

(Bee)Dancing on the Boundary between PIM and GIM

Shannon Bradshaw*, Marc Light†, and David Eichmann†
 Drew University* and The University of Iowa†

sbradsha@drew.edu, marc-light@uiowa.edu, david-eichmann@uiowa.edu

What?

In the life sciences, 40,000 new articles are published each month, adding a large number of findings to an already immense area of knowledge. Literature review is difficult, increasing the chance of duplicated effort. It is also time-consuming, reducing the amount of resources scientists have to pursue new discoveries.

Results so far:

- Our studies indicate that there exists a high degree of consensus among readers as to which passages are important in a paper and that these passages account for only one-third of the total text in a paper.
- An initial version of our system, BeeDance is in alpha.
- Our text extraction library for PDF is in beta and available for download.

Why?

One approach to this problem would be to fund a talented team of researchers to curate all published knowledge in a structured database. However, it is our position that the ordinary information gathering activity of scientists provides both an effective means of self-curation of published knowledge and an effective vehicle for communicating this information to researchers who will need it.

How?

An information system that enables researchers to:

- Read and annotate (e.g., highlight) electronic documents as easily as they do with paper.
- Organize documents, annotations, lab notes, and data for personal information management.
- Share these "knowledge artifacts" (KAs) with others in their research group and larger community.

