the *Dewey Decimal Classification* to literary warrant using the Google Books and Hathi Trust corpora (Tennis 2012). Constructing an episemantic methodology would allow for explicit links, revealing how terms were deployed in literature.

Essentially, Tennis's exploratory proposal would allow subject ontogeny researchers to connect the meaning of subjects to both the use of those subject terms (via large scale analysis of cataloging records) and the separate use of the same terms outside the context of knowledge organization (via the methods of corpus linguistics). While these methods do not eliminate the methodological concerns Tennis has identified (2016), they represent viable new lines of research with implications for concepts of aboutness and meaning within the LIS context. This would be a nod to studying the semantics and the pragmatics (in the linguistic sense) of terms alongside their role in indexing and in warrant. Analysis of the "code" of indexing languages in KO could thus be substantially supplemented by examinations of its emergent "expression" within works and records at scale. We will elaborate on this possibility below.

3.0 Hauser's episemantics: posterior projection of meaning

> Losee's conception of aboutness's role arises from a category error: while processes' output is related to both their input and the processes themselves as he claims, that relationship should not be described as aboutness until episemantic interpretation occurs. Following logical empiricism, Losee assumes that episemantic interpretation is (or: can be; should be; for science, must be) a transparent process, enabling processes' outputs to be about their inputs. I contend that aboutness only obtains in the relationship between the interpretation process and the jussive encoding process.
>
> Hauser, "Information from Jussive Processes," 303

Influenced by both the pragmatic philosophy of language and its continental critics, especially Derrida, Hauser emphasizes the lack of meaning inherent to inscriptions. For Hauser, this is encapsulated in Bowker's discussion of the jussive. Bowker views "memory practices" in light of the way in which they enact forgetting (Bowker 2006; Hauser 2018b). For Hauser, this amounts to a proposal to investigate technologies of remembering via the techniques of forgetting they enable.

These observations were sparked by a critique of Losee, who seeks to embed an informative aboutness into a domain-independent account of information (Losee 1997, 2012). Losee renders information as the result of pro-

https://doi.org/10.5771/0943-7444-2019-8-590 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

cesses and as informative "about" the process and its inputs. This is a powerful approach but problematically embeds a strong representational aboutness within the foundation of information. While scientific realists are likely to see no problems with such an arrangement, Hauser seeks to preserve the power and expansive domain of Losee's work while stripping it of its reliance on scientific realism. Scientific realism is incompatible with many of the domains we serve, so Hauser tries to preserve as a possible viewpoint while avoiding placing it at the core of our discipline.

Contra Losee, Hauser locates aboutness as subsequent to the interpretation of inscriptions rather than as inherent to processes. Episemantics is thus the posterior projection of meaning (and aboutness) onto inscriptions via interpretation. Meaning is always enacted rather than inherent. This includes both the meaning of information resources and of indexing languages. To revise Losee's formulation, information is merely subsequent-to processes; aboutness comes afterwards according to Hauser (2018a, 304): "The aboutness relationship consists of and is created by the episemantics of interpretation." Aboutness is thus not a property but a relation that arises out of interpretive acts.

Though it is inherently constructivist, Hauser takes pains to situate scientific realism within this conception. In Hauser's reading, Losee's information from processes and its embedded aboutness results from a specific account of the process of interpretation. "Following logical empiricism, Losee assumes that episemantic interpretation is (or: can be; should be; for science, must be) a transparent process, enabling processes' outputs to be about their inputs" (Hauser 2018a, 303). This framing doesn't exclude strong representationalist conceptions of aboutness but rather de-centers them. They are one amongst many potential instances of the creation of meaning. It is this de-centering which accomplishes Hauser's pluralistic goal. As, for better or worse, a metadiscipline (Bates 1999), we must serve a variety of fields and individuals, who make meaning in disparate ways. By identifying these interpretive processes, and cognizing our own, we can better align our activities with the needs of those we serve.

Losee's aims, and consequently Hauser's critique, are of course far broader than classification theory. So how does this work relate to KO? Tennis's exploratory proposal for computational analysis of indexing determinations, supplemented with corpus linguistics becomes an important way of investigating the enactment of meaning that Hauser situates at the core of information. Understanding the "posterior projection of meaning" in subject ontogeny thus becomes a project of uncovering evidence of such projections through the analysis of cataloging languages within the corpora that surrounded them.

4.0 Episemantics, recombined

Each conception could stand on its own, but we've found it generative to consider how the two conceptions might be recombined. Methodology and theory should ideally reinforce each other's strengths to form a coherent whole. Can such a project be accomplished here?

Tennis's account is much more deeply embedded within the methodology of classification research, especially subject ontogeny. This depth makes it clear how it might be applied, but obscures the true power and breadth of the idea. Hauser's approach is more general. This generality offers greater breadth but is ultimately diffuse and difficult to apply. This section will show how the two approaches can be combined to maximize their strengths and mitigate each other's weaknesses.

Tennis's analogy to epigenetics is apt, and a closer look at the field of epigenetics offers an important template for how KO might evolve like traditional genetics when confronting these ideas. Traditional genetics might simplistically be thought of as a series of sophisticated rules for labeling organisms and groups of organisms. Medical genetics uses the possession of genes as, effectively, a categorization rule to inform statistical analyses of morbidity and mortality (e.g., patients with this gene are X% more likely to develop heart disease, and live, on average, Y years less than those without). Phylogenetics uses algorithmic measures of similarity to infer ancestral relationships between species. Each of these approaches contains a step when the object of study is simply labeled genetically, and from this point on the label is all that is available. This labeling process is jussive, in Bowker's sense, and encodes a specific disciplinary technique of forgetting.

Epigenetics represents a deepened interpretation of DNA sequences by bringing their expression into view. Traditional genetics was presumed to be a method for finding the animating code behind everything but at times has devolved into a sophisticated mechanism for tagging data prior to statistical analyses of co-occurrence patterns. Epigenetics has a claim to this original promise, but must do so by abandoning a view of genetic sequences as determining the futures of the organisms that possess them in favor of a more fully contextualized account of how those genes proliferate and are expressed within an organismal and ecological context. Wendy Chun has noted the logocentrism common to biology and computing technologies (Chun 2013). She makes the novel, but convincing, claim that the kind of logodeterminism represented in works like Schrodinger's "What is Life" was an important precursor to our understanding of what code is and how computers work (Chun 2013, Ch. 3). Chun's analysis suggests a new light within which to view Tennis's analogy: that episemantics might offer a path, parallel to that of epigenet-

https://doi.org/10.5771/0943-7444-2019-8-590 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

ics, whereby we gain a greater account of context, and greater explanatory power, by abandoning an outmoded logodeterminism.

Such a potential approach is offered by a pragmatic account of aboutness as aroundness, or "in the vicinity, nearby." The occurrence of a gene within a strand of DNA is irrelevant unless the gene is expressed. The import of a gene remains unknown precisely "until" we have an account of its expression within an organism, population, or ecological context. Genetic expression is a process of interpretation within a context. Similarly, the possession of a term within an indexing language, or applied to a specific work, is irrelevant until we know how such a term is used. The meaning of a term is impossible to analyze prior to a contextualized account of use.

Thus, pragmatism forms a bridge between Tennis and Hauser's accounts of episemantics. Pragmatism implicitly animates a good deal of LIS work, and has recently gained traction as a subject of research in its own right (Dousa 2009; Buschman 2017; Sundin and Johannisson 2005). While competing accounts of pragmatism have been offered, we prefer Star's simple and concise definition: a focus on "consequences, not antecedents" (Star 2015, 133). Star here references the words of her mentor Anselm Strauss, who in turn was inspired by the work of John Dewey. This pragmatic ethos unites all three thinkers, even as the meaning of this mantra has evolved. Bowker and Star's *Sorting Things Out* would have been far less impactful for our field if it had been subtitled *Classification and its Antecedents*. A focus on consequences animates both Hauser and Tennis's approaches. Tennis's epigenetics analogy shifts focus away from the antecedent, DNA-like indexing language to the consequent, RNA-like classification records and the content of the works they classify. Hauser positions the antecedent inputs of an informative process as ultimately irrelevant to the aboutness of the consequent interpretation and enactment of output.

5.0 Why does KO need an account of episemantics?

Episemantics represents an important reminder to avoid viewing meaning as an inherent property of either indexing terms or abstract concepts. This offers the key methodological benefit of a shared account of both natural and artificial languages in a way that concepts like literary warrant cannot. The materiality of language emphasized by Hauser acts to blur the distinction between natural language, indexing languages, and computer languages. This, combined with Tennis's proposal to look for traces of use within all three kinds of languages, presents a new picture of what classification research might become. In addition to strengthening existing techniques such as subject ontogeny, our recombined concept of episemantics offers a glimpse of what larger scale, comparative "subject phylogeny" might be.

If we take episemantics seriously, we must revise our conception of aboutness. The notion of meaning somehow inhereing in the inscriptions that constitute a language (what Star might call an "antecedent" view of meaning) has proven philosophically problematic for human languages. Given this difficulty, we suggest abandoning an attempt to clarify or utilize this traditional sense of aboutness for indexing and computer languages. Instead, a turn to pragmatism about meaning and a focus on investigating use, both within narrow contexts and at scale, offers a viable way forward. "Aboutness" in this view need play no larger role than suggesting that something has been placed near something else, as librarians commonly do with cataloged books. The effects of cataloging may be deeply complex, socially embedded, and ethically significant, but the analysis need not include a strong account of aboutness as inherent meaning. Rather, we argue, an episemantic approach precludes this.

Our proposal does not seek to or need to enforce a uniform account of aboutness to succeed. Researchers who still believe that a strong account of inherent meaning is possible may continue to pursue work in that direction separately. To move forward, we need only agree to proceed with a weak aboutness within the empirically and historically oriented study of classification. When we do, Tennis's proposal of exactly how this might be studied at scale, for both subject ontogeny and the as-yet-unrealized field subject phylogeny, becomes merely a promising suggestion of many potential ways forward.

6.0 Aroundness, Dewey, and the digital

Although his classification system is often conflated with universalist classification projects, Melvil Dewey himself never considered the "aboutness" of his original classification system to be a specification of the property of the works cataloged and arranged on shelves. In the preface to the first edition of his classification, it is clear that his focus was primarily on the "effects of placing books near each other" (Dewey [1876] 1976):

> In all the work, philosophical theory and accuracy have been made to yield to practical usefulness. The impossibility of making a satisfactory classification of all knowledge as preserved in books, has been appreciated from the first, and nothing of the kind attempted. Theoretical harmony and exactness has been repeatedly sacrificed to the practical requirements of the library or to the convenience of the department in the college.

https://doi.org/10.5771/0943-7444-2019-8-590 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

The effects Dewey considered, of course, were both upon patrons when browsing the shelves, and upon the operation and maintenance of the library itself. The fact that subsequent versions of his system and its presentation became indelibly associated with universalist classification schemes need not prevent us from returning to it for inspiration. Dewey's system, embedded as it was within the late 19th century library movement's goals and cultural assumptions (for more on this, see Miksa 1998), was nevertheless a novel and pragmatic take on how to organize a newly abundant information resource for optimal use and management. Viewed in this lens, Dewey's principle was to identify a physical property of information resources, their physical location, and produce a system for manipulating this property to balance the needs of library patrons and library staff. Though this system contained subject headings, these were merely cogs in an ultimately spatial machine. In our terms, this machine manipulated aroundness rather than ascribing aboutness.

Recapitulating this approach with digital resources is non-trivial. Unlike a physical library, the interfaces, sequences, and formats that users access digital information are wildly disparate. To give a simple example, library patrons walk through the front door. Taking this into account, libraries could arrange resources in such a way as to reliably shape these first interactions. Though digital libraries still have putative "home" pages, users may land upon practically any part of the site, from practically any other digital context. What can serve the function that physical proximity did in Dewey's original system?

This, of course, is a question with proliferating answers. In a sense, the intractability of organizing the massive amounts of highly specialized knowledge, a task increasingly confronted by Dewey's successors, encourages the essentialist approach to "aboutness" that we have critiqued. For a specialist researcher seeking journal articles in her speciality, a given resource is either "about" "the desulfurization of hot coal gas with regenerable metal" or not. As Miksa notes, classification theorists who took up the devilish challenge of organizing specialist knowledge, such as Richardson, Bliss, and Rangagnathan, found themselves increasingly drawn to map a "universe of knowledge," where every specialist query could have a definite home (Miksa 1998, 56–73 et seq.). Through the lens with which we have been reading Dewey's work, this strikes us as precisely an attempt to provide an analogy to the physical location that made Dewey's system work for generalist libraries. A conceptual location within the Cartesian space of the universe of knowledge would, modernist classification theory held, allow the precise provision of the right resource for any sufficiently specified need.

The task of repeating this process without universalization and its attendant definite aboutness is one we suggest as a future research program. Methodologically, Tennis's proposal of utilizing large scale computational linguistics as a kind of window into the use and relationship of words to each other in a corpus would help ground such a project in the actual use of language rather than encouraging the invention and perfection of a crystalline representation of the universe of knowledge. Hauser's exhortation to remove meaning from classification helps us uncover the practical effects of classification activities. Dewey's pragmatism led him to focus on the physical arrangement of books. Subsequent modernists sought an ideal, universal space within which to arrange and relate classes to each other. The fragmented space of new digital technologies belies either approach. Knowledge is not a set of cartesian coordinates, waiting to be arrayed in crystalline perfection. There is no reliable experience of physical space to structure patrons' encounter with digital resources. How might we re-envision these organization practices to instead modulate properties that acknowledge the fractured nature of digital encounters but provide flexible structure for navigation and exploitation of digital resources?

7.0 Conclusion

In two separate threads, Tennis and Hauser point to a contingent and pragmatic view of aboutness. This leads us to reconsider the concept in terms of an earlier meaning, "in the vicinity, nearby; in some place or various places nearby; all over a surface." The vicinity and surface of meaning, we have argued, are epistemantically derived—both theoretically and methodologically.

Revisiting the early work of Dewey, we uncovered a new sense of aroundness, a literal one. Physical location was central to Dewey's scheme to balance the needs of patrons and library staff. Modernist classification theorists, who Miksa read as constructing "the universe of knowledge" as their domain, still employed an attenuated aroundness in their schemes relating classes, and thereby subsequently catalogued resources, to each other via their physical proximity within collections. Dewey's pragmatism centered around the realization that physical location was the primary "outcome" of his classification and the primary tool he had to influence library operations.

In a digital environment, many possible operationalizations of aroundness are possible. Commercial information systems have pioneered many of these, driven by large scale collection of user data ("Customers who viewed this also viewed"). The synthesized conception of episemantics advanced in this paper is intended to support deep engagements with these new possibilities. We hope that a pragmatic analysis of the consequences of different manifestations of aroundness might help provide guidance for continued innovation in KO.

https://doi.org/10.5771/0943-7444-2019-8-590 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

And, of course, episemantics remains an exciting methodological proposal for subject ontogeny research. Hauser's theoretical contributions are consonant with the goals of subject ontogeny. If meaning is viewed as enacted interpretation, examining traces of such interpretation in corpora and even individual cataloging decisions helps provide insights unavailable by any other means. Amongst its other goals, we hope that this paper encourages the large-scale collaborations needed to understand the complexity of semantic construction that animates KO activities, now and in the past.

References

Bates, Marcia J. 1999. "The Invisible Substrate of Information Science." *Journal of the American Society for Information Science* 50: 1043-50.

Bowker, Geoffrey C. 2006. *Memory Practices in the Sciences*. Cambridge, MA: MIT Press.

Buschman, John. 2017. "Once More unto the Breach: 'Overcoming Epistemology' and Librarianship's de Facto Deweyan Pragmatism." *Journal of Documentation* 73, no. 2: 210-23.

Chun, Wendy Hui Kyong. 2013. *Programmed Visions: Software and Memory*. Cambridge, MA: MIT Press.

Dewey, Melvil. (1876) 1976. *A Classification and Subject Index for Cataloguing and Arranging the Books and Pamphlets of a Library*. Lake Placid, NY: Forest Press Division, Lake Placid Educational Foundation.

Dousa, Thomas M. 2009. "Classical Pragmatism and Its Varieties: On a Pluriform Metatheoretical Perspective for Knowledge Organization." In *Proceedings from North American Symposium on Knowledge Organization. Vol. 2*, 1-9.

Hauser, Elliott. 2018a. "Information from Jussive Processes." In *Challenges and Opportunities for Knowledge Organization in the Digital Age: Proceedings of the Fifteenth International ISKO Conference*, ed. Fernanda Ribiero and Maria Elisa Cerveira. Advances in Knowledge Organization 16. Würzburg: Ergon, 300-7.

Hauser, Elliott. 2018b. "UNIX Time, UTC, and Datetime: Jussivity, Prolepsis, and Incorrigibility in Modern Timekeeping." In *Proceedings of the ASIST Annual Meeting*, 171-80. Hoboken, NJ: Association for Information Science and Technology.

Library of Congress. 2019. S.v., "Information Organization (Authority Record)." sh 99001059.

Losee, Robert M. 1997. "A Discipline Independent Definition of Information." *Journal of the American Society for Information Science* 48: 254-70.

Losee, Robert M. 2012. *Information from Processes: About the Nature of Information Creation, Use, and Representation*. Berlin: Springer.

Miksa, Francis L. 1998. *The DDC, the Universe of Knowledge, and the Post-Modern Library*. Albany, NY: Forest Press.

Soergel, Dagobert. 1974. *Indexing Languages and Thesauri: Construction and Maintenance*. Los Angeles: Melville.

Star, Susan Leigh. 2015. "Living Grounded Theory: Cognitive and Emotional Forms of Pragmatism." In *Boundary Objects and Beyond: Working with Leigh Star*, ed. Geoffrey C. Bowker, Stefan Timmermans, Adele E. Clarke, and Ellen Balka, 121-42. Cambridge: MIT Press.

Sundin, Olof and Jenny Johannisson. 2005. "Pragmatism, Neo-Pragmatism and Sociocultural Theory: Communicative Participation as a Perspective in LIS." *Journal of Documentation* 61: 23-43.

Tennis, Joseph T. 2012. "Emerging Concepts in Ontogenic Analysis." Paper presented at the Research Fair, University of Washington School of Information. https://digital.lib.washington.edu/researchworks/bitstream/handle/1773/37974/Tennis2012ResearchFairOntogenic.pdf

Tennis, Joseph T. 2016. "Methodological Challenges in Scheme Versioning and Subject Ontogeny Research." *Knowledge Organization* 43: 573-80.

https://doi.org/10.5771/0943-7444-2019-8-590 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY