Visual Semantic Enrichment for eReading

Michael Coleman
Computer Science, University of Waikato
New Zealand
mjc62@students.waikato.ac.nz

Annika Hinze
Computer Science, University of Waikato
New Zealand
hinze@waikato.ac.nz

The current transition from physical to electronic books opens up opportunities to present semantic information about a book's content. This paper reports on a project that aims to improve the eBook reading experience by presenting semantic information about the eBook content together with the text. The key research challenges included suitable semantic information to be presented alongside a text and manners to display semantic information in an eReader application to supplement the text. We developed a web-based prototype that incorporates analysis of semantic content elements and explored the effectiveness of these semantic elements in a user study with 30 participants. Our results indicate that dynamic semantic visualisations may assist comprehension of existing concepts within an eBook, and provide insight to information that could not be easily gleaned by simply reading the text.

1. INTRODUCTION

Metadata about a text can provide useful information about a book's content, such as a character's attributes or the locations mentioned. Although eReaders can display dynamically created information, very few visual enhancements are implemented. Common are integrations with APIs such as Wikipedia (e.g., in Aldiko and Kindle). While semantic annotation of digital texts is an active field of research, visualisation of semantic information beyond textual form is rare. Our project aims at improving the digital reading experience with dynamic semantic enhancements by building upon existing semantic knowledge bases, semantic extraction methods and data visualisation tools. A prototype eReader was developed to provide semantic visualisations that may enhance the electronic reading experience. We used book-based ontologies and natural language parsing to create semantic information (such as proximity between actors in a book, and commonly used parts of speech around selected actors) and visualised these in a range of graphs. We conducted a user study with 30 participants to assess the potential of these semantic visualisations.

2. RELATED WORK

Enrichment in eBooks. Current eReading software incorporates a variety of features that extend beyond the reading experience of physical books. The Kindle Paperwhite is a prime example of the additional capabilities of eReaders. After selecting text in an eBook, users have access to dictionary, translation, and Wikipedia search for the highlighted word or phrase. It also provides one of the rare examples of some semantic enhancement.

Its x-ray tool displays available Amazon data about characters, locations or topics in the text. On tapping such a word, the x-ray feature provides an overview of quotes from the book regarding this character or location, including page references. However, the feature only works on books that have pre-loaded such information. The coverage one can expect from the feature is not disclosed by Amazon. Currently, the available semantic data is proprietary, limited in scope, and poorly presented. The Aldiko application for eBook reading on Android devices is an example of the current standard features one encounters in most modern eReading devices and software. It provides tools for highlighting text, creating and sharing notes, dictionary lookups, and text searches on the book digest. In general, semantic annotation is limited and the use of dynamic visualisations is virtually non-existent. Other examples of readers that support these standard features are Prestigio and Cool Reader. As usual, each of them also offers additional tools (such as text-to-speech in both systems, and a correction tool in Cool Reader). None of these provide semantic enhancements.

Text Visualisations. Sirkiä and Sorva's research on providing visualisation alongside (non-fiction) text revealed that, for an overwhelming majority of participants, such visualisations assisted the reader to comprehend the material (Sirkiä and Sorva 2015). Their study used graphical representations of the programming concepts alongside a programming text in an eBook. Student interaction with the eBook was monitored, and the students were also surveyed on the usefulness they attributed to the visualisations in regards to understanding the material. They found that a large number of participants often viewed visualisations not directly related to their given task, and

saw this as support for the graphical representation of the material rather than plain-text explanations. Their study used non-fiction texts while we focus on fiction eBooks. They also manually created the visualisations to assist the text, which is much easier than attempting to automatically generate them from segments within the text. Children's books traditionally contain more illustrations than non-fiction and adult books, and are well known to be effective in supporting a child's comprehension of a story (Sipe 1998; Fang 1996). Kay (1972) proposed the concept of a *Dynabook*, calling for future eBooks that would be more engaging and interactive. Colombo et al. explored this concept of more engaging children's eBooks with a cooperative inquiry study with children, leading to six design guidelines. Guideline 4 called for books being "not too difficult to read" and suggests inline dictionaries and what they call "illustrated descriptive cards" (illustrations and text in addition to the core text). Guideline 6 calls for the non-textual elements to "enhance, not replace text" Colombo et al. (2014). Visually illustrating textual information allows both children and adults to form a better understanding of concepts within the text, however, these should accompany but not replace the textual core.

3. SEMANTIC-ENHANCED EREADER

To evaluate the usefulness of dynamic semantic visualisations within eBooks, we developed an eReading prototype. We created a web application as to avoid any technical issues when reading on different mobile devices (e.g., iOS or Android tables). Algorithms that perform *perfect* automatic extraction of semantic data do not exist and reliable ontologies of semantic data cannot be built solely by using simple natural language processing (NLP) techniques. NLP techniques are often used to identify significant entities and their relationships within texts. The acquired semantic information must be used with caution as the presentation of incorrect or incomplete data will create a negative experience for users if relied on heavily. However, linked databases allow quick retrieval of this information once discovered. We therefore decided to combine an external book ontology and semantic tagging of the eBook's full-text. The results of these are then combined in order to ensure information quality.

3.1. Design and Implementation.

Figure 1 shows the architecture and data flow of the developed eReading application. After a user loads an eBook (in *ePub* format) into the web application (step 1), the book content is extracted as text (step 2) and presented in the eReading interface (in *html* format) to the user (step 3). The initial unannotated text display is similar to an eBook view in a standard eReader. To obtain semantic information, two independent processes are started. Firstly, we query an available book ontology for semantic information about the eBook (step 4), based on bibliographic metadata (e.g., book title, author, year). A
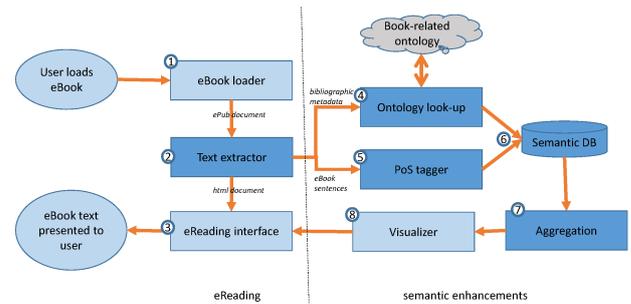

Figure 1: *eReader application architecture*

Book ontology may provide semantic information such as names of persons in a book, locations, and relationships. Secondly, we segment the text into sentences and run these one-by-one through a Part-of-Speech (PoS) tagger (step 5). The results of this process are semantic entities identified from the eBook text. Both the external ontology information as well as PoS data are stored in a server-side database (step 6) that allows the information to be retrieved quickly the next time the text is loaded. The only external queries necessary for the application are the ones accessing relevant external ontologies. Using and combining these two sources aims to ensure semantic quality. Only those semantic entities occurring in both result sets are used for visualisation. Once the appropriate semantic information has been acquired, the entity highlighting and all associated visualisations will become available to the user (step 7). A variety of web technologies were used to create the eReading application. For display and user interaction in the eBook loader component and reading interface, we used common web languages. These components run in the web application (indicated in light blue). These require a reasonably up-to-date-browser, but have no other system-specific requirements, which allows us to support a variety of devices. Interaction with the server is done through PHP code (for parsing eBook text and gathering of semantic information). The server components are shown in dark blue. All discovered semantic information is stored in a database located on the web server.

We trialled two book-related ontologies, Freebase (Bollacker et al. 2008) and Wikidata (Vrandečić and Krötzsch 2014). Freebase provided rich content and book-related structures such as bibliographic metadata as well as semantic information regarding the content (e.g. lists of characters). It covered 1,874,006 books, but is now deprecated and its content being integrated into Wikidata. Wikidata currently holds a small set of properties for the book ontology. The book-related structures are rich (e.g., bibliographic metadata, characters, real-life analog of fiction characters, stories or events that inspired the book, etc.) However, currently only a small set of useful character information is supported. Once the Freebase data is integrated, the semantic source will be much richer. Based on bibliographic metadata of the eBook a user

selects, a list of characters in the eBook is retrieved from the Wikidata ontology. For Part-of-Speech tagging, the Stanford PoS Tagger (`nlp.stanford.edu/software/`) was used to extract semantic information about a text independently of available online resources. In this way, books that lack on-line repositories of information can still be semantically enhanced. For the graph visualisations, a variety of existing libraries for JavaScript (plot, flot and ChartJS) were explored. The visualisations are created with ChartJS, which allows the greatest flexibility as we gather increasingly complex semantic data and relationships. The interface of our eReading application uses the Bootstrap framework (`www.w3schools.com/bootstrap`) to assist in creating a mobile responsive design. This enables the application interface to render correctly on multiple screen sizes, by using appropriate bootstrap styling techniques within the application.

3.2. Interface and Interaction

Our exploratory approach based on the available semantic data tests the concept of semantic visualisations in eBooks and identifies promising directions for future explorations. The eReading interface (see Figure 2) was purposefully kept simple and minimalistic; it allows the user to read the eBook, using left-to-right flipping through pages similar to most eReaders. We use here the example of Carroll's *Alice in Wonderland* as sourced from Project Gutenberg.
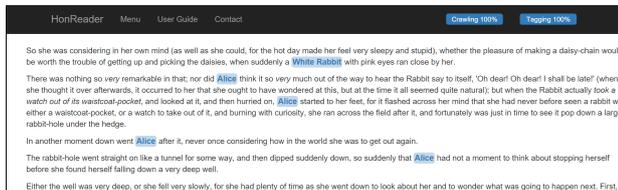
Figure 2: eReading screen (Alice in Wonderland)

Interaction with features of the semantic visualisation is through in-line semantic annotations of identified entities (blue highlights created through JavaScript callbacks) and an additional menu (for visualisations not related to a single specific entity). We also incorporated the Wikipedia integration prevalent in existing eReaders.

Character frequency. Main characters within a story are likely to be mentioned with higher frequency than minor characters. To give readers an impression of the importance of a character in the eBook, we present to the user a bar chart with the number of times each character appears. This data is aggregated from the output of the PoS tagging. For example, Alice was discovered approximately 400 times, while all other characters appear least than 100 times.

Character in eBook location. The frequency variation with which a character appears throughout the text is shown in a graph. The eBook was structured into units of text each approximately 5%, and the results of the PoS tagger are analysed while tracking progress through the
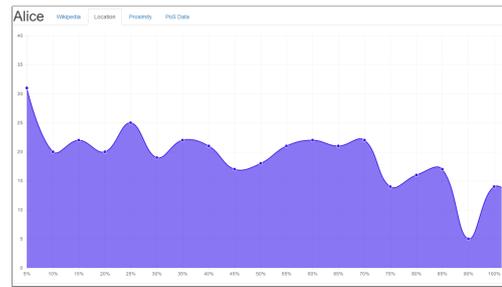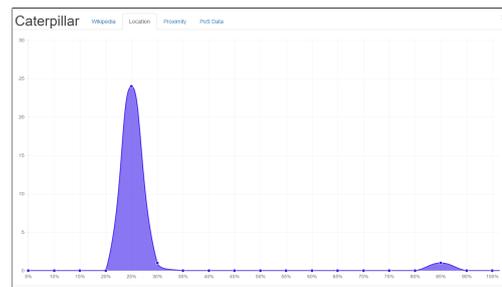
Figure 3: Locating Alice throughout the eBook

Figure 4: Locating Caterpillar throughout the eBook

eBook. Minor characters in *Alice in Wonderland* generally seem to have one or two peaks and were mentioned nowhere else (see Figure 4 for the Caterpillar), while the main character Alice appears throughout the book (see Figure 3).

Character Proximity. Two proximity graphs aim to visualise separateness and closeness of other characters in relation to the character that is being explored. The measure of *separateness* is based on the number of sentences that appear between the entities based on the sentence analysis that occurs before the PoS tagging (see Figure 5). The intention is to gain an impression of the relevance of the characters for each other. If an entity is never mentioned in close proximity to another we infer that the two entities never cross over within the text. For example, Alice is somewhat distant to the Queen of Hearts, Duck, and Eaglet. As a measure of *closeness*, we present the inverse sentence distance that other characters have from the selected character.

Verbs and Noun Clusters. The word clusters display all verbs and nouns, respectively, that are associated with the selected character. The size and colour of a word reflect the frequency with which that particular part of speech occurs in relation to the character. For example, Alice is mainly talking, but also thinking (see Figure 6). Interestingly, the most prominent nouns connected with Alice are 'time' and 'way', whereas for the Caterpillar it is, quite predictably, his 'Hookah pipe'. This visualisation is a result of the PoS classification given to words surrounding an entity. As this data is potentially prone to identification and classification errors, we do not use it for a very detailed analysis (e.g., close word by word comparison of proximity). This is an example of how
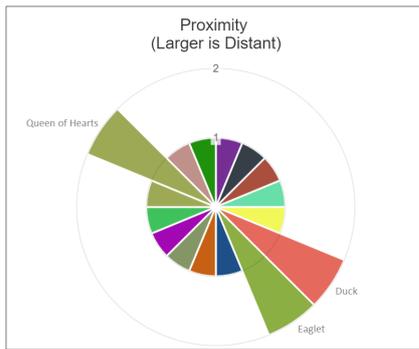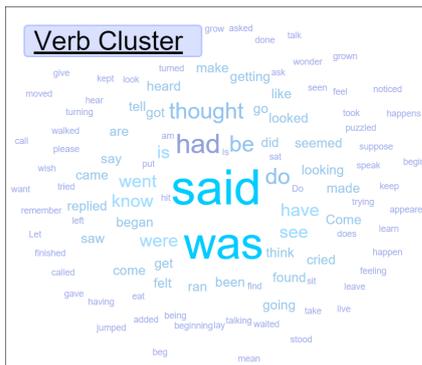
Figure 5: Separatedness (Alice)

Figure 6: Verb cluster (Alice)

Figure 7: Likert-scale and open Questions

fully automated natural language parsing can be used to gain insights without being heavily impacted by minor classification errors. The aim of this visualisation is to give an overview of the types of actions and and objects strongly associated with an entity in the story. It could be improved by more rigorous data post-processing (e.g., verb tense conversations) and by filtering more frequent words (e.g., "said" or "was") to reveal more nuanced data, or by grouping word types by chapter.

4. USER STUDY

Study Setup. To evaluate the effectiveness of the visualisations in expressing concepts about the story, we conducted a user study. The eReading application was loaded onto a tablet device with the eBook emphAlice's Adventures in Wonderland. Users were approached "in the wild" on campus. After obtaining some basic background information, participants were then introduced to the application. After a brief period of familiarisation, they were asked to find various information using the application, such as the most-mentioned character, the occupation of the White Rabbit and a common activity of the Duchess. This was to ensure the participants engaged not just with the eBook text but also the visualisations. After the participants had finished, ten questions were asked in the guided interview (see questions in Figure 7). All questions were asked by the researcher guided by the questionnaire, and answers

were recorded in audio and writing. In Questions 1–4, participants were asked to rate their experience on five-point Likert scales, with the aim of obtaining feedback on specific aspects from the participants. Questions 5–9 were designed to be open questions in which the participants were encouraged to give feedback about the application.

Participants were approached "in the wild" on the University campus. The study had 30 participants aged 18 to 37 (13 female and 17 male). Participants were predominately undergraduate and graduate students of various subjects from a variety of backgrounds. Participants were largely evenly divided between knowing the book *Alice in Wonderland* (13 report knowing the book 'well' or 'very well') or not being familiar with it (14 answering 'not at all' or 'very little') with 3 being undecided. All of the participants were familiar with reading on electronic devices, mostly for non-fiction study material. We did not ask specifically about their habits of eReading fiction.

4.1. Results from Likert-scale questions

All participants successfully located the requested information.

Enhanced Understanding. Participants were asked to rate how much they thought the graphs enhanced their understanding of the text on a scale of 1 (not at all) to 5 (a lot). The mean rating was 4.0 across the 30 participant; the mean rating was 4.0 across the 30 participants (1–5: 0, 1, 6, 16, 7). We note that no participant felt that graphs did not enhance their understanding of the text, and only a single participant gave a rating of 2 (P1). P1 was not interested in using such visualisations more as they "do not read fiction books" (more on this below).

New Information. Participants were asked if the graphs showed them new information about the story, on a scale of 1 (not at all) to 5 (a lot). We note that 24 of 30 participants gave ratings 4 or 5. Ratings 2 and 3 were selected 3 times each, and no participant felt the lowest rating (1) applied, with a mean rating of 4.1.

Entity Coverage. Participants were asked if they would be interested in information relating to other entities, specifically locations, objects and their relationships. This

question asks for interest in data that would be available when using a richer book ontology. We note that no participants selected either of the two lowest scores, score 3 only attracted 4 votes, while each of the two top scores were selected 13 times each. The mean rating was 4.3 across 30 participants. Participants were very enthusiastic about potential additional visualisations of semantic information; see discussion in Sections 4.2 & 5.

Overall Impression. Participants rated their overall impression on a scale of 1 (poor) to 5 (good). The feedback is very positive; the mean rating was 4.3 across 30 participants (3 for 3, 16 for 4 and 12 for 5).

4.2. Qualitative Results

As not all qualitative results can be listed here, we instead present common ideas and particular comments of note. Further results are available in (Coleman 2015).

Q5: What did you like about the software? Most answers were centered around the presentation of the information and the interaction when switching from eReading to semantic visualisations. 16 participants mentioned the availability and easy interpretation of the semantic data presented, and 14 participants felt the software was easy to use. Typical comments were:

- "Seems useful for research and assignments" [P2]
- "Easy to access information that would otherwise be hard to find." [P16]
- "Improvement from the software on my Kindle that only shows definitions / Wikipedia pages." [P19]

Q6: What would you change? Comments on this question involved HCI suggestions (12) and fixing of interface glitches (2), as well as suggestions for additional features (9). Comments regarding the user interface asked for changes to the way things are coloured, and the way one interacts with the book and the menus. Example comments are:

- "Ability to compare [to] other characters" [P1]
- "Include options to access character pages from info-graphics." [P7]
- "Maybe some focus on things not just main characters but a common or important keyword to eBook, e.g. the mission of Alice" [P11]

Q7: Would you be interested in this type of information when reading eBooks in the future? The 30 participants gave multiple suggestions. 25 people said they would be interested in the information in at least one specific circumstance, and 8 participants suggested its applications for text analysis purposes. The people who were not interested gave reasons such as a lack of interest in eBooks [P1], or being unable to see the use for fiction books [P7], and liking "the storyline" [P15].

- "Yes. Easier to navigate characters and what they do. Good for book reviews and analysis" [P5]
- "Not with fiction books, Doesn't provide more depth to story unless you are a 'fan' of the series" [P7]

- "Yes, often when reading books in large series eg wheel of time, sometimes it is good to refresh your memory / understanding of characters." [P9]
- "Yes I would be interested as [...] having access to all the extra information in/detail would make my reading experience more knowledgeable and enjoyable" [P18]

Q8: What else would you like to see to enhance the presentation of an eBook? Several participants did not request changes but gave positive comments ("Nothing - presentation is great!" [P6]); "Real book feel, pages, colour etc." [P10]) Seven participants requested HCI features similar to those they would expect from an established eReader product (feedback, colour schemes and page snapping). Requests for additional ways of presenting semantic data were made by 5 participants: quotes from the characters, previous sentences containing the characters, and an idea of the settings the characters are involved in.

Q9: Any other suggestions or comments? This question had a low response rate (17/30), and most responses were comments simply supporting the application ("Looks great" [P4], "Great tool" [P5]) , however six participants made suggestions for features and improvements within this section, like adding a word count number to each item in the word cloud and more links to other platforms. 4 participants made comments such as 'looks great', and 5 participants said they had no extra comments to make.

- "A way to return to a different place in the script from the pop-up segment. For use in manuals etc." [P9]
- "Concentrate on children's books." [P13]
- "Make specific stats (counts) for words available in the word clouds" [P30]

5. DISCUSSION AND CONCLUSION

The project reported in this paper aims at improving the eBook reading experience by presenting semantic information about the eBook content together with the text. The results from the study were generally supportive of semantic visualisations in eBooks, with many positive and encouraging comments. Here we discuss implications of our project work and possible future research avenues.

Semantic concepts. Most participants felt that the graphs helped them understand existing concepts within the story and also discover new information about the text. Participants felt the visualisations showed new information rather than merely supporting understanding of existing concepts. This could be attributed to the distribution of participants' preexisting knowledge of the text. In future, one may not always wish to reveal all the information about a character (e.g., in a mystery novel); hiding future character occurrences from visualisations may be desirable. Support for further semantic entities objects was also requested by participants. Results of our study support the findings of Sirkiä and Sorva (2015) about interest in visual representations of information.

Interaction and visualisation. Participants praised the quick and easy way of finding information and enjoyed the multiple visualisations. They wished for improvements in user experience ("more colour") or for the graphs to also link back into the story, allowing a more intuitive flow from the visualised information into the place it was retrieved from within the story. Improving the interactivity of the visualisations similar to the *Utopia* project (Pettifer et al. 2009) would better support the semantic nature of the data, and allow for links back to the text, as well as further manual annotation options for improving the accuracy of semantic data. Incorporating small visualisations embedded in the page margins, such as suggested by Tufte (2004), may avoid having to 'jump' between text and annotation.

Future Usage. There was an interesting division of opinions on the possible applications of visualisations in eBooks. Participants suggested many target audiences for the application when asked if they would personally be interested in using an application with semantic visualisations in the future, such as "[the application makes it] easier to navigate characters and what they do. Good for book reviews and analysis". The suggested applications covered academic texts, children's books, instructional manuals and fiction novels. The fiction novel support is not surprising as this is the type of text provided in the study, and likely the only experience participants had with semantic visualisations. The recommendations for other audiences (not directly asked in the question) is exciting, because it shows that, in general, almost every participant felt that semantic visualisations could enhance text in a variety of ways. Some participants felt they would not use visualisations in the first read they would be more interested in digesting the material, and did not want information overlaid on top of the text.

Further Semantic Information. In future, methods for acquiring semantic information may be refined to generate more possibilities for semantic enhancement. One may join multiple PoS tags to identify more complex concepts (Piryani et al. 2013). Semantic entity recognition techniques may provide another avenue for acquiring semantic data from eBook digest (Hinze et al. 2015). This would allow for easy tracking of characters who change names or have more than one name.

Fiction and non-fiction We here explored semantic visualisations for fiction eBooks but can see applicability to visualisations in non-fiction. Here the semantic links may even go beyond reference to static data, and provide live feeds and measurements (e.g., market data).

Embedded vs complementary enrichment. At current, none of the semantic enrichment can be encoded into the prevailing ePub and other similar file types. To achieve native embedded support for semantic enhancements, these would have to be expanded to embrace additional capabilities of the medium of eReader and tablet devices.

Native support would be particularly suitable for data that is directly related to, and generated from, the book. On the other hand, the possibilities of semantic enrichment for eBooks only begun to be explored. Research on manual annotation (Hinze et al. 2012) has highlighted the wide range of semantic annotations readers may be interested in, including suitability for age groups, for machine processing and analysis, cross-references with other books, comments on language, culture, citations, etc. To embrace this kaleidoscope of opportunities, external semantic annotation is more suitable.

REFERENCES

Bollacker, K., Evans, C., Paritosh, P., Sturge, T., and Taylor, J. (2008). Freebase: a collaboratively created graph database for structuring human knowledge. In *2008 ACM SIGMOD*, pages 1247–1250. ACM.

Coleman, M. (2015). *Visual Contextual Semantic Enhancement in eBooks*. University of Waikato. Honours thesis.

Colombo, L., Landoni, M., and Rubegni, E. (2014). Design guidelines for more engaging electronic books: Insights from a cooperative inquiry study. In *Conf. on Interaction Design and Children (IDC)*, pages 281–284. ACM.

Fang, Z. (1996). Illustrations, text, and the child reader: What are pictures in children's storybooks for? *Reading Horizons*, 37(2):3.

Hinze, A., Heese, R., Luczak-Rösch, M., and Paschke, A. (2012). Semantic enrichment by non-experts: usability of manual annotation tools. In *ISWC 2012*, pages 165–181.

Hinze, A., Taube-Schock, C., Bainbridge, D., Matamua, R., and Downie, J. S. (2015). Improving access to large-scale digital libraries through semantic-enhanced search and disambiguation. In *JCDL '15*, pages 147–156. ACM.

Kay, A. C. (1972). A personal computer for children of all ages. In *ACM '72 (ACM Annual Conference, Volume 1)*. ACM.

Pettifer, S., Thorne, D., McDermott, P., Marsh, J., Villéger, A., Kell, D. B., and Attwood, T. K. (2009). Visualising biological data: a semantic approach to tool and database integration. *BMC Bioinformatics*, 10(Suppl 6):S19.

Piryani, R., Uddin, A., Devaraj, M., and Singh, V. K. (2013). An algorithmic formulation for extracting learning concepts and their relatedness in ebook texts. In *Mining Intelligence and Knowledge Exploration*, pages 529–540. Springer.

Sipe, L. R. (1998). How picture books work: A semiotically framed theory of text-picture relationships. *Children's Literature in Education*, 29(2):97–108.

Sirkiä, T. and Sorva, J. (2015). How do students use program visualizations within an interactive ebook? In *International Computing Education Research*, pages 179–188. ACM.

Tufte, E. (2004). Sparklines: Intense, simple, word-sized graphics. *Beautiful Evidence*, 1:46–63.

Vrandečić, D. and Krötzsch, M. (2014). Wikidata: a free collaborative knowledgebase. *Communications of the ACM*, 57(10):78–85.