# Chapter 4
# Strategic Reading & Conceptual Modeling*

Oscar Díaz

**Abstract** "Strategic reading" is a term coined to conceive reading as a process of constructing meaning by interacting with text. While reading, individuals use their prior knowledge along with clues from the text to construct meaning, and place the new knowledge within this frame. Strategic reading is then a pivotal ability for conceptual modelers, more so if domain knowledge needs to be acquired mainly from the literature as it is the case for research projects. But this might turn problematic. In Quora and other PhD forums, students moan about their frustrating reading and literature review experiences. Traditionally, students are encouraged to annotate while reading. Digital annotations are expected to be useful for supporting comprehension and interpretation. Our belief is that strategic reading (and hence, conceptual modeling) can be more effective if annotation is conducted in direct relationship to a main research activity: root-cause analysis (RCA). RCA can provide the questions whose answers should be sought in the literature. Unfortunately, this process is not supported by current tools. When reading papers, researchers might not be all aware of the issues being raised during RCA. And the other way around, when it comes to RCA, evidences found in the literature might not be promptly accessible. This paper reports on research to develop a technical solution to this problem: a plug-in for Google Chrome that provides seamless integration between a RCA platform (i.e. MindMeister) and a reading platforms (i.e. Mendeley). The aim: improving RCA awareness while reading so that annotations can be traced back to the RCA issues.

**Key words:** Strategic Reading, Root-Cause Analysis, Annotating, Mind Mapping

Oscar Díaz
University of the Basque Country (UPV/EHU)
ONEKIN - Facultad de Informática - San Sebastián, Spain, e-mail: oscar.diaz@ehu.eus

* This work is being jointly developed by Jeremías P. Contel from the University of the Basque Contry (Spain), and John Venable from Curtin Universtity, Perth (Australia).



J. Cabot et al. (eds.), *Conceptual Modeling Perspectives*,
https://doi.org/10.1007/978-3-319-67271-7_4

## 4.1 Introduction

Some years ago, I was working on Active Databases. At this time, my interests were on formalizing database triggers and their execution model to help debugging large trigger sets. We resorted to the Event Calculus, and here, Antoni's work was most influential. Specifically, our work was strongly inspired by his paper "Validating conceptual specifications through model execution"[8]. This could have been a possible subject for this chapter. However, I would like to acknowledge here another side of Antoni's efforts: his dedication to teaching. I had the chance of chatting with Antoni in numerous occasions, and both teaching and students were a common subject. His book on Conceptual Modeling illustrates this concern . Here, I would like to report a recent work which was inspired by one of these chats with Antoni.

It was in Stockholm at CAiSE 2015. After the reception at the City Hall, and wandering along the nice canals that hug this magnificent building, Antoni observed how his students struggle with reading the literature, and particularly, the tendency of students to focus too early on the details of their PhD projects without keeping an eye on related research. This passing comment resonated one year later when I met Prof. John Venable. John has been working on Design Science Research (DSR) for more than fifteen years. DSR highlights the importance of root-cause analysis (RCA) not only at the start of the project but throughout, and how this analysis should be based on data either directly obtained or provided by the literature. The latter reminds me of Antoni's concerns about students focusing too early on their projects without keeping a wider radar at related literature. And then the bulb lighted up: if a pivotal skill for researchers is that of asking the right questions then, we can conjecture that RCA could be the means to find these questions. This paper reports on how this idea was developed[1]

DSR requires a profound understanding of the problem to be solved, the consequences to be alleviated, and the causes to be prevented. This in turn usually implies extracting evidence from the literature that warrants the project's RCA. Reading then, becomes the process of extracting evidence from the literature that sustains the project's RCA. We then conceive of RCA and reading as two inter-related processes which re-adjust and feed off each other: RCA progresses as new insights are obtained from the literature while the literature is scrutinized along the concerns that arise during RCA.

Unfortunately, this interplay lacks appropriate support in current reading tools (e.g. Acrobat Reader, Mendeley) or reference managers (e.g. Mendeley, NVivo, or End-Note). What is needed is a way to bridge the gap between conceptualizing tools – where ideas are shaped and framed – and reading tools – where ideas are sustained and opposed. We believe the challenge is not on creating brand new tools, but on coupling existing tools with minimal interference with existing practices. What is needed is for tools to keep their autonomy, but interact with a double aim:

1. to guide reading (where reading purposes are to be sought in RCA), and

[1] This is an excerpt from a paper presented at DESRIST 2017 [1].

2. to draw on and document supporting evidence for RCA issues (where evidences are obtained during reading).

These two flows are in overlapping motion: RCA concerns guide the reading while the reading comes up with new insights that confirm or refute the RCA issues. This work then addresses the following research question:

> *How can we provide seamless integration between RCA tools and reading tools to improve strategic reading for novice DSR researchers?*

To ground this research in concrete examples, we resort to MindMeister (as the RCA tool) and Mendeley (as the reference manager), being the challenge the one of coupling these tools for the sake of strategic reading (Section 4.3). DSR is defined as "research that invents a new purposeful artifact" [15]. In this case, the artifact is a Chrome plug-in, *DScaffolding*, which bridges the gap between MindMeister and Mendeley (Section 4.4). During reading in Mendeley, *DScaffolding* makes practitioners aware of the evidences being looked for. During RCA in MindMeister, *DScaffolding* makes researchers aware of the evidences that sustain/refute the causes/consequences elaborated during the RCA. *DScaffolding* is publicly available for download at the Chrome's Web Store. Next section elaborates on how Antoni's insight is not just a locale practice but a general concern.

## 4.2 Student Reading Experience

One of the most important skills for researchers to acquire is that of asking the right questions when accessing the literature. The answers you get much depend on the questions you ask. This skill is specially important for PhD students who struggle with an increasing number of papers[2] and stringent PhD deadlines. Based on Mendeley data, PhD students were the main readers of articles in 2008 for all articles [6]. This puts PhD students at the forefront of scientific literature consumption, even ahead of their supervisors! However, it is not rare to come across in Quora or other PhD forums with students moaning about their frustrating reading experiences[3]. Causes can be multifold: lack of time (increasing reading loads), lack of motivation (no prompt feedback from supervisors, reading considered an ancillary activity as opposed to programming where the real meat is) or lack of knowledge (no clear what to look for). If we focus on the latter, forums give some advices:

- "Before you start reading, have a clear idea of what information you are looking for in these papers. This by itself is about 60% of psyching yourself up for reading papers" [12]

[2] A UK study reported an average 39 scholarly readings per month, comprising 22 articles, seven books, and ten other publications \cite{Tenopir2012}, amounting to an estimated 448 hours per year spent reading (equivalent to 56 8-hour days).

[3] As a case in point, refer to `https://www.quora.com/Do-researchers-scientists-find-reading-scientific-papers-exciting` with 17 followers.

- "Make notes of how the research in the paper you're reading connects with your own" [11]
- "Reading a scientific paper should not be done in a linear way (from beginning to end); instead, it should be done strategically and with a critical mindset, questioning your understanding and the findings" [13].
- "As you read, look for the author's main points. Generate questions before, during, and after reading. Draw inferences based on your own experiences and knowledge. And to really improve understanding and recall, take notes as you read" [10].
- "If you want to make it a productive exercise, you need to have a clear idea of which kind of information you need to get in the first place, and then focus on that aspect." [9]
- "At the beginning, new academic readers find it slow because they have no frame of reference for what they are reading. But there are ways to use reading as a system of creating a mental library, and after a few years, it becomes easy to slot papers onto your mental shelves. Then you can quickly skim a paper to know its contribution." [9]

The underpinning assumption seems to be the existence of a "frame of reference". This frame serves to guide the reading, helping in having "a clear idea of which kind of information you need to get in the first place". Purposeful reading is then a distinctive feature of scientific reading as opposed to let's say, playful reading where the aim is to not to know the outcome but instead, to enjoy the poetic narrative and thrilling plot. To get the best of scientific reading, a frame of reference needs to be present.

This work addresses the case for Design Science Research (DSR), a popular approach in Information System [3]. The first question is then how will a "DSR's frame of reference" looks like. This paper's main assumption is that most of the readings during DS projects have (at least) five main foci, namely,

- finding evidence for the importance of the problem,
- ascertaining causal relationships in the problem,
- becoming acquainted with works addressing similar problems,
- becoming acquainted with work that can serve as a kernel theory or other inspiration,
- becoming acquainted with work relevant to research (method) design for the DSR project.

RCA relates to the first two of these. We can then state the problem as

*PhD students not bearing "the RCA frame" in mind when reading*

This might have a manyfold impact:

- Important facts might be overlooked when reading. This in turn, might involve a loss of opportunity for DSR projects. If not properly documented in the RCA, read-ing insights might be forgotten by the time they could impact the project's design, leading to overconfident problem analysis.

- Unfocused reading might result in boredom, lack of engagement and research effort discontinuity among PhD students,
- Literature references might not be traced back to their RCA rationales. This might cause poor reference recoverability when it comes to writing the paper, and hence, forcing re-reads

So far, we can only hypothesize those consequences. Some studies exist on the impact of reading comprehension [2, 4, 14] but this is for settings other that scientific reading. We are unaware of any study that looks into those symptoms for PhD students. That said, the frequent recurrence of this issue in the so-called grey literature (e.g. Q&A forums), provides substantial evidence of the existence of this problem. As a case in point, refer to this Quora entry [11] with 774 followers.

If the consequences are important enough to vindicate a deeper study then, next step is to delve into the causes: *why PhD students do not bear the "the RCA frame" in mind when reading*? We conjecture the following causes:

- No RCA frame available. Students might overlook the importance of RCA in DS projects. In some cases, RCA is hardly sketched, and only verbalized at the time of writing. This may already be too late to surface weak causal links or dubious if not, banal consequences that may hardly sustain the importance of the work. The importance of early RCA can not be stressed enough. This work underlines its role as a reading guideline.
- A RCA frame is available but not easy accessible. Students might have done their homework but they fail to have a presence of the RCA issues when reading. This might be due to reading and RCA being conducted through different tools. So far, the coupling falls on the shoulders of the students through the use of book-notes and copying & pasting between the tools.

This work tackles the second cause. It is not uncommon for researchers to struggle with switching back and forth between e.g. Endnote and Word, to add notes. These approaches tend to be highly manual and error prone, even if conducted through state-of-the-art reference managers. In the end, keeping track of readings represents a considerable burden for students. We then refine the research question as follows:

> *How can we bridge reading tools and RCA tools to ensure the presence of both RCA concerns when reading, and of reading evidences when conducting RCA?*

Next sections elaborates on this question, illustrating the case for MindMeister (as the RCA tool) and Mendeley (as the reference manager). For details about how this example is factored out into general meta-requirements refer to [1].

## 4.3 Coupling MindMeister & Mendeley

Students should be able to freely move between RCA and reading. The interplay between these two activities should be reflected in a tighter integration so that reading is guided by issues risen during RCA, and RCA is further elaborated as ad-

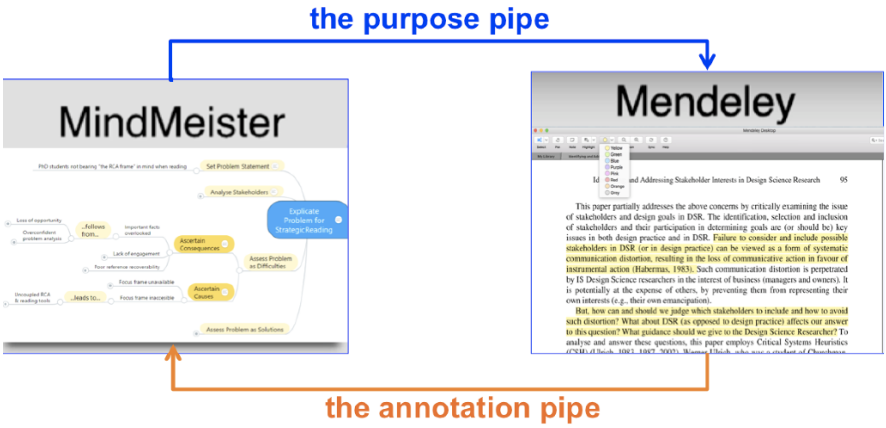


**Fig. 4.1** Coupling through pipes: the purpose pipe & the annotation pipe.

ditional insights are gained from the literature. Figure 4.1 reflects this vision. The RCA-reading coupling is achieved through two pipes. During reading, practitioners should be aware of the evidences being looked for ("the purpose pipe"). During RCA, practitioners should be aware of the evidences that sustain/refute the causes/-consequences in the RCA ("the annotation pipe"). This section introduces how has this been achieved for MindMeister and Mendeley using ***DScaffolding***, a Chrome plug-in. This plug-in is available at the Chrome's Web Store:

`https://chrome.google.com/webstore/detail/hkgmnnjalpmapogadekngkgbbgdjlnne`

Videos are provided for:

- installation: `https://youtu.be/hl6pnJGbVXY`
- the Strategic Reading Process: `https://youtu.be/jHP1MiqjVBM`

Next, we provide an outline.

### *4.3.1 The RCA tool: MindMeister*

MindMeister is a web-based collaborative mind mapping application, which allows its users to visualize their thoughts in terms of mind maps [18]. A mind map is a diagram used to visually organize information. This can be pre-set in terms of a map template, i.e. a set of labelled nodes which can be later expanded by the user by adding new child nodes. This provides a guide to gather information, especially interesting when this information is abundant and multi-sourced. This ductility together with the popularity mind maps enjoy, make mind mapping an interesting approach when it comes to explicating the problem i.e. "to formulate the initial problem precisely, justify its importance, investigate its underlying causes, provide evidences and acknowledge related work" [3].

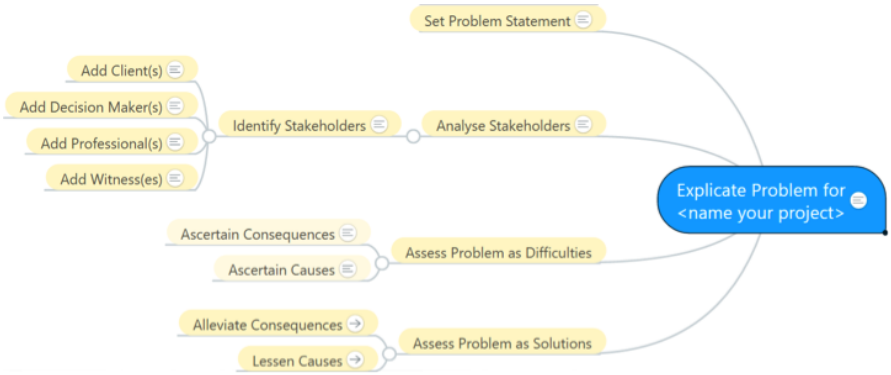


**Fig. 4.2 MindMeister**. The RCA template.

Figure 4.2 depicts the *ExplicateProblem* template at the onset. The template provides a head-start as for the information to be collected. Specifically, we resort to Coloured Cognitive Maps (CCM) [16]. The template supports the two types of CCM:

- the "Problem as Difficulties" node, which focuses on the problem, what is undesirable about it (i.e. consequences), and what causes the problem and allows it to persist, and,
- the "Problem as Solutions" node, which focuses on the solution of the problem, what benefits would accrue from solving the problem or what causes of the problem might be reduced or eliminated to solve the problem.

Details are outside the scope of this chapter. For further information refer to [16].

### *4.3.2 The reading tool: Mendeley*

Mendeley is an Elsevier-owned desktop and a Web program helping to manage and share research papers [17]. Papers can be arranged into folders, and tagged for easy retrieval. Figure 4.3 shows the content of the *StrategicReading* folder, and particularly, the metadata for the selected paper. Mendeley includes a PDF viewer with sticky notes, text highlighting and full-screen reading. Quote annotation is achieved through highlighting where different colours are available.

### *4.3.3 The coupling*

Broadly, coupling MindMeister and Mendeley involves three main challenges:

- the ability to indicate what issues risen during RCA in MindMeister need supporting literature evidence (hereafter referred to as "reading concerns"),

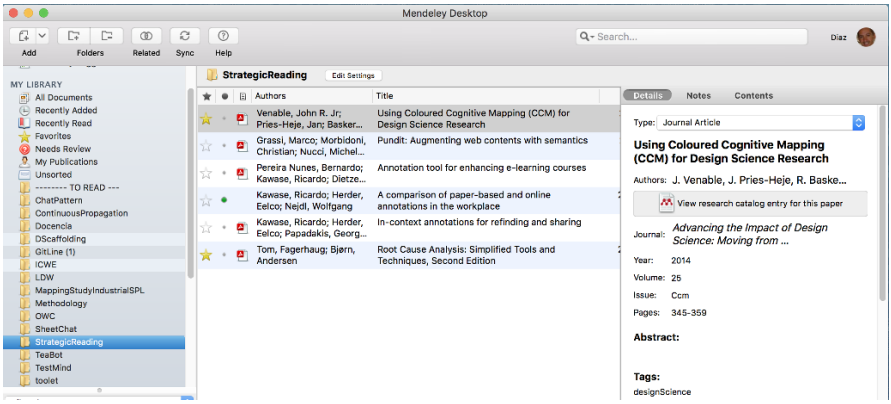


**Fig. 4.3 Mendeley**. Papers are organized along folders (e.g. *StrategicReading*).

- a way to keep the presence of reading concerns while reading in Mendeley,
- the ability to access quotes from Mendeley while ellaborating RCA in MindMeister, so that students can assess the extent causes/consequences are backed up by quotes.

Next section illustrates how is this fleshed out by *DScaffolding*.

## 4.4 DScaffolding at work

Broadly, three mains tasks interplay during RCA-based strategic reading, namely, "conducting RCA", "setting RCA issues" and "reading". The first two are conducted in MindMeister while reading takes place in Mendeley. This section illustrates these steps for the problem: *"PhD students not bearing the RCA frame in mind when reading"*, i.e. our very problem!

### *4.4.1 Conducting RCA*

To provide a head-start, MindMeister is being extended with the "Explicate Problem" template (see Figure 4.2). Students need to add the corresponding children nodes. Figure 4.4 instantiates the RCA template for the problem *"PhD students not bearing the RCA frame in mind when reading"*. We stick to MindMeister gestures for node management, so no new interaction needs to be learnt.

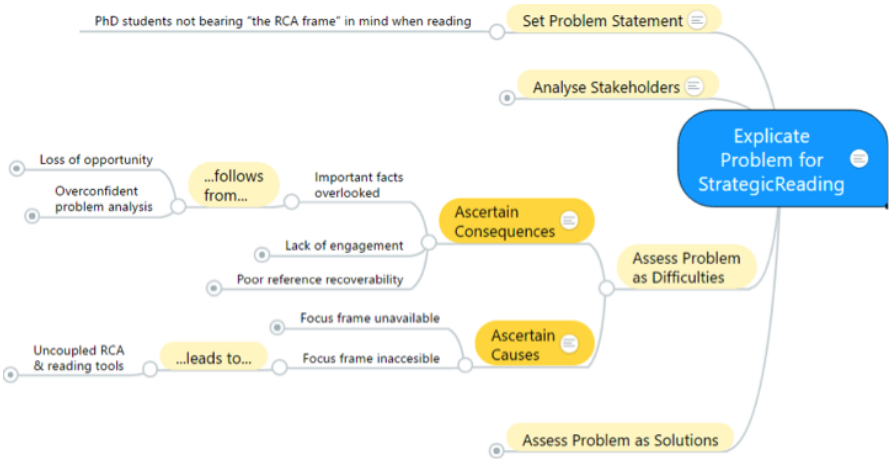


**Fig. 4.4 MindMeister**. *ExplicateProblem* template instantiated for the problem *"PhD students not bearing the RCA frame in mind when reading"*.


**Fig. 4.5 MindMeister**. Setting reading concerns by adding *"Supporting Evidences?"* nodes. The panel at the right keeps track of the current reading concerns.

### *4.4.2 Setting RCA issues*

During RCA, students might wonder: what proof do I have that this cause exists?, what proof do I have that this cause actually contributed to the problem I am looking at?, are they merely asserting causation?, is anything else needed, along with this cause, for the stated effect to occur? is it self-sufficient? Frequently, the answers to these questions should be sought in the literature. However, questions might be too numerous to be addressed simultaneously. Hence, users might decide to focus on some aspects while postponing others. Current foci are termed "reading purposes" as far as they refer to causes/consequences than need to be backed up by the literature.

Back to *DScaffolding*, RCA nodes are turned into "reading purposes" by adding a child with the label ***"Supporting Evidences?"*** Introducing such node turns the father into a "reading purpose". This is indicated by decorating the father node with

one of up to eight of the different background colours used in Mendeley (see later). Figure 4.5 illustrates the case for the running example. The user sets three reading concerns: the problem statement (in green), *"Poor reference recoverability"* (in pink), and *"Uncoupled RCA and reading tools"* (in purple). This means that students should look for quotes that somehow sustain these issues during their current readings. The current reading concerns can be obtained from the MindMeister panel. Reading concerns can be modified at any time as evidences are found or new insights advice to move the focus to another cause/consequence. To turn a node into a reading concern just extend it with the ***"Supporting Evidences?"*** child. To stop a node from being a reading concern, go to the MindMeister panel and delete it. The corresponding *"Supporting Evidences?"* node is not deleted but its associated cause/consequence is no longer track in Mendeley (see next).

### *4.4.3 Strategic Reading*

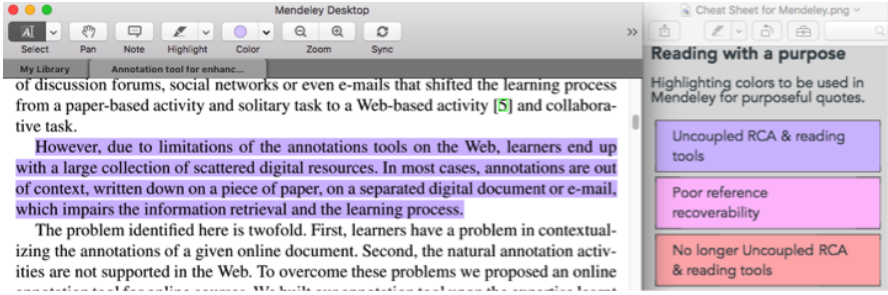


**Fig. 4.6 Mendeley**. Cheat-sheet used for RCA awareness (left side). The cheat-sheet is obtained as a screenshot through MindMeister panel.

Strategic reading implies an aim. Here, the aim is looking for evidences for the current reading concerns. Here, we resort to annotation to denote the existence of evidences. Digital annotations are expected to be useful for supporting comprehension and interpretation [5, 7]. Here, strategic reading is realized as annotation highlighting.

Mendeley provides eight different colours for annotation highlighting. Yellow is left for "structural" highlighting (i.e. attributing different levels of importance). The remaining seven are mapped to RCA-based reading concerns. Specifically, By using the very same colors in Mendeley and MindMeister, a mapping is set between Mendeley highlights and reading concerns in MindMeister. A cheat sheet about what these colours stand for can be obtained from MindMeister[4]. Researchers can

[4] A request is being posted to Mendeley to permit color legends to be configurable. This will permit *DScaffolding* to set legends based on RCA issues, avoiding to resort to the burdersome cheat-sheet.

then place this cheat-sheet by their Mendeley desktop application. Figure 4.6 provides an example. While reading a paper, a paragraph might well sustain one of the concerns risen during RCA: *"RCA & reading tool coupling"*. Since this issue is associated with the velvet color, this is the color used for highlighting. In this rudimentary way, students will keep a presence of what they are looking for.

### *4.4.4 Back to conducting RCA*

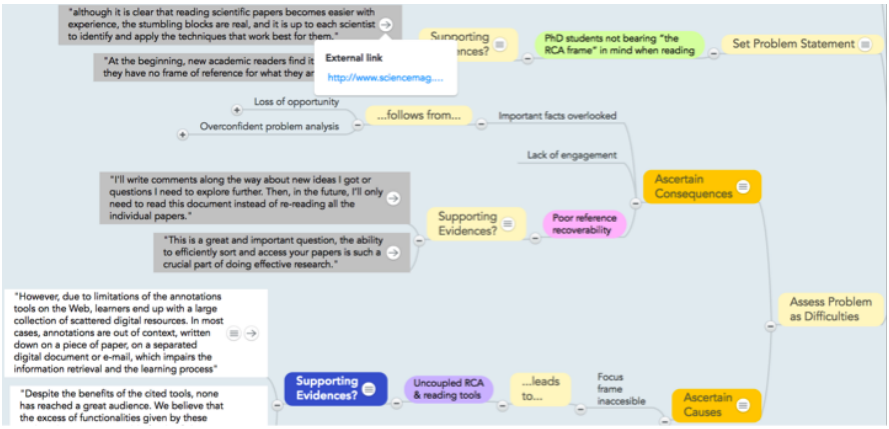


**Fig. 4.7 MindMeister**. Quotes are automatically brought at the RCA map and hung below their related issues.

At any time, students can go back to their RCA maps. *DScaffolding* tracks annotations made in papers held in Mendeley folders, to enrich the **namesake** MindMeister map. That is, students can keep different RCA maps and different Mendeley folders. The mapping between maps and folders is set based on sharing the same name.

On uploading a MindMeister map, *DScaffolding* checks out whether any new quote is being found since it was last sync with the user's Mendeley account. If so, *DScaffolding* automatically enriches the map with the new quotes. Specifically, quote nodes hang from the nodes to which they provide an evidence for. Figure 4.7 illustrates our sample map at a later stage where some new quotes have been found.

Node properties include: a label, an attached comment and a background colour. For nodes created automatically out of quotes, these properties behave as follows:

- the label holds the text being highlighted in the annotated resource,
- the comment keeps a link to the resource URL (if available). Researchers can click on the link icon to move straight to the manuscript in Mendeley, and in so doing, looking at the quote in context,

- the background colour reflects the nature of the source: "white" for annotations coming from journals and conferences, and "grey" if coming from the grey literature (not discussed here).

In addition, quotes inherit the reputation of their sources. Annotations coming for reputable sources add a "star" icon to their labels. So far, the reputation is set by users. For instance, Mendeley allows users to tick a "star" to mark sources as favourites. Although "favourite" is quite an elusive notion (no clear rationale for ticking this off), *DScaffolding* interprets the star as a sign of the source's reputation and soundness. This reputation travels together with the reference.

To conclude, this work considers reading and RCA as two inter-related processes. If this is so, practitioners should be helped in moving between the reading realm (e.g. Mendeley) and the RCA realm (e.g. MindMeister). *DScaffolding* aims at assisting in seamlessly moving between Mendeley and MindMeister, and in so doing, helping students to improve focus while reading as well as fostering RCA throughout their projects.

## 4.5 Conclusions

Antoni cared about how to engage students in reading. Inspired by this concern, this work sets a Design Theory whereby RCA may provide main drivers of attention when reading as well as supporting the importance of RCA throughout the whole project. The theory states that this can be achieved by sustaining both

- "RCA awareness" while reading (i.e. the purpose pipe that channels RCA issues to reading platforms), and
- "literature awareness" while conducting RCA (i.e. the annotation pipe that channels literature quotes towards RCA platforms).

We built *DScaffolding* to assess the extent to which this theory holds. First evaluations indicate that not only reading but also RCA might benefit from a tight coupling between these two processes (refer to [1] for further insights). We do hope Antoni like the approach!

## References

1. Díaz, O., Contell, J.P., Venable, J.R.: Strategic reading in design science: Let root-cause analysis guide your readings. In: Designing the Digital Transformation - 12th International Conference, DESRIST 2017, Karlsruhe, Germany, May 30 - June 1, 2017, Proceedings. pp. 231–246 (2017), https://doi.org/10.1007/978-3-319-59144-5_14
2. Israel, S.E., Duffy, G.G.: Handbook of research on reading comprehension. Routledge (2014)
3. Johannesson, P., Perjons, E.: An introduction to design science. Springer (2014)
4. Margolin, S.J., Driscoll, C., Toland, M.J., Kegler, J.L.: E-readers, computer screens, or paper: Does reading comprehension change across media platforms? Applied Cognitive Psychology 27(4), 512–519 (2013)

5. Marshall, C.C.: Toward an ecology of hypertext annotation. In: Proceedings of the ninth ACM conference on Hypertext and hypermedia: links, objects, time and space—structure in hypermedia systems: links, objects, time and space—structure in hypermedia systems. pp. 40–49. ACM (1998)
6. Mohammadi, E., Thelwall, M., Haustein, S., LariviÃĺre, V.: Who reads research articles? an altmetrics analysis of mendeley user categories. Journal of the Association for Information Science and Technology 66(9), 1832–1846 (2015), http://dx.doi.org/10.1002/asi.23286
7. O'hara, K., Sellen, A.: A comparison of reading paper and on-line documents. In: Proceedings of the ACM SIGCHI Conference on Human factors in computing systems. pp. 335–342. ACM (1997)
8. Olivé, A., Sancho, M.R.: Validating conceptual specifications through model execution. Information Systems 21(2), 167–186 (1996)
9. Pain, E.: How to (seriously) read a scientific paper, http://www.sciencemag.org/careers/2016/03/how-seriously-read-scientific-paper. Accessed 6 Feb 2017
10. Purugganan, M., Hewitt, J.: How to read a scientific article, http://www.owlnet.rice.edu/ cainproj/courses/HowToReadSciArticle.pdf. Accessed 6 Feb 2017
11. Quora: How do you keep your notes while reading scientific papers?, https://www.quora.com/PhD-students-How-do-you-keep-your-notes-while-reading-scientific-papers. Accessed 6 Feb 2017
12. Quora: I am a robotics phd student and i have a hard time reading research papers. i am very slow at it and find the task kind of boring. is there any way i can make paper reading fun and become faster at it?, https://www.quora.com/Iam-a-robotics-PhD-student-and-I-have-a-hard-time-reading-research-papers-Iam-very-slow-at-it-and-find-the-task-kind-of-boring-Is-there-any-way-I-can-makepaper-reading-fun-and-become-faster-at-it. Accessed 6 Feb 2017
13. Rodriguez, N.: Infographic: How to read a scientific paper, https://www.elsevier.com/connect/infographic-how-to-read-a-scientific-paper. Accessed 6 Feb 2017
14. Stern, P., Shalev, L.: The role of sustained attention and display medium in reading comprehension among adolescents with adhd and without it. Research in developmental disabilities 34(1), 431–439 (2013)
15. Venable, J., Baskerville, R.: Eating our own cooking: Toward a more rigorous design science of research methods. Electronic Journal of Business Research Methods 10(2), 141 – 153 (2012), http://search.ebscohost.com/login.aspx?direct=true{&}db=bsx{&}AN=87404098{&}lang=es{&}site=eds-live
16. Venable, J.R.: Using coloured cognitive mapping (ccm) for design science research. In: International Conference on Design Science Research in Information Systems. pp. 345–359. Springer (2014)
17. Wikipedia: Mendeley, https://en.wikipedia.org/wiki/Mendeley. Accessed 6 Feb 2017
18. Wikipedia: Mindmeister, https://en.wikipedia.org/wiki/MindMeister. Accessed 6 Feb 2017