



Thank you for downloading this document from the RMIT Research Repository.

The RMIT Research Repository is an open access database showcasing the research outputs of RMIT University researchers.

RMIT Research Repository: <http://researchbank.rmit.edu.au/>

Citation:

See this record in the RMIT Research Repository at:

Version:

Copyright Statement: ©

Link to Published Version:

PLEASE DO NOT REMOVE THIS PAGE

Building an analysis of new venture startup with Leximancer

Heather Douglas

School of Management RMIT University, Melbourne, Australia

Email: heather.douglas@rmit.edu.au

ANZAM Conference, Adelaide, December 2010

ABSTRACT

Leximancer is a valuable qualitative analysis tool. Analysing qualitative data requires expertise on the part of the researcher. Leximancer is a computer assisted text analysis program that examines the actual language of research participants. The software generates visual concept matrix maps that display principal concepts and themes and relationships among them. Leximancer provides a system for discovering the underlying core associations in a body of text while reducing any bias that may occur with grounded thematic coding. A study examining the startup of entrepreneurial ventures combined Leximancer and thematic analyses. The dual analysis verified major concepts and themes generated in Leximancer, thus increasing the quality of the study by providing coherence, credibility and confirmability of findings.

Keywords: methods in entrepreneurship research, new ventures, narrative analysis, Leximancer

Building an analysis of new venture startup with Leximancer

ABSTRACT

Leximancer is a valuable qualitative analysis tool. Analysing qualitative data requires expertise on the part of the researcher. Leximancer is a computer assisted text analysis program that examines the actual language of research participants. The software generates visual concept matrix maps that display principal concepts and themes and relationships among them. Leximancer provides a system for discovering the underlying core associations in a body of text while reducing any bias that may occur with grounded thematic coding. A study examining the startup of entrepreneurial ventures combined Leximancer and thematic analyses. The dual analysis verified major concepts and themes generated in Leximancer, thus increasing the quality of the study by providing coherence, credibility and confirmability of findings.

Keywords: methods in entrepreneurship research, new ventures, narrative analysis, Leximancer

Qualitative data offers researchers valuable insights into meanings of complex situations based on the actual understandings of research participants. Innovative and inductive research methods are appropriate way to investigate emerging phenomenon (Hine & Carson, 2007) such as social entrepreneurship which has materialised recently as a new academic field (Weerawardena & Sullivan Mort, 2006). Without doubt, high quality analysis is vital to generate reliable explanations from qualitative data, yet analysing qualitative data is acknowledged a complex task (Strauss & Corbin, 1998). Quality in the qualitative research process requires expertise on the part of the researcher to select suitable cases, collect data appropriately, minimise possible errors, conduct rigorous data analysis, provide coherence, and generate trustworthy explanations. Denzin and Lincoln (2003) maintain quality in qualitative research is based on soundness and authenticity. For qualitative research, credibility, transferability, dependability and confirmability replace the usual positivist criteria of internal and external validity, reliability and objectivity (Denzin & Lincoln, 2003). Triangulation, or the addition of a third source, assists confirmability of research findings by checking and cross checking data collected from different sources, in different ways, or by using different analysis techniques (Stake, 1995). Triangulation improves credibility, coherence, and the corroboration of research findings (Creswell, 1998; Miles & Huberman, 1994).

Case studies are a useful method to analyse new or complex situations where the concepts are not well established (Yin, 2003). Eisenhardt (1989) offered a comprehensive overview of the process of compiling meaning and described the challenges of building theory from case study data. It is a common practice in analysing qualitative data to generate meanings through a grounded theory approach (Strauss & Corbin, 1990). In grounded theory, the actual language used by research participants is transcribed and then systematically reviewed either manually or in a computer data program such as NVivo. Lists of concepts are compiled into meaningful notions that construct understandings. These notions are assembled into themes, explanatory models, and possibly new theory. This method of qualitative data analysis required researchers to become very familiar with the content of participant interviews and construct meanings in an iterative process across all the data (Mason, 1996; Silverman, 2000). Analysing qualitative data in this way is an interpretative task requiring diligence, skill and reflection (Silverman, 2001).

This paper examines Leximancer as a data analysis tool and considers its application in a study of the startup of social entrepreneurial ventures. Leximancer is proposed as a useful tool to examine written data (Smith, 2003), but it is not used widely in business research (Bumbieris, Kittel-Wegner & Hine, in Hine & Carson, 2007; Sullivan Mort, Weerawardena, & Liesch, 2008). Leximancer and NVivo are compared as a way of extracting meaning. This paper demonstrates the process of extracting meaning from Leximancer maps to illustrate the advantage of combining Leximancer analyses with a grounded theory techniques when analysing interview data. The paper argues Leximancer is a helpful tool that provides clear understandings and valuable insights.

The paper proceeds in the following manner. First, the research situation is described briefly to set the context of the study. Second, the method of incorporating Leximancer analyses is described and a Leximancer concept map is presented to illustrate how meaning can be generated. Third, the advantages of incorporating Leximancer analysis are discussed.

THE STUDY

The research investigated the startup of social entrepreneurial ventures in Australia. Social entrepreneurship is a relatively new field of study in which concepts are not well established (Chell, 2007; Haugh, 2007; Nicholls, 2006). Thus, a case study method is appropriate (Stake, 1998). Social entrepreneurship aims to achieve social innovation (Peredo & McLean, 2006). Founders adopt entrepreneurial techniques to establish and maintain their venture. The goal is to generate social value and achieve a virtuous mission rather than maximising profit to benefit founders (Weerawardena & Sullivan Mort, 2006). This study aimed to document the startup process.

LEXIMANCER

Leximancer is a specialised computer assisted text analysis program to examine printed language such as transcribed interviews or documents (Smith, 2003). A major goal of using Leximancer is to increase the researcher's awareness of the context and significance of concepts as they occur in text (Smith & Humphreys, 2006). Leximancer performs a style of semiautomatic content analysis to examine the structure and content of the actual language used by research participants. The program examines patterns of language through algorithms, machine learning and statistical processes (Smith & Humphreys, 2006). Leximancer provides a system for discovering the underlying core associations within a body of text while reducing biases that may occur with manual coding (Dann, 2009). This helps researchers by avoiding fixation on particular speech topics or patterns that may be atypical.

Leximancer generates concept matrix maps that identify the principal concepts and themes. The size, brightness and centrality of each concept in the map indicates its importance in relation to other concepts. Leximancer provides a visual display of the relationships among themes and concepts whereby if respondents link issues in their interviews, concepts in the map overlap or are located in close proximity. The main concepts are listed in an Entities table as an occurrence percentage. Each entity may be opened to investigate the actual language in the document used by the Leximancer program to compile the concept. Opening the entity allows the researcher to examine meanings of all concepts, the topics being discussed, and how each participant discussed the concept. This enables the

researcher to examine the relationship between the language and other concepts compiled by Leximancer whether displayed in the concept map or not. The mapping process can be controlled to increase or reduce the number of concepts. Increasing the number of concepts increases the complexity and helps understanding by showing the dynamics and relationships of different words. Reducing the number of concepts in the output map helps to show the concepts in a more simplified and visually accessible format. This is useful for visual representation in papers or conference presentations.

METHODOLOGY

Three cases were selected for this study by theoretical sampling and replication logic in accordance with the process described by Eisenhardt (1989). Cases were selected that had started within the previous three years. This helped to ensure data was recent which would improve the authenticity of information. Two digitally recorded interviews were conducted with founders of each case. Interviews ranged in length from 45 minutes to 1.5 hours. Each interview was transcribed verbatim and imported into Leximancer 2.25 as a single file for each case. Each of the three case files was imported into Leximancer 2.25. Standardised settings were used for all analyses. Learning iteration was maintained at 1000 (first learning level). Theme size was standardised at 33% to maintain a high number of concepts without visual complexity. The number of points was set at 3%, and the rotation was not adjusted. This standardisation of analytical settings maintained parity among the cases and allowed effective comparisons. It enabled clear identification and comparison of concepts and themes without confounding overlays of variation in theme size or other notions among the different files.

Leximancer concept maps were generated and used in this study in two waves of data analysis. The first investigated each of the three cases. The resulting Leximancer concept map for each of the three cases were considered. Then the three concept maps were compared and contrasted for conceptual similarities, differences among relationships between the cases. Following this, the transcribed files from the three cases were combined into a single file and imported into Leximancer. The resultant single Leximancer concept map for the combined cases was analysed to consider the startup process

in a more holistic viewpoint and illustrate common concerns among all participants. This combined map was compared and contrasted with each of the single cases.

In the second stage of analysis, each of the interviews was printed and analysed to generate themes in the grounded theory method described by Strauss and Corbin (1990; 1998). It is possible in small studies to generate sets of concepts and related themes with manual coding. Alternatively, many researchers use NVivo for this process to produce open and axial coding. Initially in this study, manual open coding was conducted, then the data were imported into NVivo and a second process of axial coding was conducted. In an iterative process, the Leximancer concept maps were compared and contrasted with the thematic analysis for each case and the combined cases. Unresolved issues were re-examined in a second round of thematic analysis, and compared with the detailed entities listings from the Leximancer analyses to build a strong picture of social entrepreneurship startup.

LEXIMANCER ANALYSIS

The Leximancer concept map of the combined cases demonstrates the overall picture of concerns as identified from the actual language used by the founders (see Figure 1).

INSERT FIGURE 1 ABOUT HERE

This map shows thirteen concepts in three main clusters. A feature of this map is the clear distinction between a few central concepts and a series of disconnected, secondary issues. The map is organised on a vertical axis centred on *think* and *organisation*, dominated by the large balancing concepts of *work* and *people* reflecting the major concerns of these founders.

The cluster of *organisation*, *work* and *think* is central and therefore the most important of the concepts. *Organisation* is located in the centre of the map, indicating the focus of the founders on the organisation as a site to achieve their social mission. *Work* is very large, indicating the tasks involved in starting the venture were viewed as matters of great concern. *Work* overlaps with *years*, indicating that the founders of these ventures understood the long term process of startup and viewed the effort

as *work*. *Think* overlaps with *organisation* and *work* indicating the review and reflection the founders undertook in relation to the *work* of their *organisation*. The association of *sector* and *organisation* suggests that the founders saw themselves as part of a broader social innovation process. *Sector* indicates the overlap of nonprofit and commercial operations involved with the hybrid nature of social entrepreneurship.

A second small cluster is closely associated with *organisation*. *People*, that is, those involved in the venture to deliver services, is a very large concept, indicating their importance in the venture and its work. Unsurprisingly, *people* overlaps with *skills* required to organise and deliver services. A third small cluster is somewhat isolated from other concepts. *Strategic* overlaps with and *year* reflecting the importance these ventures placed on taking a *strategic* approach to their ventures within a multi-year timeframe. Several concepts are isolated. *Business* is quite separate from other concepts, suggesting commercial relationships were not a primary concern of founders. Likewise, the concepts of *group*, *run*, and *back* are isolated and are of secondary importance. The meaning of these isolated concepts is not immediately clear. The associated interview text in the Entities listings was examined (see Figure 2). The entities listings indicated *back* was associated with founders reviewing and reflecting on the startup, *run* was connected with getting the venture ‘up and running’, and *group* was an often discussed concern of one venture that had a small, closely knit group of founders actively involved in the startup.

INSERT FIGURE 2 ABOUT HERE

Overall, the map suggests a strong attention to the *organisation* and its operation, and the *strategic work* required during the startup process. The many disconnected concepts suggest these ventures concurrently attended to numerous tasks, highlighting the complexity of the startup process. This conclusion can be confirmed by opening the entities listings and examining the participants’ statements. The predominant concepts of *People*, *Think*, *Business*, *Work*, and *Organisation* are linked in a multifaceted web of tasks, concerns, and successful outcomes. Founders statements indicated not

everything went according to plan, but overall they organised the startup as a systematic process to achieve the desired objective of social innovation.

Comparing the Leximancer output and thematic analyses revealed an important difference. Each case discussed their relationships with government as a primary concern. Government involvement, or the lack of support, had a significant influence on the startup. In particular, founders viewed government as having the potential to offer funding that might stabilise the venture during the important first years when startup was most precarious. Government was a major source of opportunities in one case and highly valued the service offered. The founder negotiated a substantial three year grant allowing the venture to expand its mission related activities. Government policy also was important for a second case by offering a small annual grants. This was a precarious advantage. The venture was faced with a difficult decision when the government changed and the funding policy along with it. Either the venture could alter its mission to maintain funding, or it could abandon government funding and struggle to maintain its activities independently. Government offered in kind support of office space and advice while the third venture was operating with a small resource base. Having a stable base legitimised the venture and its activities. This was a very important resource that facilitated startup by reducing financial constraints on the venture operation.

Undoubtedly, government was important, yet it did not appear on the Leximancer concept map and 'government' as a term was positioned third lowest in the Entities listing at 14%. Examination of the interview data and entities listing showed interviewees used a variety of terms when discussing government, such as policy, public sector, agencies, Minister, and department. Leximancer offers a capacity to combine terms when they are recognised as having a common association and government would have appeared in a reconfigured map had this been done. This highlights the need to be aware of overlapping terminology and manually reconfiguring the Leximancer analysis process accordingly. This is an easy, but vital process to produce accurate concept maps.

DISCUSSION

Analysis of qualitative data analysis is recognised as a somewhat messy process (Denzin & Lincoln, 2000; Mason, 1996). In the traditional mode of grounded theory originally described by Glaser and Strauss (1968), different data, such as interviews, documents, and websites, may be analysed with traditional manual thematic analysis techniques. To do this, the researcher must become deeply immersed in the data to analyse, understand, interpret and construct meaningful theory (Silverman, 2000). Reading all data closely enables familiarity that assists the researcher to clarify concepts and highlight potentially important issues or discrepancies among participants or documents. In small studies thematic analysis may be conducted manually, but coding is more difficult in large studies where large quantities of data are involved.

Storing documents in electronic programs such as NVivo enables easy storage and rapid retrieval of data. On the other hand, importing data into NVivo places some restrictions on the researcher. Data must be inspected on screen in NVivo and this reduces the capacity for a researcher to become profoundly familiar with the material. Electronic data is less visible and not as easy to inspect. Searching across documents is more difficult when the data is stored electronically. Some qualitative researchers like to spread transcripts over a floor and cut and past or draw lines from one aspect to another so as to connect potentially related concepts in different cases. In this study, the researcher found NVivo was a complex program that was not especially intuitive to use. Instead, researchers need to learn systematically how to use NVivo most productively. This takes considerable time and so is not justified for small research projects. NVivo offers some benefits for researchers, mainly in easy data storage and display, but it has some practical problems.

In contrast, the researcher found Leximancer was very easy to use. Data was easily imported into the program, and concept maps could be produced rapidly even without instruction. Leximancer was especially valuable in the first phase of analysis to formulate concepts rapidly and display these in a way that can be reviewed and analysed quickly. The capacity of Leximancer to display analyses

rapidly as a confirmable format is a distinct benefit of the program. The concept constructions can be verified easily by inspecting the Entities listings. This was helpful in this study to understand how participants were using indeterminate words, such as 'think' or 'back'. NVivo also has this facility to examine raw data, and understanding vague concepts such as these would not be a problem when analysing data in the traditional grounded theory approach when concepts and themes are constructed deliberately by the researcher. Leximancer, however, offers an additional advantage over thematic analysis in NVivo: a capacity to combine data in various ways very quickly for analysis. For example, six interviews in this study were combined to allow individual cases to be compared with other viewpoints. This was a simple process of pasting the six interviews together and importing the file into Leximancer. The resulting map was produced within a couple of minutes, whereas the same process of reviewing data in NVivo would have taken considerable time. Producing analytical maps by combining data in different ways enables the researcher to test assumptions easily and this so is a valuable aspect of Leximancer.

Two additional advantages of Leximancer over NVivo are the authenticity of the analysis, and the visual output as attractive multicoloured concept maps. Leximancer offers a real advantage of providing detached, objective review of the actual language contained in documents or used by study participants in interviews. By interrogating the authentic language, Leximancer improves the potential for researchers to attend core issues in the data rather than becoming distracted by interesting but minor matters. Leximancer produces readily interpretable output from default settings, but it also offers an opportunity to adjust settings and combined concepts to highlight more complex issues if required. Files may be combined easily in different ways for analysis.

An important element of Leximancer is the visual output which offers an interesting, readily accessible, pictorial view of the actual language contained in the research documents. This researcher found the visual concept maps could be readily interpreted to highlight issues in the data. The associational relationships of concepts presented visually in the Leximancer analyses improved

understanding of links among concepts that could be identified in the thematic analysis, but more slowly. A visual output is useful for explanatory display in publications or conference presentations. Presentation of findings is an important aspect of research. Researchers have a responsibility to share the understandings we generate from our investigations via papers and presentations. In general, qualitative findings are presented as thick descriptions, with quotes embedded in the discussion. This approach enables readers to understand a narrative of the findings, and offers an authentic voice of the participants. Narratives, however, are somewhat clumsy to portray in presentations. They may be compiled into association tables, but these are still somewhat static as a way of illustrating associations among themes. A concept map is a way displays core concepts quickly. Representing data visually assists readers to access information readily. The inclusion of Leximancer concept maps adds visual representations and a pictorial dimension to research findings.

Interpreting findings from Leximancer concept maps requires the researcher to delve into the deeper meanings in the data and verify via the Entities listings. This is a form of triangulation. Triangulation assists a researcher to establish authenticity (Denzin & Lincoln, 2000). The study could have reached similar conclusions without a dual analysis, and undoubtedly this would have been a faster way to conclude the study; however triangulating undoubtedly improved the soundness of findings which is very important for publication. The dual analysis processes in this study of Leximancer and thematic analysis in NVivo improved the soundness of the findings. Combining visual and textual interpretation allowed concepts to be viewed in different ways, thereby improving the researcher's interpretation. Comparing the two analyses enabled a review to confirm findings, or raise new issues that were not apparent. In this study, it was a concern initially that some concepts identified in the thematic analysis were not present in the Leximancer concept maps. The capacity to analyse the actual language in Entities listings added new understandings. Thus, interpreting Leximancer concept maps needs to proceed with care and diligence.

CONCLUSION

Leximancer analysis provided a considerable benefit for this study. It increased the quality of the analysis, added understandings that could be verified through thematic analysis, improved the dependability of the findings, and provided valuable visual output for reporting. Analysing narrative data in the Leximancer system provided useful insights in this study that would have been slower to locate, or may not have been as clear if other techniques had been used. An important aspect of Leximancer as a qualitative research tool is its ease of use. Leximancer analyses are generated rapidly. The same can not be said for NVivo as a research tool. Despite its many advantages, NVivo is not an intuitive tool to use. Researchers must learn how to use it, and it takes much longer to conduct analyses in NVivo than Leximancer.

As a data analysis tool, Leximancer improves the richness of understandings that may come from a study. It contributes to the capacity for the researcher to generate concepts rapidly or confirm grounded thematic analyses. In addition to providing a valuable means of data analysis, Leximancer offers an important communication tool. Leximancer concept maps are especially valuable as a means of presenting easily accessible research results that improves the capability of researchers to reach their audience. Qualitative research may be enhanced by using Leximancer at different stages of the research analysis process. Leximancer allows a rapid review of concepts early in the analysis. Later in the research process Leximancer offers a triangulation capacity to review concepts against other forms of analysis thus enriching the interpretation of findings. As an easy to use content analysis tool, Leximancer could be used more widely for to greater benefit in qualitative research.

REFERENCES

- Chell, E. (2007). Social enterprise and entrepreneurship - towards a convergent theory of the entrepreneurial process. *International Small Business Journal*, 25(1), 5-26.
- Creswell, J. W. (1998). *Qualitative inquiry and research design: choosing among five traditions*. London: Sage.
- Dann, S. (2009). Redefining social marketing with contemporary commercial marketing definitions. *Journal of Business Research*, *In press*.
- Denzin, N. K., & Lincoln, Y. S. (2000). *Handbook of qualitative research* (2nd ed.). Thousand Oaks, Ca: Sage.
- Denzin, N. K., & Lincoln, Y. S. (2003). *Strategies of qualitative inquiry* (2nd ed.). Thousand Oaks, Calif: Sage.
- Eisenhardt, K. M. (1989). Building theories from case study research. *Academy of Management Review*, 14(4), 532-550.
- Glaser, B. G., & Strauss, A. L. (1968). *The discovery of grounded theory: strategies for qualitative research*. London: Weidenfeld and Nicolson.
- Haugh, H. (2007). Community-led social venture creation. *Entrepreneurship Theory and Practice*, 31(2), 161-182.
- Hine, D., & Carson, D. (Eds.). (2007). *Innovative methodologies in enterprise research*. Cheltenham, UK: Edward Elgar.
- Mason, J. (1996). *Qualitative researching*. London: Sage.
- Miles, M. B., & Huberman, M. A. (1994). *Qualitative data analysis: an expanded sourcebook* (2nd ed.). Thousand Oaks: Sage.
- Nicholls, A. (Ed.). (2006). *Social entrepreneurship: new models of sustainable social change*. Oxford: Oxford University Press.
- Peredo, A. M., & McLean, M. (2006). Social entrepreneurship: A critical review of the concept. *Journal of World Business*, 41(1), 56-65.
- Silverman, D. (2000). *Doing qualitative research: a practical handbook*. London: Sage.
- Silverman, D. (2001). *Interpreting qualitative data: methods for analysing talk, text and interaction*. London: Sage.
- Smith, A. E. (2003). *Automatic extraction of semantic networks from text using Leximancer. Demonstration paper presented at the HLT-NAACL Conference, Edmonton, Canada. May-June.*
- Smith, A. E., & Humphreys, M. S. (2006). Evaluation of unsupervised semantic mapping of natural language with Leximancer concept mapping. *Behavior Research Methods*, 38(2), 262-279.
- Stake, R. E. (1995). *The art of case study research*. Thousand Oaks, Calif: Sage.
- Stake, R. E. (1998). Case studies. In N. K. Denzin & Y. S. Lincoln (Eds.), *Strategies of qualitative inquiry* (pp. 86-109). Thousand Oaks, Ca: Sage.
- Strauss, A. L., & Corbin, J. M. (1990). *Basics of qualitative research: grounded theory procedures and techniques*. Newbury Park, Ca: Sage.
- Strauss, A. L., & Corbin, J. M. (1998). *Basics of qualitative research: techniques and procedures for developing grounded theory* (2nd ed.). Thousand Oaks, Ca: Sage.
- Sullivan Mort, G., Weerawardena, J., & Liesch, P. (2008). *Strategising for international market entry in born global firms: An analysis of narratives using Leximancer. ANZMAC Conference, 13-Dec, Sydney.*
- Sullivan Mort, G., Weerawardena, J., & Williamson, B. (2007). Branding in the Non-profit Context: The Case of Surf Life Saving Australia. *Australasian Marketing Journal*, 15(2), 108-119.
- Weerawardena, J., & Sullivan Mort, G. (2006). Investigating social entrepreneurship: a multidimensional model. *Journal of World Business*, 41(1), 21-35.
- Yin, R. K. (2003). *Case study research design and methods*. Thousand Oaks, Ca: Sage.

Figure 1: Leximancer concept map of the combined cases

Iterations = 1000

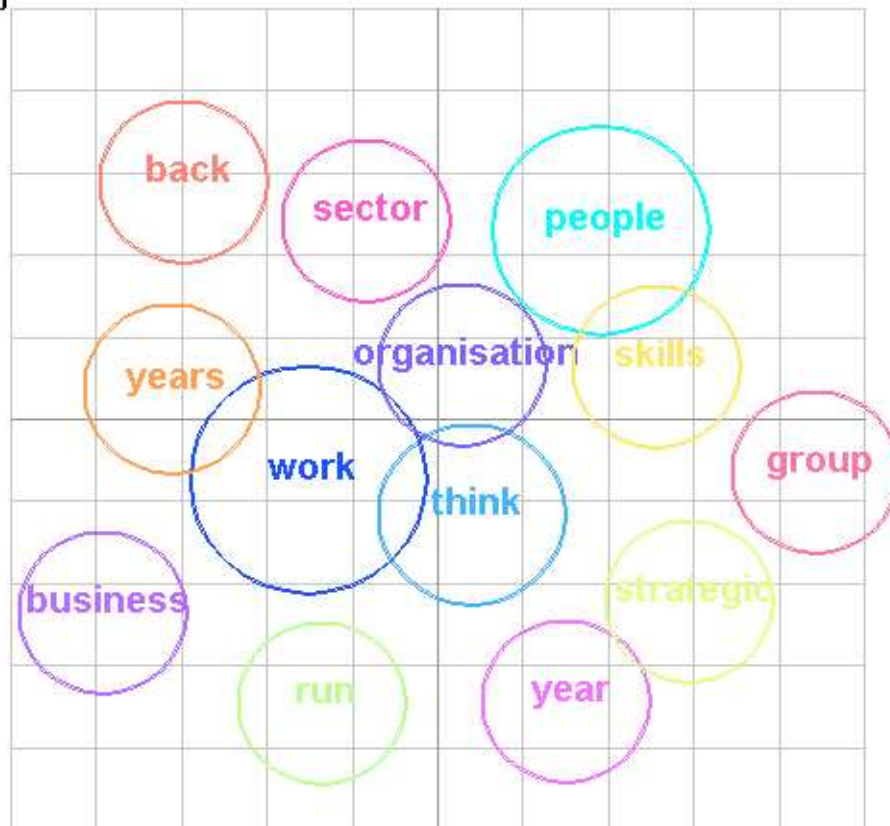


Figure 2: Entities listing combined continuing organisations

Concept	Absolute Count	Relative Count	
people	250	100%	
think	211	84.3%	
business	134	53.6%	
work	127	50.8%	
organisation	124	49.6%	
year	90	36%	
time	89	35.6%	
group	82	32.8%	
back	75	30%	
sector	74	29.6%	
years	57	22.8%	
person	57	22.8%	
skills	55	22%	
board	51	20.4%	
money	48	19.2%	
organisations	47	18.8%	
talk	47	18.8%	
strategic	46	18.4%	
started	45	18%	
profit	44	17.5%	
project	43	17.2%	
talking	42	16.8%	
structure	41	16.4%	
working	40	16%	
kind	40	16%	
involved	40	16%	
run	39	15.6%	
months	39	15.6%	
running	39	15.6%	
set	39	15.6%	
income	38	15.2%	
government	36	14.4%	
worked	36	14.4%	
few	29	11.6%	