

# An Exploratory Analysis of Mind Maps

Joeran Beel

Docear & OvGU/FIN/DKE  
Magdeburg, Germany  
beel@docear.org

Stefan Langer

Docear & OvGU/FIN/DKE  
Magdeburg, Germany  
langer@docear.org

## ABSTRACT

The results presented in this paper come from an exploratory study of 19,379 mind maps created by 11,179 users from the mind mapping applications 'Docear' and 'MindMeister'. The objective was to find out how mind maps are structured and which information they contain. Results include: A typical mind map is rather small, with 31 nodes on average (median), whereas each node usually contains between one to three words. In 66.12% of cases there are few notes, if any, and the number of hyperlinks tends to be rather low, too, but depends upon the mind mapping application. Most mind maps are edited only on one (60.76%) or two days (18.41%). A typical user creates around 2.7 mind maps (mean) a year. However, there are exceptions which create a long tail. One user created 243 mind maps, the largest mind map contained 52,182 nodes, one node contained 7,497 words and one mind map was edited on 142 days.

## Categories and Subject Descriptors

I.7.5 [Document and Text Processing]: Document Capture – document analysis

## General Terms

Measurement, Design.

## Keywords

mind maps, concept maps, content analysis, document analysis, mind mapping software, information retrieval

## 1. INTRODUCTION

Millions of people are using mind maps for brainstorming, note taking, document drafting, project planning and other tasks that require hierarchical structuring of information. Figure 1 shows a mind map which was created as draft for this paper. As all mind maps, it has a central node (the root) which represents the main topic the mind map is about. From this root node, child-nodes branch out, in order to describe sub-topics. Each node may contain an arbitrary number of words. This way, a mind map is comparable to an outline but with stronger focus on the graphical representation. Mind maps created on a computer may also contain links to files, hyperlinks to websites (in Figure 1 indicated by red arrows), pictures, and notes (indicated by yellow note icons).

In this paper we present the initial results of an exploratory study of 19,379 mind maps. The overall research objective was to find out how mind maps are structured and what information they

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

DocEng'11, September 19–22, 2011, Mountain View, California, USA.  
Copyright 2011 ACM 978-1-4503-0863-2/11/09...\$10.00.

contain. To our knowledge this is the first study of its kind. We therefore aimed at a broad overview to determine further areas of interesting research.

## 2. RELATED WORK

There is lots of research on content and structure of other documents: Web pages, emails, academic articles, etc. have all been analyzed thoroughly in the past (e.g. [1-3]). With respect to mind maps there is mostly research about the effectiveness as learning tool (e.g. [4]).

The lack of analyses of mind maps is not surprising. Emails, web pages, etc. had to be thoroughly researched to make information retrieval tasks, for instance, indexing and spam detection, effectively possible. Such information retrieval tasks have never been applied to mind maps, and therefore the need for knowledge about mind map content and structure was low.

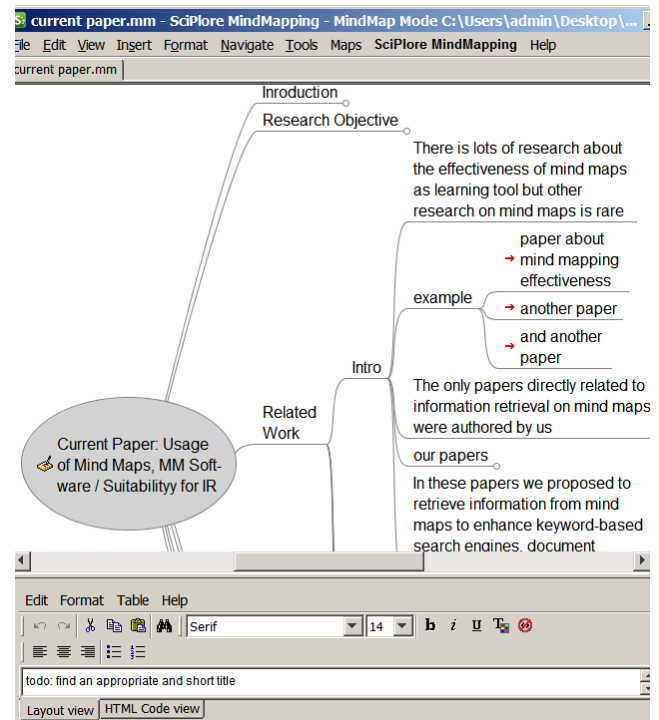


Figure 1: Screenshot of a mind mapping software

However, recently we proposed to apply information retrieval tasks to mind maps to enhance keyword-based search engines, document recommender systems, and user profile generation [5]. To do this effectively, knowledge about the content and structure of mind maps is required.

There was only one paper we found that is somewhat related: a survey from the *Mind Mapping Software Blog* [6]. For this survey 334 participants answered questions about their use of mind mapping software. However, the survey was based on 334 self-

selected participants from a single source (readers of the *Mind Mapping Software Blog*). Accordingly, it seems likely that predominantly very active mind mapping users participated in the survey and results are not representative. In addition, the survey focused on the usage of mind mapping software rather than the content and structure of mind maps.

### 3. METHODOLOGY

We conducted an exploratory study on 19,379 mind maps created by 11,179 users from the two mind mapping applications *Docear*<sup>1</sup> and *MindMeister*<sup>2</sup> (the latter one is abbreviated as ‘MM’ in figures and tables).

Docear is a mind mapping application for Windows, Linux and Mac, focusing on academic literature management, and developed by ourselves [7]. 2,779 users agreed to have their mind maps analyzed. They created 7,506 mind maps between April 1, 2010 and March 31, 2011.

MindMeister is a web based mind mapping application. 8,400 users published 11,873 mind maps in MindMeister’s public mind map gallery<sup>3</sup> between February 2007 and October 2010. For our study these public mind maps were downloaded in XML format via MindMeister’s API<sup>4</sup>, parsed, and analyzed.

Numbers include only mind maps containing six or more nodes<sup>5</sup>, and that were not being edited between April 1, 2011 and the day of the analysis (June 2, 2011). This way it is ensured that mind maps in the beginning of their life-cycle do not spoil the results but only “mature” mind maps were analyzed.

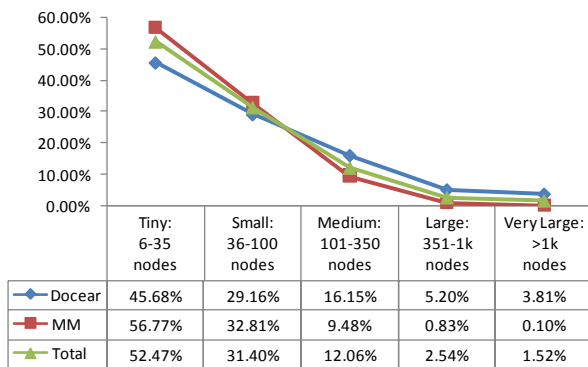


Figure 2: Distribution of mind maps based on size (number of nodes)

We were particularly interested in finding out whether differences existed for different types of mind maps and between the two mind mapping applications. Therefore, mind maps were grouped based on their size, measured by the number of nodes. Mind maps with 6 to 35 nodes were considered as ‘tiny’, with 36 to 100 nodes as ‘small’, with 101 to 350 nodes as ‘medium’, with 351 to 1000 nodes as ‘large’ and with more than 1000 nodes as ‘very large’. In the data set, the majority of mind maps were tiny (52.47%) or small (31.40%) as shown in Figure 2.

## 4. RESULTS & INTERPRETATION

### 4.1 Mind Maps per User

Figure 3 shows the number of mind maps users created. The majority of MindMeister users created, or we should say published, exactly one mind map (81.26%). Only 2.32% of MindMeister users published five or more mind maps. In contrast, 56.75% of Docear users created one mind map and 11.36% created five or more mind maps. On average (mean), users created 2.7 mind maps (Docear) during the 12 month period of data collection, respectively 1.4 (MindMeister) during ~3.5 years. The highest number of mind maps created by one user was 243 for Docear and 73 for MindMeister.

It has to be noted that numbers of MindMeister and Docear are only limitedly comparable, as we did only analyze MindMeister mind maps that were published by their users. It can be assumed that most users who published mind maps on the Web, created further private mind maps that were not publicly available.

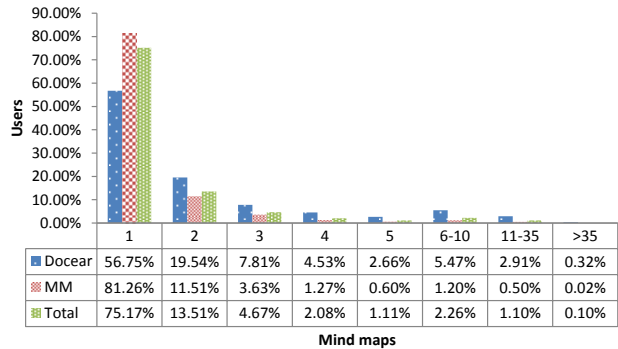


Figure 3: Number of created mind maps per user

### 4.2 Nodes per mind map

As mentioned in the methodology and shown in Figure 2, most mind maps were rather small. On average, Docear mind maps contained 232 nodes (mean), respectively 41 nodes (median). MindMeister mind maps contained 51 nodes (mean), respectively 31 (median). Docear mind maps tended to be larger than MindMeister mind maps. For instance, while only 0.10% of MindMeister mind maps were ‘very large’, 3.81% of Docear mind maps were. The largest Docear mind map contained 52,182 nodes (and there are several more mind maps containing 10,000+ nodes); the largest MindMeister mind map contained 2,318 nodes.

### 4.3 File Links

In a mind map, users may link to files on their hard drive. Figure 4 shows the distribution of mind maps containing a certain number of links (for Docear mind maps only since MindMeister does not provide this feature). Well over half of mind maps do not contain any links to files (63.88%).

Table 1: File types linked in mind maps

PDFs	Images	Documents	HTML	Excel/CSV	PowerPoint	MP3s	Other
89.58%	1.26%	0.53%	0.47%	0.42%	0.34%	0.27%	7.14%

However, some users make heavy use of the feature. 2.94% of mind maps contained more than 1,000 links to files and 2.97% of mind maps contained between 351 and 1,000 links. The highest number of links in a mind map was 52,138 and all 7,506 Docear mind maps together contained 1,184,547 links to files on the users’ hard drives. This does not mean that 1,184,547 different files were linked. Most users linked the same file multiple times in a mind map.

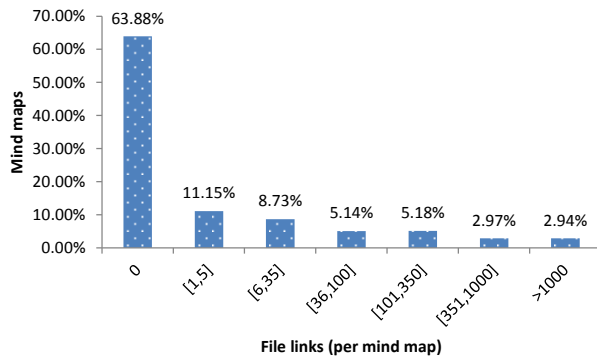
<sup>1</sup> <http://docear.org>

<sup>2</sup> <http://mindmeister.com>

<sup>3</sup> <http://mindmeister.com/maps/public>

<sup>4</sup> <http://mindmeister.com/services/api>

<sup>5</sup> A random sample of 50 mind maps showed that the vast majority of mind maps with five or fewer nodes were created for testing purposes and did not contain valuable content.

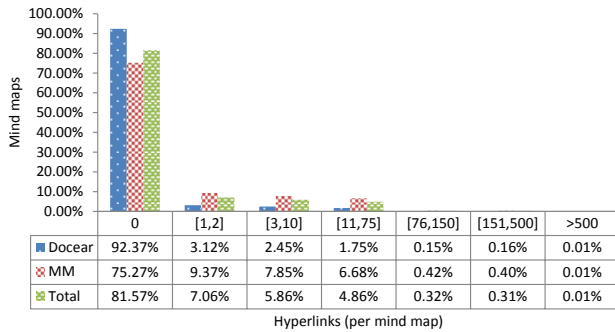


**Figure 4: Number of file-links in Docear mind maps**

From all links, 89.58% pointed to PDF files (see Table 1). Other files being linked included images (.gif, .png, .jpeg, .tiff), MP3s and text documents (.doc, .docx, .odt, .rtf, .txt), but with much smaller frequency.

#### 4.4 Hyperlinks

Looking at all mind maps, 81.57% do not contain a single hyperlink to a website (see Figure 5). However, there are differences between Docear and MindMeister. While 92.37% of Docear mind maps do not contain hyperlinks at all, only 75.27% of MindMeister mind maps do not contain any hyperlinks. In other words: 7.63% of Docear mind maps and 24.73% of MindMeister mind maps contain at least one hyperlink.



**Figure 5: Number of hyperlinks in mind maps**

Larger mind maps more often contain hyperlinks when compared to smaller mind maps. For instance, around 20% of Docear's (very) large mind maps but only 3.94% of tiny mind maps contain hyperlinks. Similarly, around 40% of MindMeister's (very) large mind maps but only 22% of tiny mind maps contain hyperlinks<sup>6</sup>.

#### 4.5 Notes

Most mind mapping software tools (such as Docear and MindMeister) allow users to add notes to a node.

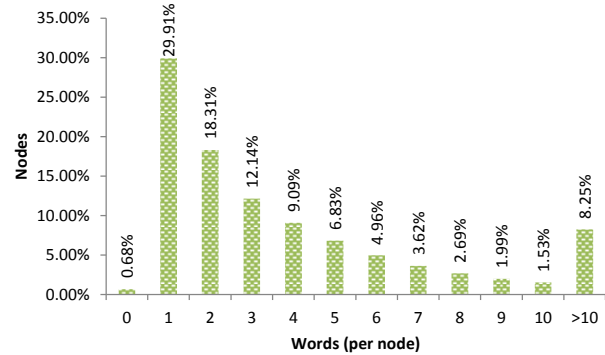
**Table 2: Number of notes in mind maps**

Mind Maps Size	Amount of Notes						
	0	[1,2]	[3,10]	[11,75]	[76,150]	[151,500]	>500
<b>Tiny</b>	68.66%	19.53%	8.00%	3.81%	0.00%	0.00%	0.00%
<b>Small</b>	65.72%	15.02%	9.58%	9.53%	0.15%	0.00%	0.00%
<b>Medium</b>	59.58%	13.73%	9.92%	14.37%	1.97%	0.43%	0.00%
<b>Large</b>	52.15%	11.86%	11.66%	17.59%	4.29%	2.25%	0.20%
<b>Very large</b>	61.74%	6.04%	7.38%	16.44%	3.69%	3.36%	1.34%
<b>Total</b>	66.12%	17.01%	8.81%	7.42%	0.45%	0.16%	0.03%

Many users do not use this feature – 66.12% of mind maps do not contain any notes (see Table 2). Results are similar for both, MindMeister and Docear mind maps<sup>6</sup>.

#### 4.6 Words per node

Figure 6 shows the distribution of words per node (everything separated by whitespace characters was assumed to be a word). Nodes in mind maps generally contain few words. Nearly 1/3 of all 2,352,584 nodes contained a single word (29.91%). Only 8.25% of nodes contained more than ten words.



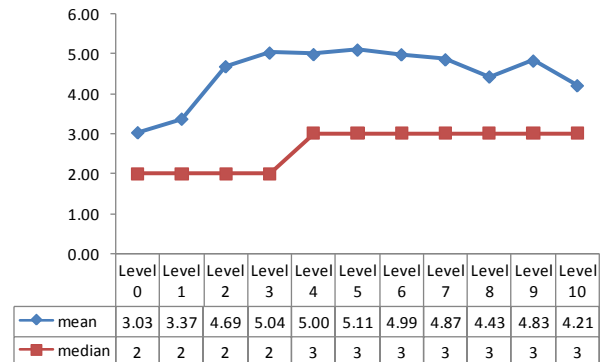
**Figure 6: Number of words per node**

However, there is a long tail in the distribution – the maximum word count for a node was 7,497 for Docear and 1,184 for MindMeister. Although the most frequent word count per node is one, mean is 4.80 words per node and median is 3. There is a slight tendency that the larger mind maps are, the more words their nodes contain. Details are provided in Table 3.

**Table 3: Number of words per node by mind map size**

	Word count per node			
	Mean	Median	Modal	Max
<b>Tiny maps</b>	4.67	2	1	1,874
<b>Small maps</b>	4.45	2	1	687
<b>Medium maps</b>	5.07	2	1	1,463
<b>Large maps</b>	5.76	3	1	2,723
<b>Very large maps</b>	4.60	3	1	7,497

Also, the deeper a node is in a mind map (further out on the branch), the more words it tends to contain. While root nodes (level 0) contain 3.03 words on average (mean), respectively 2 (median), nodes in level 5 contain 5.11 words on average (mean), or 3 (median) respectively (see also Figure 7).



**Figure 7: Number of words per node based on node level**

<sup>6</sup> Detailed results are not provided due to space restrictions.

Results are similar for both, Docear and MindMeister mind maps<sup>6</sup>. Except, the median word count for Docear is three, and for MindMeister two.

## 4.7 Days Edited

The majority of mind maps seem to be used for rather short term activities such as brainstorming or maybe taking meeting-minutes. Figure 8 shows on how many days mind maps were edited<sup>7</sup>. 60.76% of mind maps were edited only during a single day<sup>8</sup>. However, also a large proportion of mind maps were edited on several days, and a small fraction (0.55%) even on more than 25 days. On average, mind maps were edited on one day (median), respectively 2.36 days (mean). The maximum was 142 days.

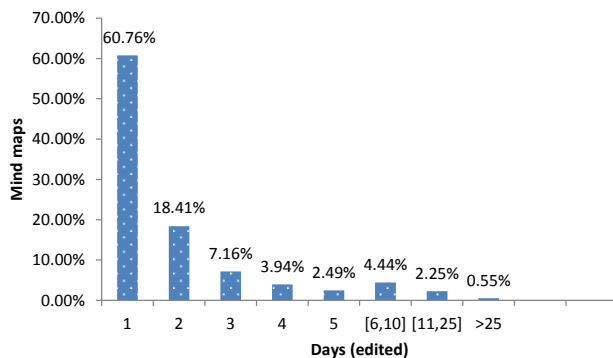


Figure 8: Number of days mind maps were edited

## 5. INTERPRETATION & SUMMARY

For some features, there appear to be significant differences between mind maps created with Docear and those created with MindMeister. However, most of the differences can be attributed to the special functionality of the corresponding software. For instance, Docear offers special features for literature management such as automatically importing PDF bookmarks as new nodes to a mind map. Accordingly, it was expected that Docear mind maps would be larger, in terms of number of nodes. Concerning this case, probably MindMeister numbers are more representative for other mind maps than Docear's are.

On the other hand, when estimating the number of mind maps per user, Docear's numbers are probably more suitable for generalizations, as we could only analyze public mind maps of MindMeister users.

The study showed that a 'typical' (average) mind map is rather small, with a few dozen nodes (31 was the median for MindMeister mind maps), whereas each node contains probably between one to three words (more for large mind maps or nodes deeper in a mind map). The mind map probably contains few if any notes (66.12%). The number of hyperlinks depends on the mind mapping application and tends to be rather low, too. Probably the mind map was edited only on one (60.76%) or two days (18.41%) and it is expected that a typical user creates around 2.7 mind maps a year (mean, Docear).

However, these are only averages. Most results followed a power-law distribution with a long tail. There was one user who created 243 mind maps (and several users more created 10+ mind maps).

<sup>7</sup> Data was available for Docear mind maps only.

<sup>8</sup> Creation of a mind map was counted as one edit. All edits made during one day were combined.

The largest mind map in the data set contained 52,182 nodes (and several more with 10,000+ nodes existed), there was one node containing 7,497 words (and several more nodes with 100+ words existed), one mind map was edited on 142 days (and several more were edited a few dozen times) and several mind maps contained a few hundred notes.

## 6. OUTLOOK

For future research, analysis of the evolution of mind maps could be interesting. Maybe there are different patterns how mind maps evolve and are used by users. Also, differences between user types should be analyzed. In addition, the content of mind maps has only been analyzed superficially, yet. It would be interesting to know what exactly the content is and what mind maps are used for exactly (brainstorming, literature management, etc.). A more detailed analysis should also look at the extremes and outliers (e.g. the node with 7,497 words).

Most importantly, mind maps need to be compared to other types of documents and consequences for information retrieval needs to be drawn. What does it mean when nodes usually contain one to three words? Are they comparable to search queries which usually consist of a similar number of terms? If so, can approaches for search query recommender easily be adopted to create a 'node recommender'? Are mind maps with a few dozen nodes comparable to a user's collection of social tags which usually also consist of a few dozen tags each with one or two words? If so, can approaches for user modeling based on social tags easily be applied to model the interests of mind map users? And are mind maps, which contain a few thousands nodes or words, comparable to web pages, academic articles, or emails? If so, what does this mean for the ability to apply information retrieval on mind maps? All these questions need to be answered in further research.

## 7. ACKNOWLEDGEMENT

This work was supported by grant No. 03EGSST033 of the 'German Federal Ministry of Economics and Technology', 'German Federal Ministry of Education and Research', and 'European Social Fund' (ESF).

## REFERENCES

- [1] Alexandros Ntoulas, Marc Najork, Mark Manasse, and Dennis Fetterly. Detecting spam web pages through content analysis. In *15th International Conference on World Wide Web*, pages 83–92, 2006.
- [2] Dean F Sittig. Results of a content analysis of electronic messages (email) sent between patients and their physicians. *BMC Medical Informatics and Decision Making*, 3 (11), October 2003..
- [3] Ken Hyland and Polly Tse. "i would like to thank my supervisor". acknowledgements in graduate dissertations. *International Journal of Applied Linguistics*, 14 (2): 259–275, July 2004.
- [4] Glennis Edge Cunningham. *Mindmapping: Its Effects on Student Achievement in High School Biology*. PhD thesis, The University of Texas at Austin, 2005.
- [5] Joeran Beel, Bela Gipp, and Jan-Olaf Stiller. Information Retrieval on Mind Maps - What could it be good for? In *Proceedings of the 5th International Conference on Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom'09)*, 2009. IEEE.
- [6] Chuck Frey. Mind mapping software user survey, March 2010. URL <http://mindmappingsoftwareblog.com/2010-survey-form/>.
- [7] Joeran Beel, Bela Gipp, Stefan Langer and Marcel Genzmehr. Docear: An academic literature suite for searching, organizing and creating academic literature. In *Proceedings of the 11th ACM/IEEE Joint Conference on Digital Libraries (JCDL'11)*, 2011.