

Edith Cowan University
School of Business and Law



ANZAM 2021 Doctoral Workshop

**The use of Leximancer in academic research:
Qualitative data analysis exemplars.**



Presented by: Dr Violetta Wilk
Lecturer & Researcher in Digital Marketing

What is Leximancer?

Welcome



Turn your text insight out.



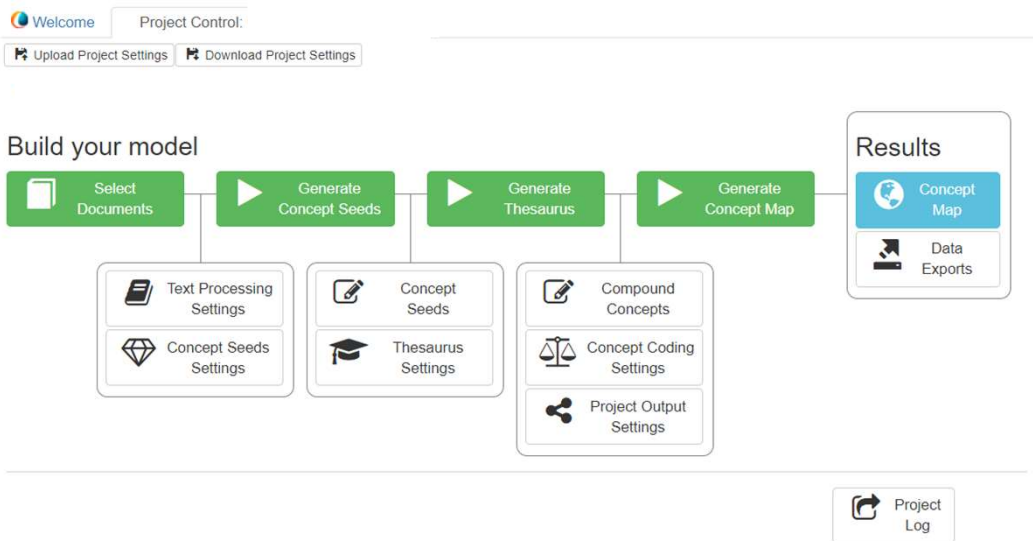
Getting
Started



User
Guide

- ❑ Created by the University of Queensland, Dr Andrew Smith and his team.
- ❑ Automated, machine-learning driven qualitative data analysis.
- ❑ Uses blocks of text to identify concepts and themes through an iterative process of seeding word definitions from frequencies and co-occurrences.
- ❑ Files: MS Word, pdf, html, csv, plain text.
- ❑ Literature review, Annual reports, Policy documents, Open ended questions, Social media monitoring, Interview transcripts etc.

Leximancer Interface



Comparison of Leximancer with QSR NVivo

QSR NVIVO	LEXIMANCER
Strengths	
<ul style="list-style-type: none"> • Researcher-driven, at researcher's discretion which nodes to identify for data classification. This may be perceived as specific and efficient, in identifying only those concepts, themes or keywords, which the study is interested in exploring. • Code and analyse data to specific concepts. As above. • Suitable for interpretative approach, in studies where a specific conceptual model is of interest to be investigated and the concepts/constructs are known. • Researcher assigns meaning to the data at coding stage. • Akin to manual handling of data. A manual way of coding with the aide of a computer program. • Content analysis, data linking and data display based on nodes (codes). • Various analyses and visual output produced (e.g. Word Frequency Query, Coding Stripes Analysis, Matrix Coding Query). • Linking function allowing access from nodes to the original data. 	<ul style="list-style-type: none"> • Program-driven, automated analysis which may be perceived as more objective. • Automatic identification of key words (concepts) and clusters (themes). • Seeding based on frequencies and co-occurrences of words (concepts). • Suitable for exploratory study as "themes" emerge via automated lexical analysis. • Content analysis, data linking and data display based on emergent themes and concepts. • Efficient for large volumes of data. Quick, automated analysis. • Researcher able to manually seed (define) concepts required for the program to identify. Akin to setting up queries. • Program develops Concept Map and Insight Dashboard with Prominence Scores, to highlight key "themes" and "concepts" within them. • Linking function allowing access from the Concept Map to the original data.

Wilk, V., Soutar, G. N., & Harrigan, P. (2019). Tackling social media data analysis: Comparing and contrasting QSR NVivo and Leximancer. *Qualitative Market Research: An International Journal*. 22 (2), pp. 94-113.

Comparison of Leximancer with QSR NVivo

QSR NVIVO	LEXIMANCER
Weaknesses	
<ul style="list-style-type: none"> • Subjective and researcher bias possible. Limited by researcher's epistemological position. • Time consuming in identifying what the concepts could be, thus what the nodes should be, and in the coding of data. • Questionable reliability. Due to the researcher's heavy involvement through the whole process, it is arguable to what extent the results are reliable. • <u>Auto-Coding</u>, should be used with caution. Albeit efficient, did not prove to be effective at capturing the correct information required for the purposes of this study, resulting in manual coding of data. 	<ul style="list-style-type: none"> • Lack of human insight during lexical analysis which is driven solely by the program. • Input data needs specific formatting and spelling checking prior to input into the program, which may be time consuming for large quantities of data. The program will only recognise correctly spelt words, and in the right format. • Lexical analysis occurs in 2 sentence blocks, which may be adjusted up or down. This is problematic in particular for data sourced online, where, unlike any other written form, the form of online expression can be short or long winded. • Researcher assigns meaning after program-analysis (i.e. at Concept Map configuration stage). • Sentiment lens is only suitable for at best compound concept (two 'affective' words) type analysis but it is not suitable for in-depth affective-type analysis. • Unable to capture the online communication style (e.g. acronyms) or implied tone of voice (e.g.!!!), common to online written form. • "Theme" names after the most prominent concept but do not necessarily represent the essence of the other "concepts" within the theme. • Unexpected or unexplained emergent concepts and relationships. • Misleading terminology, i.e. "concepts" = keywords and "themes" = concepts. • Somewhat awkward and problematic researcher-driven identification of the themes' size, the number of clustering attempts and rotation of the Concept Map.

Wilk, V., Soutar, G. N., & Harrigan, P. (2019). Tackling social media data analysis: Comparing and contrasting QSR NVivo and Leximancer. *Qualitative Market Research: An International Journal*. 22 (2), pp. 94-113.

What types of data can you analyse in Leximancer?

- Journal articles
- Annual reports
- Policy documents
- Open ended questions' answers
- Interview transcripts
- Social media posts and comments
- Big unstructured data
- Any text-based data

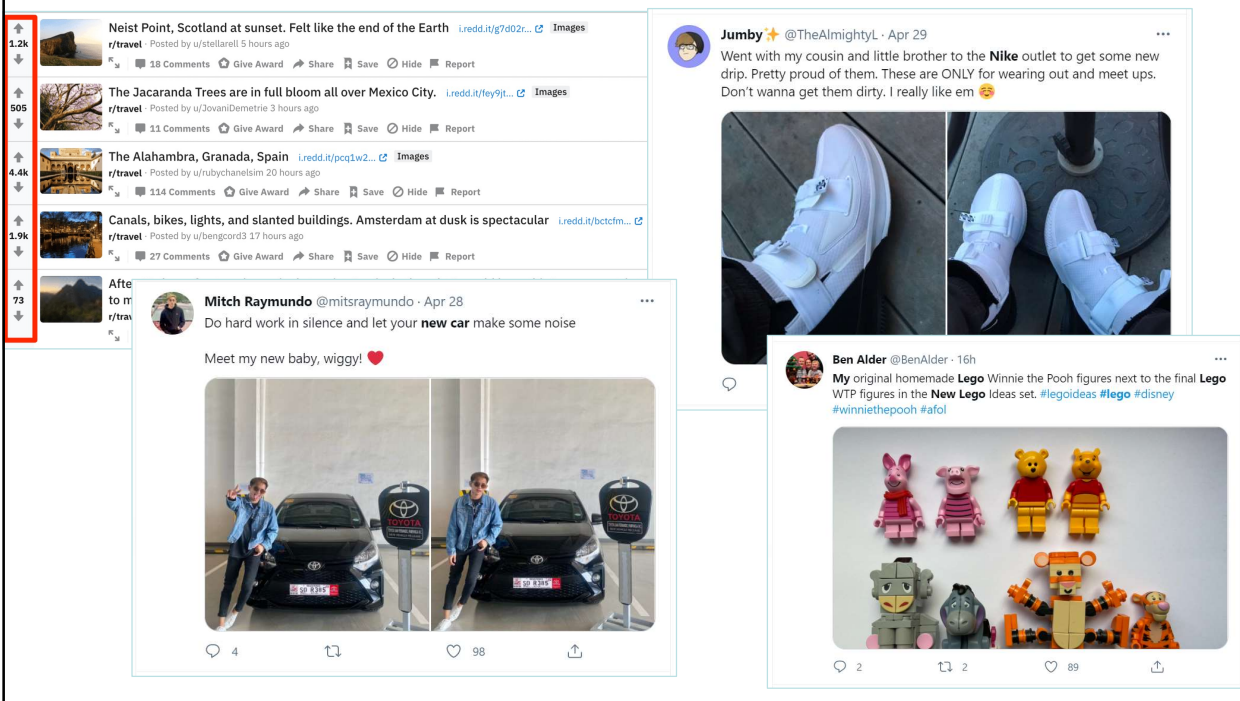


What is Big Unstructured Data?

- Big data is characterised by:
 - volume (the amount of data being managed),
 - velocity (the rate at which the data is created),
 - variety (the form of data); and
 - veracity (the validity of the data) (White, 2016).
- Unstructured data is information that either does not have a pre-defined data model or is not organized in a pre-defined manner (Balducci & Marinova, 2018).
- Much of big unstructured data is generated via digital touchpoints, such as user-generated content (UGC) (Flynn et al., 2015).

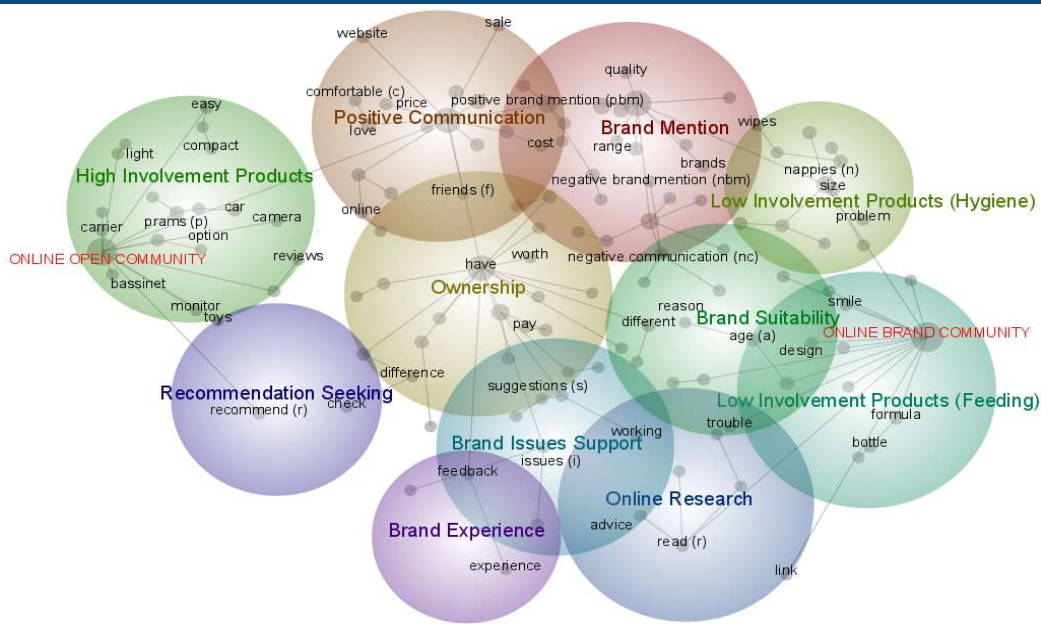


User-Generated Content (UGC) on Social Media is one of many sources of Big Unstructured Data



The screenshot displays a collage of social media content. On the left, a vertical list of travel-related posts from the subreddit r/travel is shown, with a red box highlighting the first post: "Neist Point, Scotland at sunset. Felt like the end of the Earth". Below this, a tweet from Mitch Raymundo (@mitsraymundo) dated April 28 features a photo of a man with a new Toyota car and the text "Do hard work in silence and let your new car make some noise". To the right, a tweet from Jumbo (@TheAlmightyL) dated April 29 shows two photos of white sneakers with the text "Went with my cousin and little brother to the Nike outlet to get some new drip...". Below that, a tweet from Ben Alder (@BenAlder) dated 16h shows a collection of custom-made LEGO figures of Winnie the Pooh characters.

Content Analysis through Big Data Visualisation in Leximancer



Wilk, V., Harrigan, P., and Soutar, G., (2019), Navigating Online Brand Advocacy (OBA): An exploratory analysis, *Journal of Marketing Theory & Practice*, 26 (1-2), pp. 99-116.

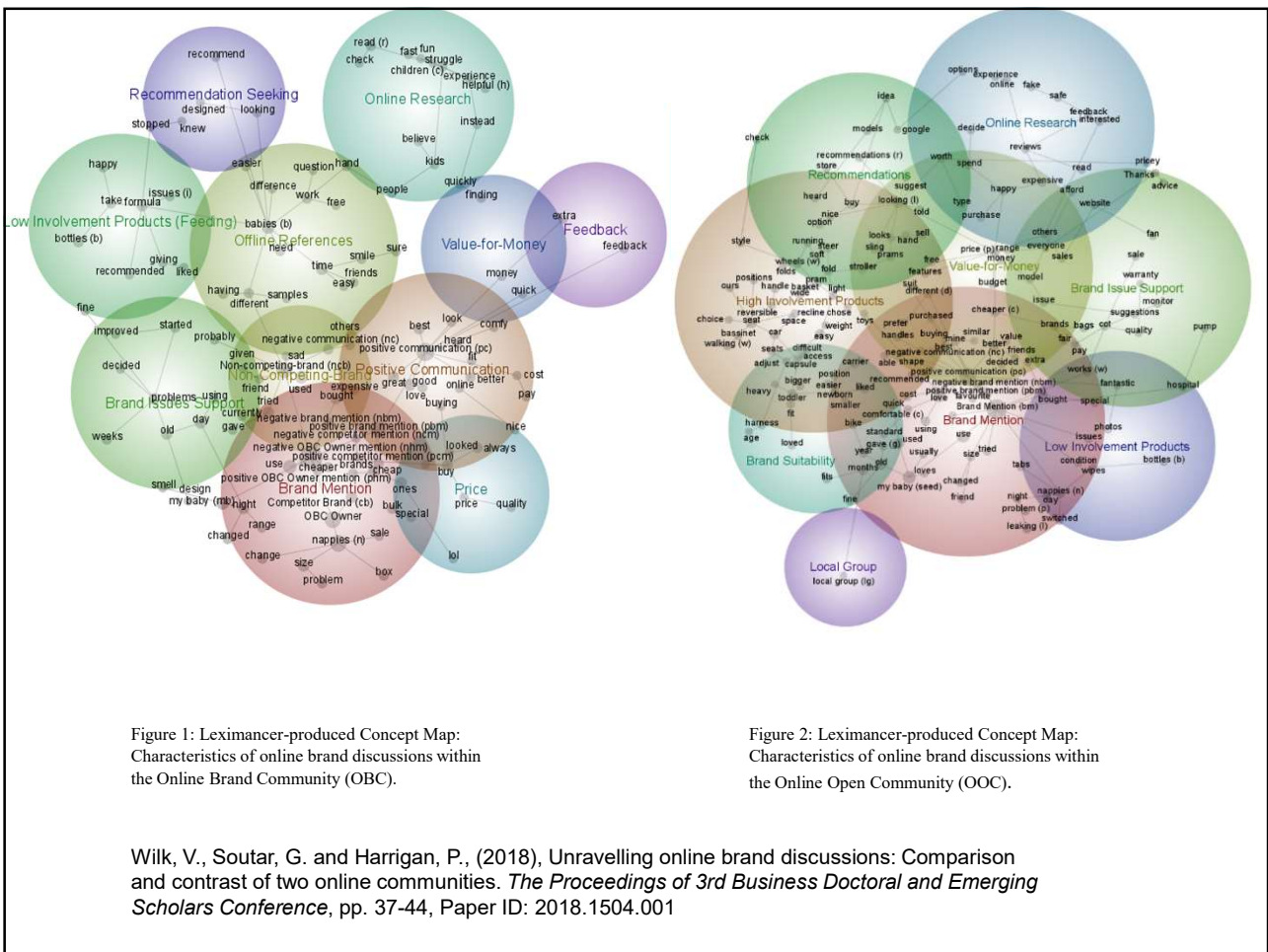


Figure 1: Leximancer-produced Concept Map: Characteristics of online brand discussions within the Online Brand Community (OBC).

Figure 2: Leximancer-produced Concept Map: Characteristics of online brand discussions within the Online Open Community (OOC).

Wilk, V., Soutar, G. and Harrigan, P., (2018), Unravelling online brand discussions: Comparison and contrast of two online communities. *The Proceedings of 3rd Business Doctoral and Emerging Scholars Conference*, pp. 37-44, Paper ID: 2018.1504.001

Leximancer: Insight Dashboard Report output

Brand mention		OBC		OOC	
Concept	PS	Concept	PS	Concept	PS
Tried	1.7	OBC brand name	1.7	Love	1.4
Use	1.6	Read	1.7	Price	1.3
Love	1.4	Tried	1.7	Looking	1.3
Better	1.4	Problem	1.6	Bought	1.1
Buy	1.3	Best	1.4	Great	1.1

Table III.
Top ranking concepts and their prominence scores (PS) against three categories of interest: Brand mention; OBC; and OOC

Brand mention		OBC		OOC	
Compound concept	PS	Compound concept	PS	Compound concept	PS
Tried and brands	17.4	Read and understand	22.0	Easy and fold (product functionality)	39.3
Positive brand mention and negative brand mention	16.8	Tried and brands	12.1	Price and range	11.5
Better and cheaper	11.6	Best and brands	8.4	Looking and reviews	8.5
Positive brand mention and reviews	10.0	OBC brand name and brands	8.1	Issues and pay	8.2
Problem and never	9.3	Positive brand mention and negative brand mention	7.9	Recommend and compact (product feature)	7.6

Table IV.
Top ranking compound concepts and their prominence scores (PS) against three categories of interest: Brand mention; OBC; and OOC

Wilk, V., Soutar, G. N., & Harrigan, P. (2019). Tackling social media data analysis: Comparing and contrasting QSR NVivo and Leximancer. *Qualitative Market Research: An International Journal*, 22 (2), pp. 94-113.

Example: Leximancer Theme Ranking

Table 3: Comparison of the characteristics of online brand discussions between the two online community types (OBC and OOC).

Theme Ranking	OBC Themes	OOC Themes
1	Brand Mention	Brand Mention
2	Positive Communication	High Involvement Products
3	Non-Competing Brands	Value-for-Money/Price
4	Offline References	Brand Issue Support
5	Brand Issues Support	Recommendations
6	Low Involvement Products	Brand Suitability
7	Online Research	Online Research
8	Price/ Value-for-Money	Low Involvement Products
9	Recommendation Seeking	Local Group
10	Feedback	-

Wilk, V., Soutar, G. and Harrigan, P., (2018), Unravelling online brand discussions: Comparison and contrast of two online communities. *The Proceedings of 3rd Business Doctoral and Emerging Scholars Conference*, pp. 37-44, Paper ID: 2018.1504.001

Edith Cowan University
School of Business and Law

ot: Cricket Australia (Crisis Week) * Map Explorer: Cricket Australia (Crisis Week) * Project Control: Digital Entrepreneurship 4 * Map Explorer: Digital Entrepreneurship 4 *

Concepts:

Synopsis Concepts Thesaurus Query Summaries Log

Analyst Synopsis Detail Level - Spreadsheet (CSV) Export

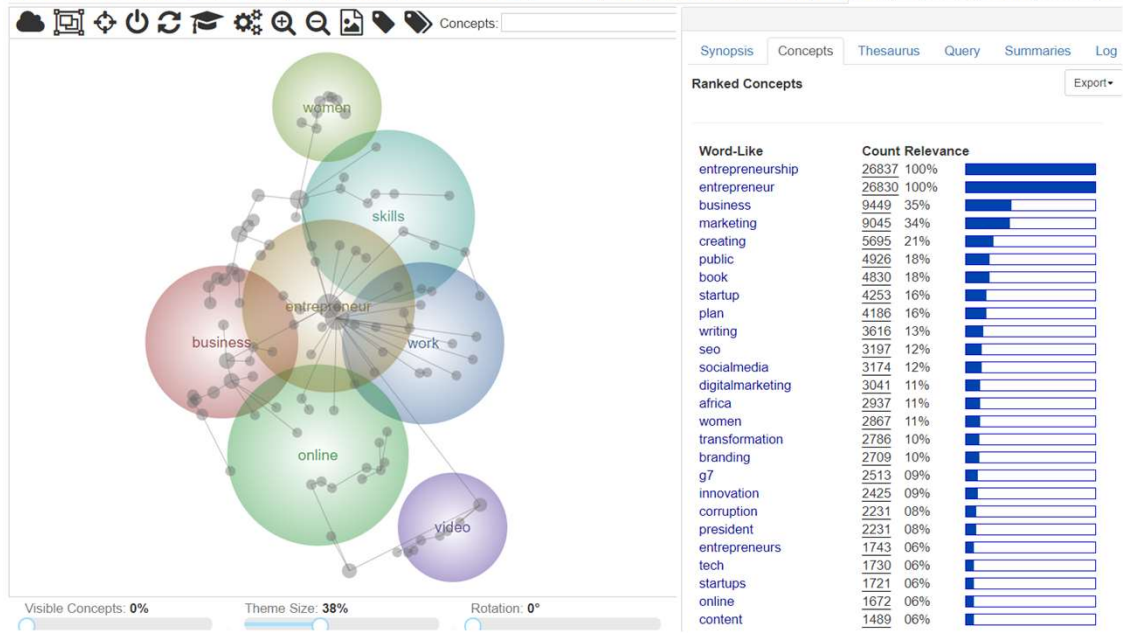
Theme	Hits
entrepreneur	27028
business	15384
online	7099
women	5593
skills	1336
video	972
work	838

Theme: **entrepreneur**
 Concepts: entrepreneur, entrepreneurship, innovation, tech, technology, training, development, world, build
 Hits: **27028**

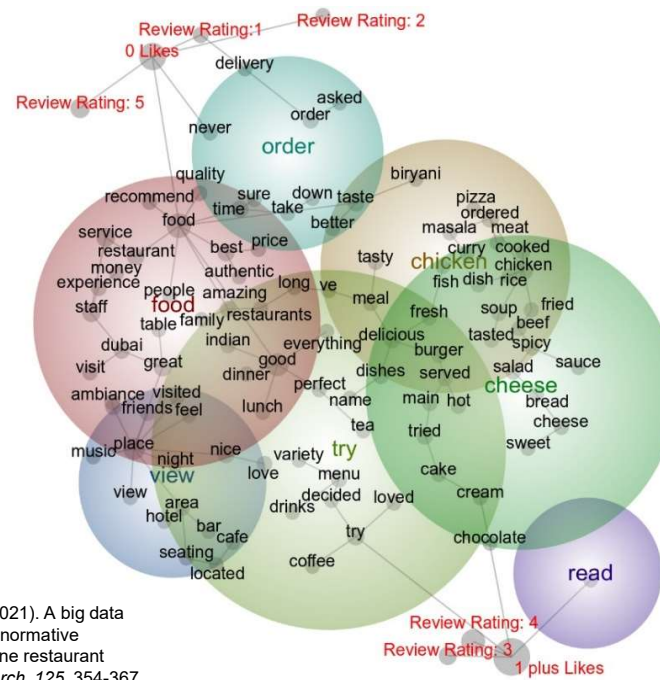
COMMENT_COUNT: 0
 SENTIMENT: positive
 FOLLOWERS: 0
 FOLLOWING: 0
 VIEW_COUNT: 0
 COLUMN: content
 understanding of the operational activities of the Navies of two countries and to get acquainted with advanced technology through various theoretical and practical training. 8) MSDE, IBM jointly launch SkillsBuild platform The Ministry of Skill Development & Entrepreneurship (MSDE) launched the SkillsBuild platform in...
 Opositive00content

Visible Concepts: 0%
Theme Size: 38%
Rotation: 0°

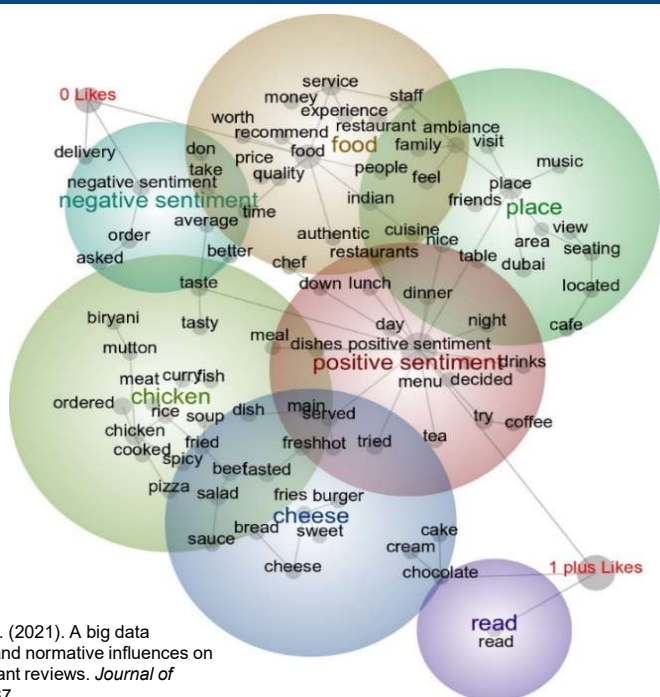
o: Cricket Australia (Crisis Week) * Map Explorer: Cricket Australia (Crisis Week) * Project Control: Digital Entrepreneurship 4 * Map Explorer: Digital Entrepreneurship 4



Leximancer: The use of tags to understand emergent themes



Leximancer: The use of tags to understand emergent themes



Meek, S., Wilk, V., & Lambert, C. (2021). A big data exploration of the informational and normative influences on the helpfulness of online restaurant reviews. *Journal of Business Research*, 125, 354-367.

Leximancer-produced Insight Dashboard results

Comparison of top-ranking compound concepts between reviews with “no Likes” and “Like(s)” and positively and negatively framed reviews.

Compound concept “no Likes”	PS “no Likes”	Compound concept “Like(s)”	PS “Like(s)”
Order & delivery	11.1	Dubai & hotel	9.6
Order & asked	7.3	Restaurant & hotel	7.5
Order & take	6.9	Try & decided	6.7
Time & long	6.9	Menu & variety	5.3
Time & delivery	6.5	Dubai & best	4.9
Compound concept ‘Positive Sentiment’	PS	Compound concept ‘Negative Sentiment’	PS
‘best & (in) Dubai’	48.3	‘bad & experience’	307.3
‘amazing & view’	37.5	‘bad & service’	205.4
‘best & restaurant’	29.0	‘bad & delivery’	204.6
‘delicious & soft’	27.9	‘tasted & bad’	162.1

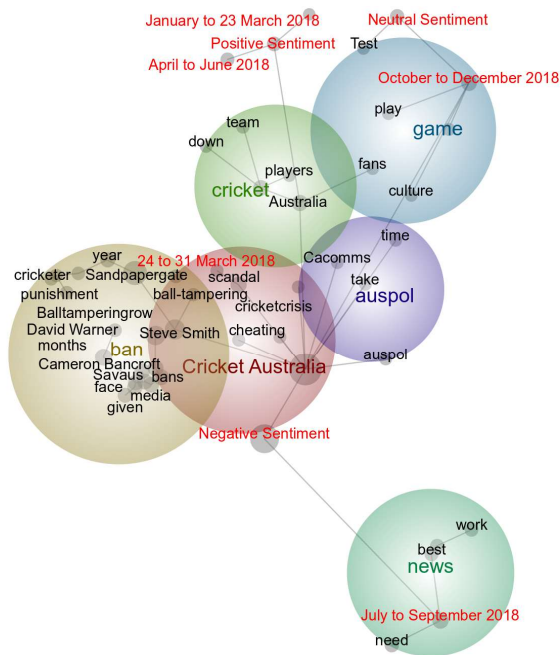
Meek, S., Wilk, V., & Lambert, C. (2021). A big data exploration of the informational and normative influences on the helpfulness of online restaurant reviews. *Journal of Business Research*, 125, 354-367.

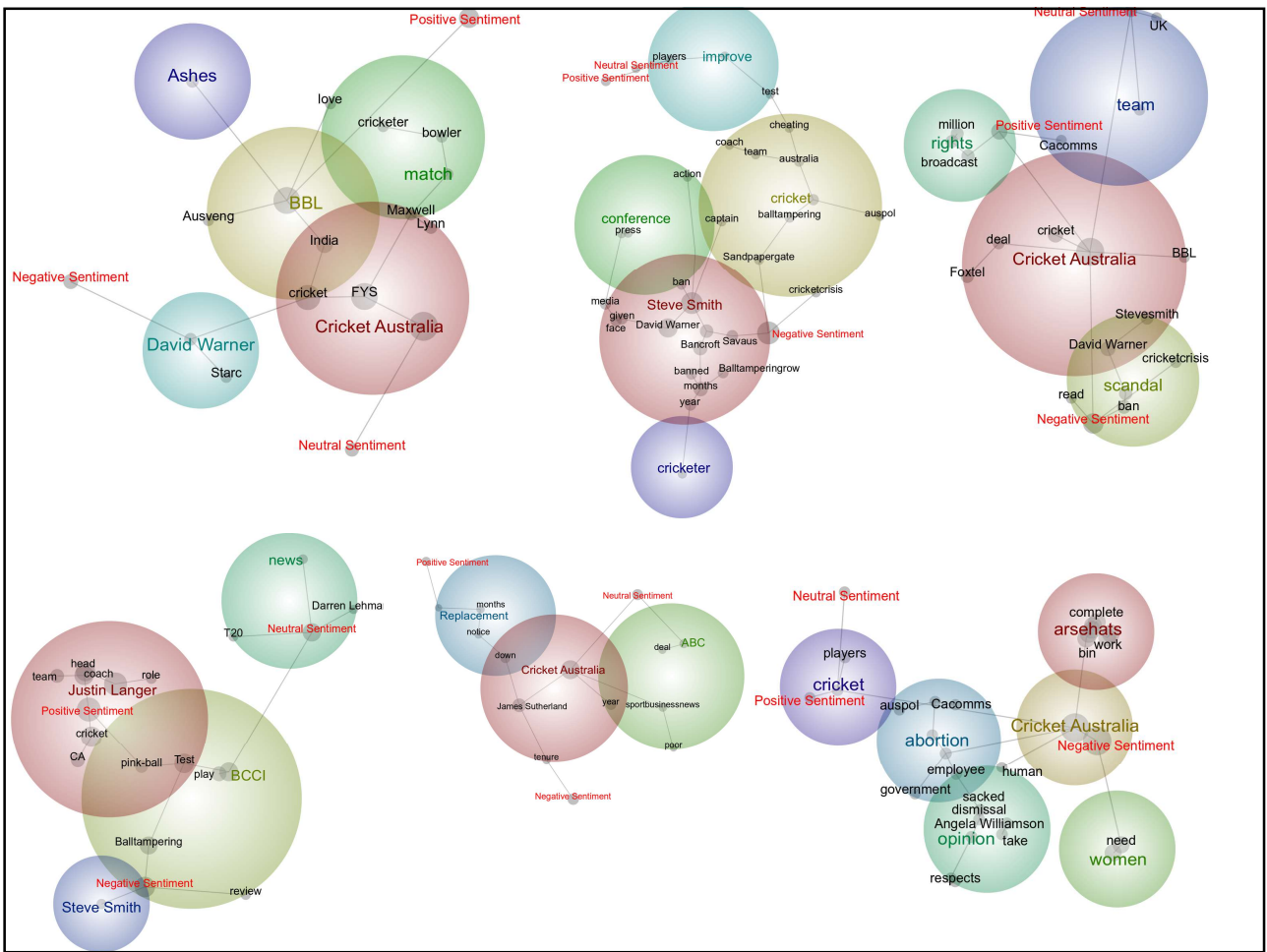
Examples of content to supplement Leximancer-driven findings

Review	Content
# 3446	<i>Great place and good food quality but delivery time takes too long. (Received '0' "Likes")</i>
# 8	<i>Good south Indian food with reasonable prices. Good taste, good seating arrangement and good service. I think no better keralian restaurant than this in terms of quality and prices. (Received '0' "Likes").</i>
# 10	<i>...I thought to give this place a try and I must say I am not disappointed at all with the food and service here, if you are nearby, do try this place. Ambiance is more of mix of colors; they have used green color with brown, which highlights the decor. They have private seating and corporate seating as well, which I found very innovative. Service is prompt and staff is helpful and polite. Other thing impressed me here is the menu; menu is mix of many must have Indian dishes, also some unique dishes. So in every dish it focuses on authenticity and very Indian touch, so even if the innovation is done with names". (Received '32' "Likes").</i>
# 58157	<i>Having their Kifaya pastry was such a splendid experience. This soft and moist cake was light green in colour and topped with fruits. Definitely recommended. (Received '17' "Likes").</i>
# 20	<i>Awesome Hyderabad Dum Biryani... Placed an order and delivered on the said 15 mins. (Received '0' "Likes").</i>
# 8087	<i>Went there with a group of friends for dinner on Thursday night. We took the set menu. However, most of us were vegetarians and few Vegan as well. The interiors of this place are very royal and so is the service...made us feel very special with their extremely warm service. Chef Pradeep Khullar also catered to Vegan dishes even though it was not on the menu.... The main course and the desserts made our complete meal. The entire experience was memorable. A must visit place! (Received '13' "Likes").</i>
# 12	<i>Order came wrong, chicken is not that good when compared to other restaurants that offer the same meals with same price. coleslaw is just beyond terrible and fries are somehow dry and not good....fast delivery and terrible experience. (Received '0' "Likes").</i>
# 3588	<i>Horrible experience! I should have taken the empty tables and chairs on a Thursday night as a warning sign. They couldn't even get the most basic of dishes right and I won't be coming back to see what their main dishes would be like... When I wanted to ask for the bill, two of them acknowledged me yet no one attended to me. Would not recommend and will definitely not be coming back. (Received '2' "Likes").</i>

Meek, S., Wilk, V., & Lambert, C. (2021). A big data exploration of the informational and normative influences on the helpfulness of online restaurant reviews. *Journal of Business Research*, 125, 354-367.

Morgan, A., & Wilk, V. (2021) Social Media Users' Crisis Response: A lexical exploration of social media content in an international sport crisis, Public Relations Review. (ahead-of-print).





Morgan, A., & Wilk, V. (2021) Social Media Users' Crisis Response: A lexical exploration of social media content in an international sport crisis, Public Relations Review. (ahead-of-print).

Table A1: Prominence Scores (PS) for Singular Concepts for each time period, by sentiment type.

Sentiment (Concepts)	Pre-Crisis: January-23 March 2018	Crisis Week: 24-31 March 2018	Post Crisis: April 2018	Post Crisis: May 2018	Post Crisis: June 2018	Post Crisis: July-September 2018	Post Crisis: October-December 2018
Negative	"David Warner" (3.3) "Starc" (1.6) "Cricket Australia" (1.5)	"cricketcrisis" (1.3) "balltamperingrow" (1.2) "savvas" (1.2)	"scandal" (1.7) "ban" (1.7) "David Warner" (1.3) "Steve Smith" (1.2)	"Steve Smith" (2.8) "balltampering" (2.5) "Cricket Australia" (1.4)	"Tenure" (2.1) "James Sutherland" (1.3) "Cricket Australia" (1)	"Sacked" (1) "Angela Williamson" (1) "CACOMMS" (1)	"damning" (2) "slammed" (2) "Cricket Australia" (1.9) "review" (1.7)
Positive	"Maxwell" (1.9) "bowler" (1.9) "match" (1.9)	"players" (3.2) "improve" (2.6) "press" (2.9)	"million" (3.5) "Broadcast" (3.5) "rights" (3) "CACOMMS" (2.5)	"pink-ball" (4) "team" (4) "Justin Langer" (1.8)	"replacemement" (4) "months" (1.7) "notice" (1.6)	"Players" (16.1) "Cricket" (9.5) n/a	"game" (2.1) "time" (1.3) "play" (1.2)
Neutral	"Cricket Australia" (1.5) "India" (1.5) "FYS" (1.4)	"coach" (1.9) "conference" (1.1) "Auspol" (1)	"UK" (6.3) "team" (3.2) "Cricket Australia" (1)	"Darren Lehmann" (2.5) "news" (2.5) "BCCI" (1.3)	"deal" (2.5) "ABC" (2.4) "sportsbusinessnews" (1.3)	"Players" (11) "Cricket" (8.7) n/a	"year" (2.5) "Virat Kohli" (1.6) "game" (1.5)

Table A2: Prominence Scores (PS) for Compound Concepts for each time period, by sentiment type.

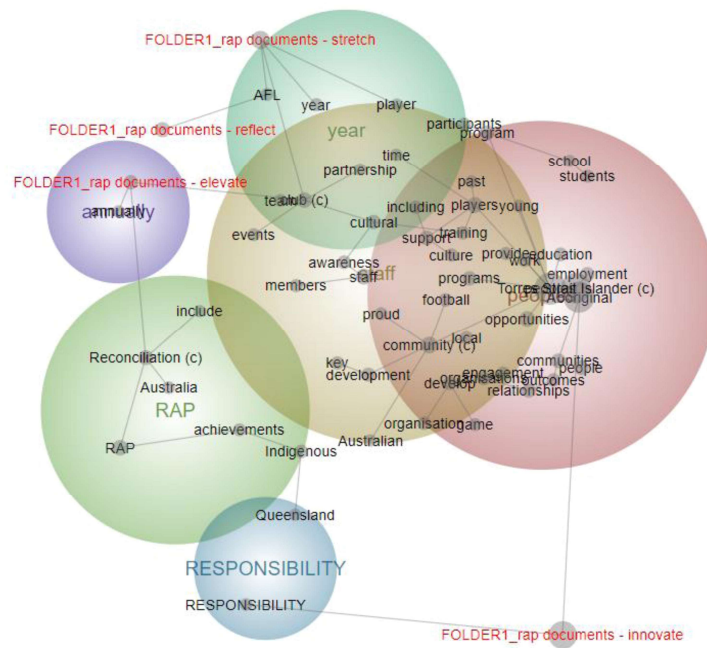
Sentiment (Compound Concepts)	Pre-Crisis: January-23 March 2018	Crisis Week: 24-31 March 2018	Post Crisis: April 2018	Post Crisis: May 2018	Post Crisis: June 2018	Post Crisis: July-September 2018	Post Crisis: October-December 2018
Negative	"Starc & ausveng" (48.3) "David Warner & India" (8.1)	"cricketcrisis & Auspol" (18.4) "given & face" (14.3) "banned & balltamperingrow" (8.2)	"scandal & ban" (15.5) "ban & David Warner" (11.5)	"Steve Smith & Cricket Australia" (11.3) "Steve Smith & balltampering" (8.5)	"tenure & James Sutherland" (4.9) "tenure & Cricket Australia" (2.4)	"Opinion & Making" (34.9) "Angela Williamson & Dismissal" (33)	"win & pressure" (54.2) "Cricket Australia & cheating" (46.2)
Positive	"Starc & cricket" (3.1) "Maxwell & Lynn" (37.4) "bowler & match" (28.1) "Maxwell & match" (18.7)	"press & conference" (129.6) "players & improve" (60.5) "players & ban" (6.3)	"broadcast & million" (55.7) "broadcast & rights" (47.2) "CACOMMS & deal" (16.1)	"BCCI & test" (4.8) "coach & head" (30.7) "pink ball & test" (27.4) "coach & Justin Langer" (17.1)	"James Sutherland & Cricket Australia" (1.5) "replacemement & months" (41.1) "replacemement & notice" (38.4) "replacemement & James Sutherland" (4.6)	"CACOMMS & Auspol" (32.9) "Players & Cricket Australia" (16.5) "Cricket & Cricket Australia" (9.7)	"play & pressure" (39.1) "time & play" (17) "report & Cricket Australia" (6.5) "play & players" (5.7)
Neutral	"Cricket Australia & India" (5.8) "cricket & FYS" (3.9) n/a	"Press & conference" (38) "Coach & players" (22) "test & improve" (14.5)	"team & Cricket Australia" (3.9) n/a n/a	"Darren Lehmann & news" (12.1) "news & T20" (6.7) "BCCI & news" (3.7)	"deal & ABC" (18) "ABC & sportsbusinessnews" (4.8) "deal & Cricket Australia" (2.9)	"Players & Cricket" (39.8) "Cricket & Auspol" (38.8) "Players & Cricket Australia" (12.1)	"Virat Kohli & year" (33.6) "culture & review" (6.6) "players & culture" (3.5)

Morgan, A., & Wilk, V. (2021) Social Media Users' Crisis Response: A lexical exploration of social media content in an international sport crisis, *Public Relations Review*. (ahead-of-print).

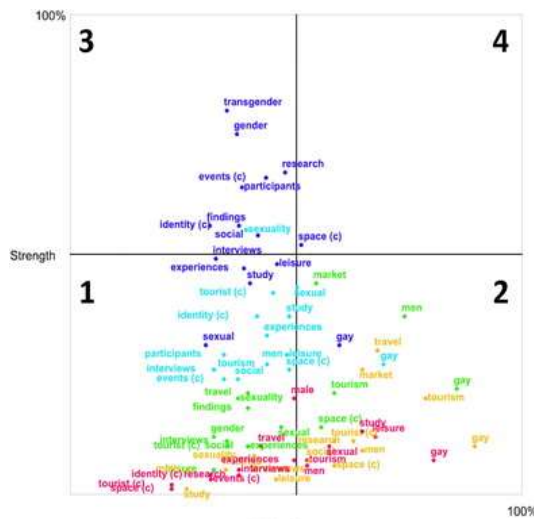
Table 2: Social Media Users' Crisis Response Framework (SMUCR)

Stage	Description
Framing and Identity Development	<ul style="list-style-type: none"> ▪ Giving crisis an identity (labelling it by a descriptive name) ▪ Activating the online "tribe" ▪ Villainising: naming and attacking suspected culprits
Sense Making	<ul style="list-style-type: none"> ▪ Individual culprit blame (focus on specific individuals involved) ▪ Spreading "truth" (brand protection) ▪ Broadening the discussion beyond the crisis ▪ Accepting the brand's response
Accountability	<ul style="list-style-type: none"> ▪ Brand accountability ▪ Focus shifts to the changes occurring or being implemented by the brand
Vigilant Rebuilding	<ul style="list-style-type: none"> ▪ Criticism of any emergent instability with the brand ▪ Acknowledging what has happened and accepting the fall out ▪ Cautiously moving forward and optimism for the future ▪ Brand in the spotlight and under on-going public scrutiny ▪ Amplification of any future brand-related issues and crises

Morgan, A. & Wilk, V. (2021), Sports organisations and reconciliation in Australia, Sport in Society.



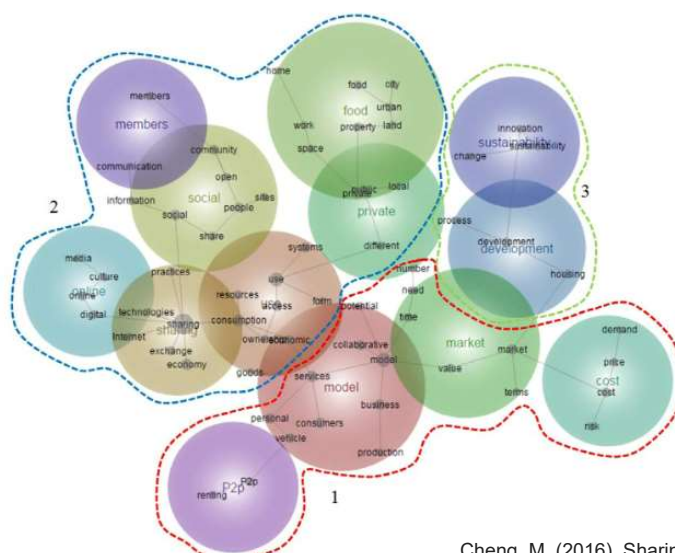
Quadrant Report (from Insight Dashboard)



Ong, F., Vorobjovas-Pinta, O., & Lewis, C. (2020). LGBTQ+ identities in tourism and leisure research: a systematic qualitative literature review. *Journal of Sustainable Tourism*, 1-24.

- Legend**
- FOLDER1_1995-1999
 - FOLDER1_2000-2004
 - FOLDER1_2005-2009
 - FOLDER1_2010-2014
 - FOLDER1_2015-2019
- Relative Frequency**
1. Occur seldom and not unique to category
 2. Occur often and not unique to category
 3. Occur seldom and unique to category
 4. Occur often and unique to category

Groupings of themes within Concept Map

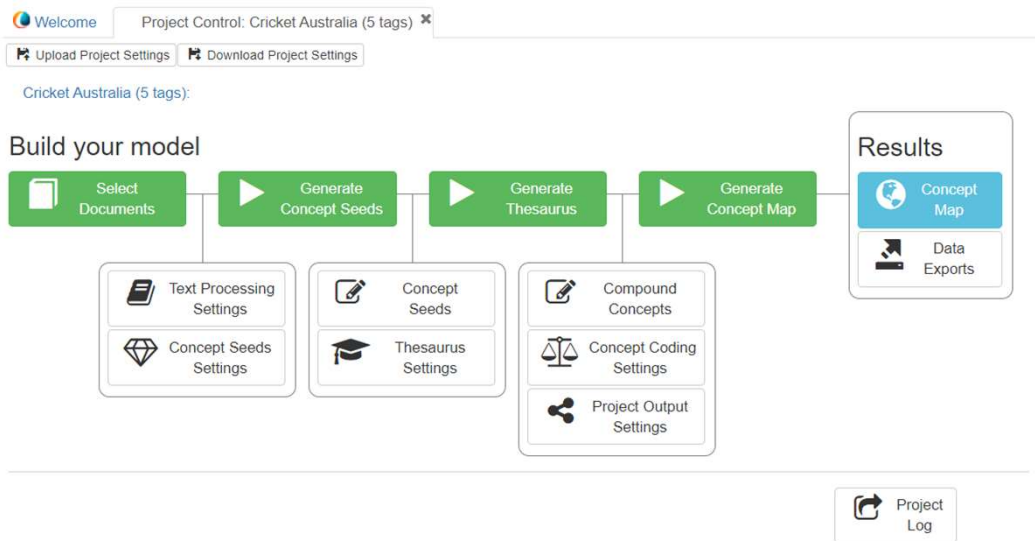


- 1: SE's business models and its impacts
- 2: Nature of SE
- 3: SE's sustainability development

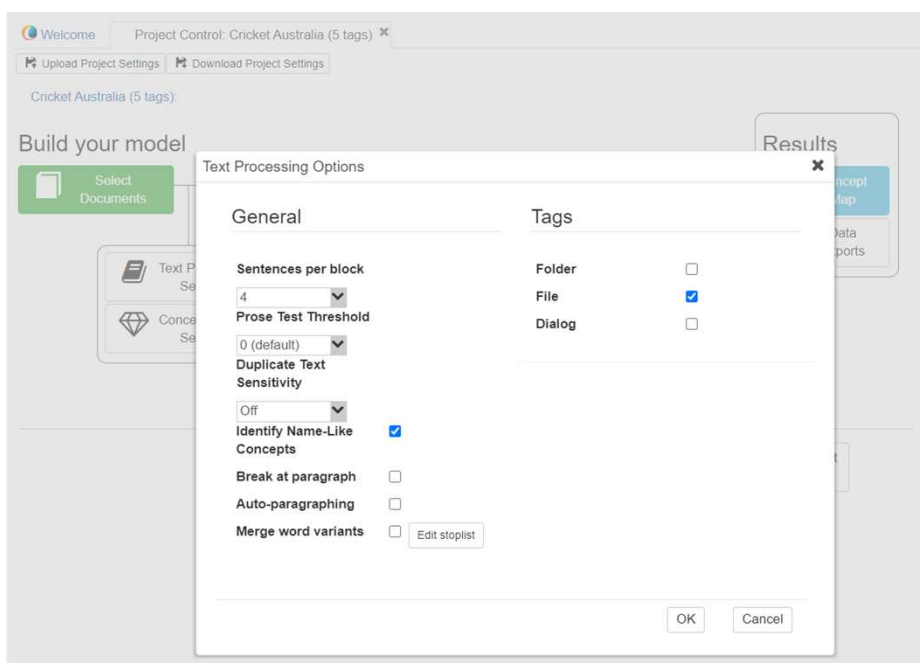
Cheng, M. (2016). Sharing economy: A review and agenda for future research. *International Journal of Hospitality Management*, 57, 60-70.

Fig. 3. Conceptual map of SE literature in general.

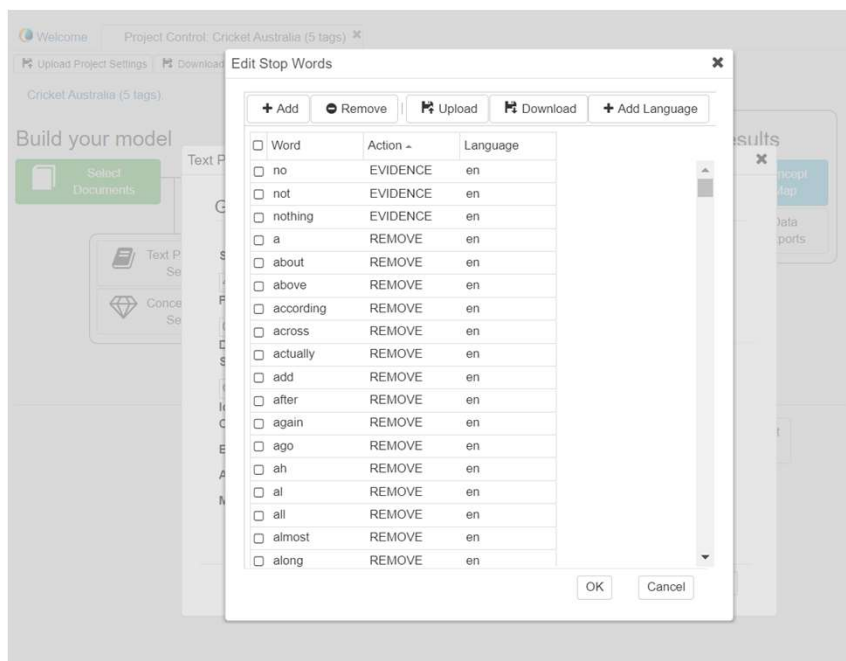
Tips for using Leximancer



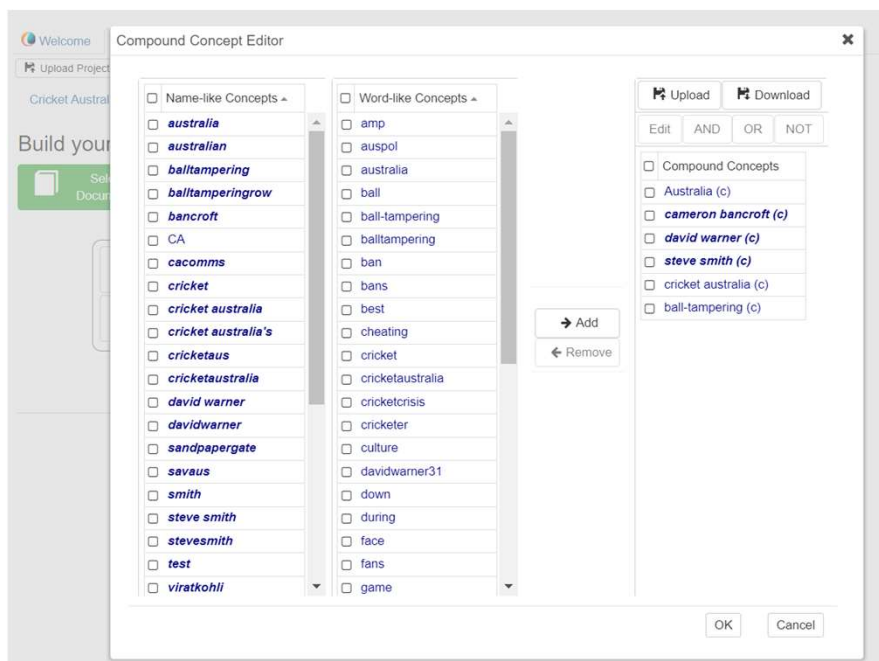
Text Processing Options



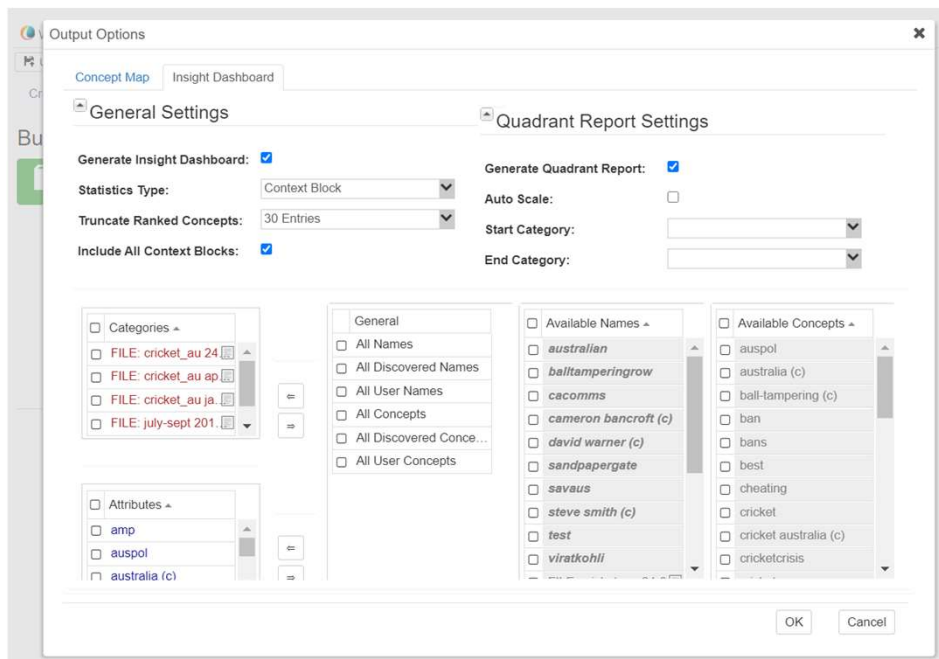
Edit Stop Words



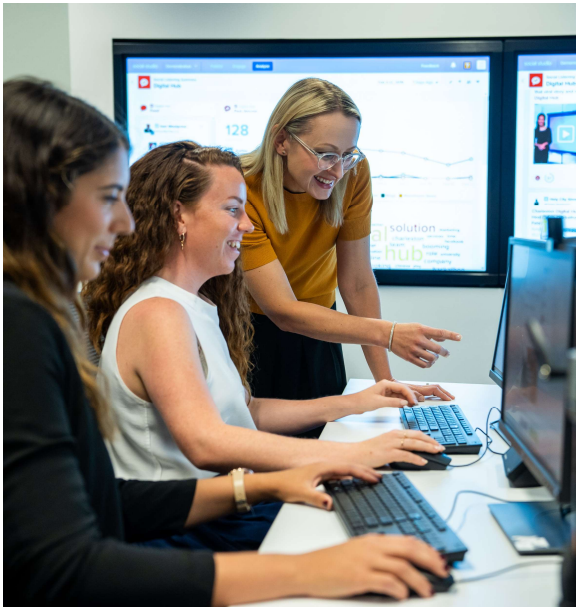
Compound Concept Editor



Output Options



Questions & Thank You



Dr Violetta Wilk
Lecturer & Researcher in Digital Marketing
v.wilk@ecu.edu.au
+61 8 6304 2663

Please connect with me:

 drviolettawilk

 @wilkvioletta