

Supporting Collaborative Information Seeking Through Shared Annotations

Samuel Dodson
School of Library, Archival and Information Studies
University of British Columbia
Vancouver, BC Canada
dodsons@mail.ubc.ca

This paper presents an argument that information systems may support deeper engagement between readers and the text by realizing the capacity for readers to take more active roles while seeking information. Information systems could support collaborative information seeking by displaying shared annotations. Subsequent readers would be able to leverage the semantic information encoded within the shared annotations to interact with the text in more efficient and effective ways. Different types of collaboration and contexts in which it occurs will likely influence the ways in which shared annotations are used by readers and information systems to support collaborative information seeking. Directions for exploring the applications of shared annotations are discussed, concluding with four research questions for future work.

Annotation. Collaborative Information Seeking. Reading. User Generated Content.

1. INTRODUCTION

An understudied area of research within the information access community is information interaction, in which the reader's engagement with, and use of, the text is of primary focus. User generated content is created from this interaction, such as comments, reviews, and tags, which can then be used by the reader and subsequent readers. This content can also be used by information systems to determine the relevance of search results. There is an opportunity for information systems to support collaborative information seeking by leveraging the user generated content produced throughout the reading experience. By collaborative information seeking, I refer to the ways in which an individual reader's information interaction can inform subsequent readers' interactions with the text.

Annotations—i.e., highlights, notes, and other content that readers add to the text—are especially useful in supporting the experience of reading. When shared, annotations may be used by subsequent readers and information systems. Digital reading environments, such as the Amazon Kindle platform, now display *Popular Highlights*, which are passages that have been highlighted by multiple readers. *Popular Highlights* use the consensus of many readers to signal important passages. The presence of consensus highlights,

for example, allows subsequent readers to jump from one highlighted passage to the next. By enabling the creation and sharing of annotations, information systems can support readers as they make sense of information individually and collaboratively (Cabanac, Chevalier, Chrismont, & Julien, 2007). Sense-making, in this context, is the process by which people give meaning to the reading experience.

In this paper, I argue that information systems may support deeper engagement between readers and the text by realizing the capacity for readers to take more active roles while seeking information. The transactional model, which represents information as an object, or commodity, to be transferred from the system to the reader, treats individuals only as consumers. This model fails to recognize that individuals have the potential of taking more active roles as readers *and* writers of information. By not providing readers with the tools to add content to the text, information systems miss an opportunity to support readers' information interactions, and to collect user generated content.

2. ANNOTATION

To support greater levels of engagement between readers and the text, information systems could be designed for active reading. Tashman & Edwards (2011a) define active reading as "reading activities

that involve more interaction with the reading media than simple sequential advancement through the text” (p. 2928). While reading actively, people continuously evaluate the meaning of the text and its use for their task. This often requires readers to move through the text non-linearly to compare and contrast passages and make sense of the information within the text. Many systems have been designed to support active reading, such as PapierCraft (Liao, Guimbretière, & Hinckley, 2005), LiquidText (Tashman & Edwards, 2011b), and RichReview (Yoon, Chen, Guimbretière, & Sellen, 2014), but active reading has received less attention within the information access community.

A common way in which readers become, and remain, deeply engaged while reading is by annotating. This is also a way in which people transition from readers to writers. The annotations may be structured, such as ratings of the text, or unstructured, such as free-form doodles and notes in the margin. Annotations, such as comments, tags, and up or down votes, are commonly used in social media platforms. Whatever their form, annotations are an important part of the reading experience. O'Hara (1996) stated, “[annotations] are not just a tag on to the reading process but an integral part of it, changing the way in which reading occurs and the way information is processed” (p. 6).

2.1 Annotation as Process and Thing

Annotating is both a process and an outcome of the reading experience, a distinction that MacMullen (2005) calls annotation-as-process and annotation-as-thing. The process of annotating may only serve people as they read, but the outcome of annotating can support subsequent readers and information systems. Annotations, as objects, may be useful to subsequent readers, because they reflect the cognitive work that goes into reading and understanding the text. By sharing annotations, for example, subsequent readers could learn about others' interpretations of the text by reading through notes written in the margin. I suspect that annotation has received limited attention within the information access community because it is primarily considered a process rather than as content in its own right.

The activity of annotating serves several purposes for readers. Marshall (1997) identified six functions of highlights, notes, and other annotations. Of these functions, three could be useful to subsequent readers and information systems: 1) signals for passages to be read or referenced later, 2) records of sense-making, and 3) traces of readers' attention. These functions allow readers to engage with the text by assessing the semantic cues left behind by others (Agosti & Ferro, 2005; Kopak, Freund, & O'Brien, 2011). For example, subsequent readers may assess the relevance of the text to their

information need by jumping from one highlight to the next.

3. COLLABORATIVE INFORMATION SEEKING

Reading is often thought of as a solitary activity, although there have been efforts to understand the social dimension of the reading experience (e.g., Stein, 2010). The ability to create and share annotations with others makes the experience of reading more social. The text is a social space where questions and understandings can be exchanged. In today's digital reading environments, reading need not be considered a solitary activity, but rather a social one. With social spaces for reading now possible, the information access community can begin to explore collaborative information seeking between readers.

Golovchinsky, Pickens, & Back (2008) identified four dimensions of collaboration, which can be used to investigate the contexts in which collaboration can occur and ways in which annotations can be created, shared, and used.

3.1 Intent

The information seeking task, which Golovchinsky, Pickens, & Back (2008) refer to as intent, is either explicit or implicit. Explicit collaboration involves groups of readers in which every reader works towards an information need negotiated by the group. Implicit collaboration, on the other hand, involves groups of readers who use one another's contributions without an understanding of their intent. An example of implicit collaboration is the serendipitous experience of encountering pre-existing annotations in a used book.

3.2 Depth of Mediation

The depth of mediation refers to the visibility of user generated content within the digital reading environment. If user generated content is visible, readers have direct access to shared annotations either in their raw or processed form (see section 4.3 for further discussion of the visualization of shared annotations). Shared annotations could also be used by information systems out of public view. For example, readers may not be aware of the ways in which user generated content determines the rank of their search results, and in turn influences their information interaction.

3.3 Concurrency

Readers of the same text can either collaborate synchronously or asynchronously. In asynchronous collaboration, readers do not work at the same time. Those who read the text later can benefit from the work of earlier readers, though earlier readers do not necessarily benefit from the work of subsequent

readers. Synchronous collaboration, however, involves reading of the text at roughly the same time. Synchronous collaboration does not necessarily imply that readers intend to work towards a shared information seeking task (i.e., readers do not share an explicit intent to aid one another). For example, many readers can add content to the text at the same time, but the individual readers are not necessarily working towards a shared goal.

3.4 Location

Location refers to the physical location of the reader in relation to other readers of the same text. Thus, readers can be co-located or distributed. Co-located reading is more conducive to explicit, synchronous collaboration than distributed reading. For example, reading groups often meet physically to share their questions and understandings about the text.

4. FUTURE WORK

In the face of information overload, collaborative information seeking is a way to distribute the cognitive load required to make sense of information. Collaborative information seeking is a promising area of research; it is a way in which information systems can leverage the potential for readers to assume more active roles in interpreting the meaning of the text and its application to their information need. Collaborative information seeking can be supported by shared annotations. I am also exploring the relationship between active reading—including its processes and outcomes—and collaborative information seeking. Active reading, including shared annotation, is an understudied area of research within the information access community. I see at least four open research questions within collaborative information seeking that I hope to investigate in my PhD studies.

4.1 Reading Processes and Outcomes

RQ1: How can the effects of shared annotations on collaborative information seeking be measured?

There are many different outcomes of collaborative information seeking, including learning. The searching as learning framework, which has emerged in part from the Searching as Learning workshops at IliX and CHIIR, may be appropriate to use. This framework measures learning through the assessment of reading processes, outputs, and other information behaviours (Freund, Dodson, & Kopak, 2016). Other outcomes of collaborative information seeking include engagement, participation, and a sense of community.

4.2 Reading Behaviour

RQ2: How do readers interact with and assess shared annotations?

Some work has applied information foraging theory (Pirulli & Card, 1995) to shared annotations. Chi, Gumbrecht, & Hong (2007), for example, found that highlights attracted readers' attention. In their study, readers completed an information seeking task faster and more accurately using shared annotations than just the text. More research is needed to explore how other types of shared annotations affect the experience of reading.

Furthermore, while some work has investigated readers' perceptions of shared annotations (e.g., Hemminger & TerMaat, 2014), it is not clear how readers assess the quality of shared annotations. What attributes of shared annotations, for example, are most important to readers when they assess quality? By exploring which attributes are most important to readers, information systems can present these cues to readers to facilitate their evaluation of shared annotations. This information can also be used by information systems to retrieve shared annotations that are relevant to readers' information tasks (Kopak et al., 2011).

4.3 Information Retrieval

RQ3: How might shared annotations be retrieved and presented to facilitate collaborative information seeking?

A challenge in designing information systems for collaborative information seeking is the potential for interrupting the reading experience. This challenge can be addressed by information systems and readers. First, information systems could retrieve shared annotations that are relevant to readers' information needs. The retrieved shared annotations could be aggregated and displayed using a visualization to suit readers' preferences, such as a heat map or tag cloud (e.g., Shipman, Price, Marshall, & Golovchinsky, 2003; Tashman & Edwards, 2011a). The annotations could also be used to retrieve additional information (e.g., Price, Golovchinsky, & Schilit, 1998).

Second, readers could assess the relevance and usefulness of shared annotations. Previous work suggests that readers would like the ability to search within shared annotations to manage information overload, which is caused by accessing too much user generated content (Marshall & Brush, 2004; Tashman & Edwards, 2011a).

4.4 Collaboration

RQ4: How do the dimensions of collaboration discussed in section 3 affect readers' collaborative information seeking?

Previous work by Marshall & Brush (2004) suggests that readers annotate differently when sharing their annotations with others. How does, for example, depth, location, and concurrency affect collaboration?

5. CONCLUSION

Information systems could be improved by supporting active reading and collaborative information seeking. Annotating is a common process and output of active reading. The ability to share annotations with others makes the experience of digital reading more social and allows for collaboration between readers. Shared annotations can be used by readers to efficiently and effectively seek information. Information systems could also leverage user generated content to determine the relevance of information within the text. Collaborative information seeking is a promising area for future research with open research questions that have practical and theoretical implications.

6. REFERENCES

- Agosti, M., & Ferro, N. (2005). Annotations as context for searching documents. In *Context: Nature, Impact, and Role – 5th International Conference on Conceptions of Library and Information Sciences* (pp. 155–170).
- Cabanac, G., Chevalier, M., Chrisment, C., & Julien, C. (2007). Collective annotation: Perspectives for information retrieval improvement. In *Large Scale Semantic Access to Content* (pp. 529–548). Paris, France: Le Centre de Hautes Etudes Internationale D'Informatique Documentaire.
- Chi, E. H., Gumbrecht, M., & Hong, L. (2007). Visual foraging of highlighted text: An eye-tracking study. In *Human-computer interaction: HCI intelligent multimodal interaction environments* (pp. 589–598). Berlin, Germany: Springer.
- Freund, L., Dodson, S., & Kopak, R. (2016). On measuring learning in search: A position paper. In *Proceedings of the Second International Workshop on Search as Learning*.
- Golovchinsky, G., Pickens, J., & Back, M. (2008). A taxonomy of collaboration in online information seeking. In *1st International Workshop on Collaborative Information Retrieval* (pp. 1–3).
- Hemminger, B. M., & TerMaat, J. (2014). Annotating for the world: Attitudes toward sharing scholarly annotations. *Journal of the Association for Information Science and Technology*, 65(11), 2278–2292.
- Kopak, R., Freund, L., & O'Brien, H. L. (2011). Digital information interaction as semantic navigation. In A. Foster (Ed.), *Innovations in information retrieval: Perspectives for theory and practice*. London, England: Facet Publishing.
- Liao, C., Guimbretière, F., & Hinckley, K. (2005). PapierCraft: A command system for interactive paper. In *Proceedings of the 18th Annual ACM Symposium on User Interface Software and Technology* (pp. 241–244). New York, NY: ACM.
- MacMullen, W. J. (2005). Annotation as process, thing, and knowledge: Multi-domain studies of structured data annotation. Chapel Hill, NC: School of Information; Library Science, University of North Carolina at Chapel Hill.
- Marshall, C. C. (1997). Annotation: From paper books to the digital library. In *Proceedings of the Second ACM International Conference on Digital Libraries* (pp. 131–140). New York, NY: ACM.
- Marshall, C. C., & 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* (pp. 349–357). New York, NY: ACM.
- O'Hara, K. (1996). Towards a typology of reading goals. Cambridge, United Kingdom: Rank Xerox Research Centre.
- Pirolli, P., & Card, S. (1995). Information foraging in information access environments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 51–58). New York, NY: ACM.
- Price, M. N., Golovchinsky, G., & Schilit, B. N. (1998). Linking by inking: Trailblazing in a paper-like hypertext. In *Proceedings of the Ninth ACM Conference on Hypertext and Hypermedia* (pp. 30–39). New York, NY: ACM.
- Shipman, F., Price, M., Marshall, C. C., & Golovchinsky, G. (2003). Identifying useful passages in documents based on annotation patterns. In *Research and advanced technology for digital libraries* (pp. 101–112). Berlin, Germany: Springer.
- Stein, B. (2010). A taxonomy of social reading: A proposal. Retrieved from <http://futureofthebook.org/social-reading/>
- Tashman, C., & Edwards, W. K. (2011a). Active reading and its discontents: The situations, problems and ideas of readers. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2927–2936). New York, NY: ACM.
- Tashman, C., & Edwards, W. K. (2011b). LiquidText: A flexible, multitouch environment to support active reading. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 3285–3294). New York, NY: ACM.
- Yoon, D., Chen, N., Guimbretière, F., & Sellen, A. (2014). RichReview: Blending ink, speech, and gesture to support collaborative document review. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology* (pp. 481–490). New York, NY: ACM.