



From Social Media to Expert Reports: The Impact of Source Selection on Automatically Validating Complex Conceptual Models of Obesity

Mannila Sandhu¹, Philippe J. Giabbanelli²(✉), and Vijay K. Mago¹

¹ Department of Computer Science, Lakehead University, Thunder Bay, ON, Canada
{msandhu3,vmago}@lakeheadu.ca

² Computer Science Department, Furman University, Greenville, SC, USA
giabbanelli@gmail.com

Abstract. Models are predominantly developed using either quantitative data (e.g., for structured equation models) or qualitative data obtained through questionnaires designed by researchers (e.g., for fuzzy cognitive maps). The wide availability of social media data and advances in natural language processing raise the possibility of developing models from qualitative data naturally produced by users. This is of particular interest for public health surveillance and policymaking, as social media provide the opinions of constituents. In this paper, we contrast a model produced by social media with one produced via expert reports. We use the same process to derive a model in each case, thus focusing our analysis on the impact of source selection. We found that three expert reports were sufficient to touch on more aspects of a complex problem (measured by the number of relationships) than several million tweets. Consequently, developing a model exclusively from social media may lead to oversimplifying a problem. This may be avoided by complementing social media with expert reports. Alternatively, future research should explore whether a much larger volume of tweets would be needed, which also calls for improvements in scalable methods to transform qualitative data into models.

Keywords: Conceptual modeling · Network analysis · Social web mining · Theme mining · Twitter mining

1 Introduction

Overweight and obesity is now a global phenomenon, found in economically developed or developing countries (e.g., United States [1], European countries [2], South Africa [3], China [4]) as well as in regions that experience a double burden with the concomitant problem of malnutrition [5]. While there are ongoing

Research funded by MITACS Globalink Research Award, Canada.

debates on a possible plateau or even decrease of overweight and obesity in the next generation, updated prevalence data for children suggests that severe obesity is on the rise [6]. There is a plethora of interventions to prevent overweight and obesity in both children [7] and adults [8], and an equally impressive number of interventions for treatment [9, 10]. Yet, individual struggles to achieve a health weight over a sustained period of time. For example, a review of weight management interventions found a weight loss over two years of 1.54 kg [11], which is far from the 5% weight loss recommended to produce health benefits [12]. These challenges have led to the realization that a simple solution would not suffice [13]: the health system needs to cope with the *complexity* of obesity [14–16].

The notion of complexity covers multiple characteristics, such as the vast individual differences (or *heterogeneity*) between weight-related factors [17, 18], or the nonlinear ways in which factors interact to form a system. The obesity system has been the subject of numerous studies [19–22]. This system involves factors from a broad array of sectors (e.g., built environment, eating disorders, weight stigma [23, 24]), with interactions within as well as across sectors. Accurately modeling this system facilitates the development of integrated policies building on cross-sectoral efforts [25, 26]. If policies are developed separately along traditional themes (e.g., public planning works on the environment, doctors work on diseases and physiology, mental health experts work on psychology), then we have a heavily fragmented approach to obesity (Fig. 1a). Efforts such as the Foresight Obesity Map [20, 27], or the Public Health Services Authority’s series of maps [24, 28, 29] thus support the development of synergistic policies working on integrated thematic clusters (Fig. 1b).

Given the importance of developing accurate models of the obesity system, the modeling process often seeks to be comprehensive by including experts and community members [19, 24, 30–34]. While many qualitative modeling processes can produce models in the form of maps [35] (e.g., cognitive/concept mapping, causal loop diagrams), they are generally conducted with a facilitator. Some of the limitations (e.g., costs, trained facilitator) may be addressed through emerging technologies [36]. However, one limitation remains: participants may not openly express their beliefs (e.g., weight discrimination) when perceiving that they may not be well received by a facilitator or the research team. In contrast, the naturally occurring exchange of perspectives in social media provides an unobtrusive approach to collecting beliefs on causes and consequences of obesity. Mining social media may thus provide the views of community members [37–40].

While obtaining a model via social media can inform policymakers about popular support for possible policies [41], the model may stand in stark contrast with an expert-based model [34]. Identifying and reconciling these differences is an important step to integrate social computing (and specifically social web mining) with policy making. In this paper, we contrast how mining social media instead of expert reports affects the validation of a large conceptual model of obesity. This overarching goal is achieved through three consecutive steps. First, we assemble a social media dataset (consisting of several million tweets) and several expert reports (totaling hundred of pages). Second, we employ an innovative

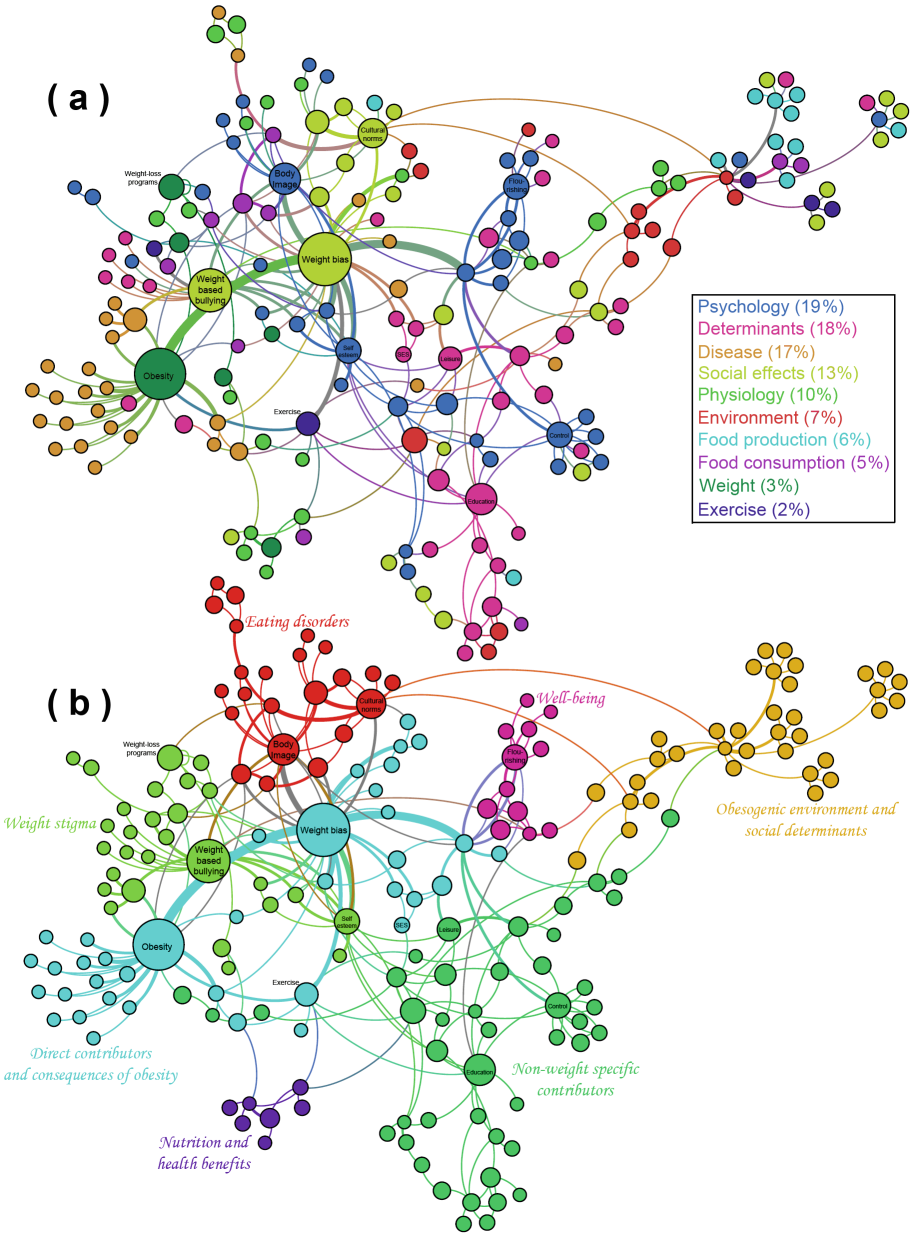


Fig. 1. The Public Health Services Authority’s series of maps [24, 28, 29] suggests that typical categories lead to fragmented approaches (a) whereas themes specific to overweight and obesity can support more integrated options (b). These maps are *conceptual maps* as they articulate how concepts (labeled circles) are related (curves).

multi-step process to examine a conceptual model using both the social media dataset and the expert reports. Finally, we contrast the structure of these models using network methods.

The remainder of this paper is organized as follows. In Sect. 2, we provide background information on the application of social web mining to health, and on the use of conceptual models in obesity research. In Sect. 3, we briefly explain our approach to validate a conceptual model from text. In Sect. 4, we perform this inference on both expert reports and tweets, and we examine how the conceptual models differ. Finally, these differences are discussed and contextualized in Sect. 5.

2 Background

2.1 Social Web Mining for Health

The social media of interest in this paper is Twitter, in which users post and interact through short messages known as ‘tweets’. Twitter has been used for many studies on obesity and weight-related behaviors. For instance, Harris and colleagues collected 1,110 tweets and read them to understand how childhood obesity was discussed [42], while Lydecker *et al.* read 529 tweets to identify the main themes related to fatness [43]. Similarly, So and colleagues analyzed the common features of 120 tweets that were most frequently shared (i.e., retweets) to understand what information individuals preferred to relay when it came to obesity [37]. Reading the tweets to identify themes (i.e., content analysis) is a typical task to understand the arguments that a specific population uses on a subject of interest. Broader examples in health include the content analysis of 700 tweets [44] and 625 tweets [45] to examine the type of claims that health professionals make online, or an examination of 8,934 tweets documenting cyberincivility among nurses and nursing students [46]. While such content analyses make a valuable contribution to the body of knowledge on arguments in public health¹, they do not employ computational methods to automate (parts of) the analysis and thus scale it to a larger dataset. Automation can be as simple as counting how many times keywords of interest appear across tweets. Turner-McGrievy and Beets used Hashtagify.me to automatically count keywords in tens of thousands of tweets on weight loss, health, diet, and fitness. By dividing the analysis across time periods, they were able to examine if there are times of the year when individuals would be likely to consider weight loss, thus contributing to the timing of interventions [48]. Similarly, Sui *et al.* used the intensity of topics on Twitter as part of an effort to identify the public interest in intensive obesity treatment [49]. Such studies illustrate the important shift from having

¹ While our focus is on analyzing the *text* provided by tweets, studies on Twitter that are primarily human- rather than computer-based are not exclusively content analyses. In the study of May and colleagues, the researchers created twitter accounts for fictional obese and non-obese characters. They evaluated whether the weight status mediated how other users would interact with them [47].

humans read and code all tweets to relying on a machine to handle most of a (much larger) dataset. The latter is the focus of data mining applied to the ‘social web’ (i.e. social web mining) which includes social networking sites such as Twitter but also encompasses blogs and micro-blogging. As Twitter has been the social platform of interest for many studies, the term of ‘Twitter mining’ has also emerged to refer specifically to the application of social web mining to Twitter [50].

Social web mining started to garner attention in the late 2000’s to early 2010’s. The application of social web mining to health was discussed in 2010 by Boulos *et al.* [51] and in 2011 by Paul and Dredze [52], showing how a broad range of public health applications could benefit from mining Twitter. Studies have been able to mine a staggering volume of data, going well over what a team of humans could handle. For example, Eichstaedt *et al.* mapped 148 million tweets to counties in an effort to relate language patterns to county-level heart disease mortality [53]. At an even larger scale, Ediger and colleagues used a Cray computer to approximate centrality within two hours on a dataset of interactions between Twitter users comprising 1.47 billion edges [54]. While these cases are noteworthy by their volume of data, studies employing social web mining for obesity research typically involve millions of tweets². Using 2.2 million tweets, Chou and colleagues found that tweets (as well as Facebook posts) often stigmatized individuals living with overweight and obesity [38]. In two studies on obesity and weight-related factors, Karami analyzed 6 million [39] and 4.5 million tweets [40]. In a study of health-related statistics, Culotta mined 4.3 million tweets and found that the data was correlated with obesity [56]. Given that obesity is driven by many factors (e.g., eating behaviors, physical activity behaviors), there is also a wealth of large-scale studies on such factors, such as the work of Abbar *et al.* on 503 million tweets regarding food [57]. Finally, the value proposition of several new platforms is not the analysis of one particular dataset, but rather the ongoing ability to monitor diet or physical activity. This is particularly the case for the Lexicocalorimeter, which measures calories in each US state via Twitter [58], and to a lesser extent for the National Neighborhood Dataset of Zhang *et al.* which tracks diet and physical activity through Twitter [59].

Several commentaries [60] and reviews [61–63] have explored whether this abundance of studies has contributed to public health. Findings depend on what specific aspect of health is concerned. Social media has yet to impact practices in public health surveillance [62], but a review centered on chronic disease found a benefit on clinical outcomes in almost half of the studies [61], and a review specific to obesity highlighted a modest impact on weight [63].

² There are several exceptions of studies employing smaller dataset. However, their objectives may not be to identify themes (which necessitates a large volume of tweets), thus they can accomplish their goals with a smaller dataset. A case in point is the work of Tiggemann and colleagues, who used 3,289 tweets to examine interactions between Twitter communities that promoted either a ‘thin ideal’ or health and fitness [55].

2.2 Conceptual Models in Obesity Research

Although our work will involve the identification of themes, we have a very different endeavor from studies reviewed in the previous section, which focused on identifying themes and their variations across time, places, or communities of users. Our objective is to contrast *conceptual models* that have been automatically extracted from tweets and expert reports. As evoked in the introduction, models of complex systems such as obesity support several important policy-making and analytical tasks. In this section, we briefly review the features that models often seek to capture when it comes to complex health systems, and how models are used in obesity research specifically. Penn detailed key characteristics of complex health systems that justify the development of models (emphases added):

“Many problems that society wishes to address in population health are clearly problems of *managing complex adaptive systems*. They involve making interventions in systems with *multiple interacting causal connections*, which *span domains* from physiological to economic. Additionally, of course, the individuals whose health we ultimately wish to improve adapt and *change their behavior* in response to medical or policy interventions.” [64]

Several of these points were echoed by Silverman in justifying the use of systems-based simulation for population health research [65]. Modeling changes in the heterogeneous health behaviors of individuals often uses the simulation technique of Agent-Based Modeling, and has been done in obesity research on multiple occasions [66–70]. Such models can be very detailed and use widely different architectures to capture the cognitive processes of the agents. Validating them using text is thus an arduous task. Modeling interacting causes across domains has been achieved in obesity research through a variety of techniques. System Dynamics (SD) allows to represent nonlinear interactions between weight-related factors over different time scales and at different strengths [71, 72]. However, much like agent-based modeling, the great level of details supported by SD makes it difficult to derive or validate such models from text. Fuzzy Cognitive Maps (FCM) are a simpler alternative that eliminates the notion of time to focus on the different strengths of causal relations [34, 73–75]. Such models can be compared [34], but validating them from text still requires a trained analyst [76]. An even greater simplification is to use conceptual rather than simulation models. Conceptual models cannot run scenarios or what-if questions, and cannot ‘generate’ numbers. Instead, their focus is to capture relevant factors and whether they are connected [77]. Conceptual models can be compared [78] and validated using text as shown in our previous work [77].

There are several types of conceptual models [35]. We recently detailed the differences between causal maps, mind maps, and concept maps [36]. In short, this paper focuses on concept maps (Fig. 1), which are undirected networks representing concepts as nodes and relationships as edges. Similarly to the other forms of conceptual models aforementioned, a concept map supports policy-oriented

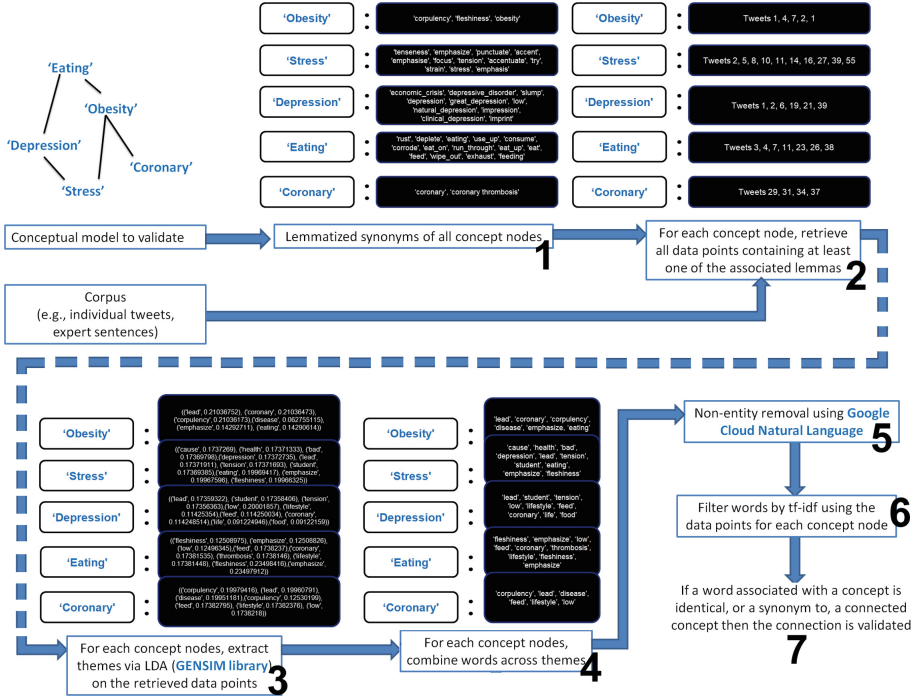


Fig. 2. Our process in seven steps to validate a conceptual model using textual data. The high-definition figure can be zoomed in for details.

tasks such as identifying clusters [27] (e.g., to coordinate actors across domains on one problem such as food) or finding feedback loops [24, 28, 29] (e.g., to use as leverage points in an intervention).

3 Validating a Conceptual Model from Text

The process starts with a conceptual model that we seek to validate, and the text corpus is used to validate. Intuitively, our process uses the concepts’ names to find relevant parts of the corpus and find which concepts tend to co-occur. Technical aspects include handling variations in language (as we cannot rigidly assume that a concept’s name will appear as such), identifying themes, and mapping themes from the corpus back to concepts in the conceptual model. Our process uses seven steps, illustrated on a theoretical example in Fig. 2. The first two steps are performed for each concept node:

- (1.a) We replace all concepts’ names and words from the corpus with their base form (i.e., lemma). This is accomplished through *lemmatization*, which uses a morphological analysis to remove inflectional endings. This step ensures that minor variations of a term are all mapped to the same one (e.g., ‘flooding’ and ‘floods’ are all mapped to ‘flood’).

- (1.b) Each lemmatized concept names is expanded with derivationally related forms. For instance, instead of only searching for ‘flood’ in the corpus, we will also accept words such as ‘deluge’.
- (2) For each concept (i.e., the expanded lemma), we retrieve all parts of the corpus that contain it. For instance, the concept ‘flooding’ will lead to retrieving all tweets include the lemmas ‘flood’ or ‘deluge’.

Upon completion of step 2, we have related a portion of the corpus to each concept node. We then find the themes in each portion of the corpus using three parameters:

- (3) We apply the Latent Dirichlet Accuracy (LDA) model to find prevalent themes. The two parameters for this step are the number of themes and number of words per theme.
- (4) We gather words across themes into a single set of words. This set is cleaned by removing words that are already present in the set of derivationally related form of the node. In other words, we only look for concepts that the node could be *associated* with but not *equivalent* to.
- (5) Since concepts’ names are entities, a concept can only be associated with an entity. Consequently, we remove all non-entities from the words.
- (6) At this step, we have a set of entities that a concept node could be associated with. However, some of the entities may be noise rather than meaningful associations. We thus sort the entities by tf-idf (term-frequency inverse-document-frequency) computed over the set of tweets in which each word appears. We use a threshold parameter to identify which entities have a sufficient tf-idf to be selected.

Upon completion of step 2, we found entities that a concept node could be associated with. The final step goes back to the conceptual model to see if the association exists:

- (7) For each node, we compare its associated entities with its connected nodes and derivationally related forms. If there is a match, then the text corpus has confirmed an association between the two concepts. If no match is found, the association is not confirmed. Note that associated entities that do not match any connected nodes suggest additional connections, which is a different from validation as we seek to confirm existing connections.

This process is also depicted in Fig. 3, listing the libraries that can be used for each step. The specific versions of the libraries used in our experiments are included in Sect. 4.

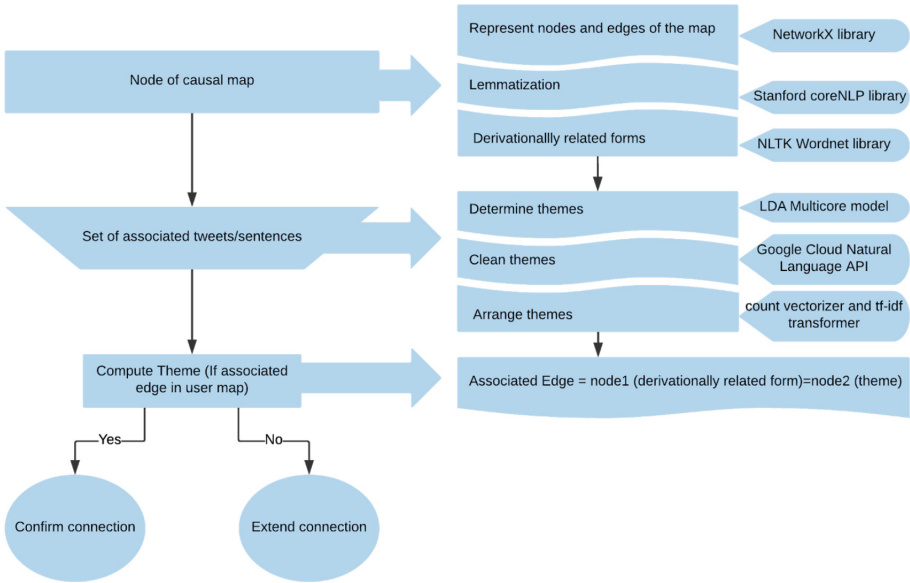


Fig. 3. Alternative view of our process, including libraries and APIs.

4 Comparing Conceptual Models from Twitter and Expert Reports

4.1 Datasets and Pre-processing

The conceptual model that we seek to validate was developed with the Provincial Health Services Authority (PHSA) of British Columbia to explore the interrelationships involved in obesity and well-being. The model was presented in 2015 at the Canadian Obesity Summit [24] and tested with policy makers in 2016 [29]. The model is now part of the **ActionableSystems** tool [28] can be downloaded at <https://osf.io/7ztwu/> within ‘Sample maps’ (file *Drasic et al (edges).csv*). The model consists of 98 nodes and 177 edges. From here on, we will refer to it as ‘the PHSA map’.

To validate the PHSA map, we used two datasets. Our first dataset (‘the twitter dataset’) consists of 6,633,625 tweets in the English language on obesity collected from Oct. 2, 2018 to Oct. 4, 2018. The number of tweets was chosen to be in line with comparable studies at the interface of natural language processing and obesity research [38–40]. The keywords to collect the tweets included each of the 98 concept names in the PHSA map as well as their synonyms automatically retrieved through WordNet. For instance, we used not only ‘obesity’ but also words such as ‘fatness’, ‘corpulent’, ‘embonpoint’ and ‘fleshiness’. Similarly, physical activity was expanded to include many forms such as calisthenics, isometrics, jogging, jump rope, and so on. The rationale is that the map contains abstract concepts, but individuals may speak of specific instances or use a vari-

ety of words to describe the same abstraction. After collecting a large number of tweets, natural language applications require extensive pre-processing. The impact of each options (and their interactions) on results obtained from Twitter has been extensively described when performing sentiment analysis [79–81] and in more generic tasks such as classification [82]. Some of these options are summarized in Fig. 4 and include the removal of parts deemed unnecessary for analysis (e.g., hashtags, URLs, numbers, non English words) or the mapping of data into forms that can be more conveniently processed (e.g., expanding acronyms and abbreviations, replacing emojis, spell checking). The pre-processing options used for our dataset are depicted in Fig. 5. These options are chosen specifically for our research question: for instance, we remove stop words because they cannot be meaningful concept names in a model, but other analyses (e.g., attributing tweets to specific writers) may have kept such words. The order of the steps also matters: for instance, we cannot perform part-of-speech tagging and lemmatization (step 5) before ensuring that all the words have been corrected (step 3). After pre-processing, our dataset included 1,791,333 tweets.

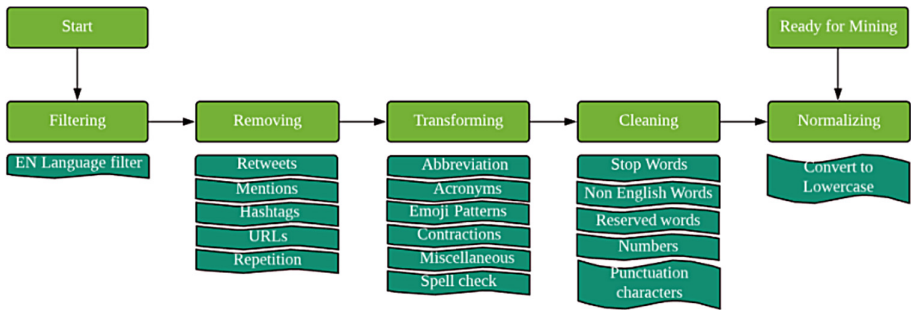


Fig. 4. Typical pre-processing techniques applied to tweets.

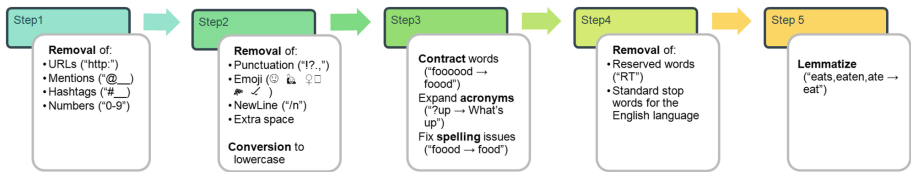


Fig. 5. Pre-processing techniques applied to our tweeter dataset in a specific order. We used a Spell Checker library in step 3, the Natural Language Toolkit (NLTK) for steps 1–4, and the Stanford coreNLP library for step 5.

The second dataset is formed of three reports on obesity: the 2010 report from the white house task force on childhood obesity [83], the 2013 report to the Provincial Health Services Authority [84] and its 2015 update (whose findings

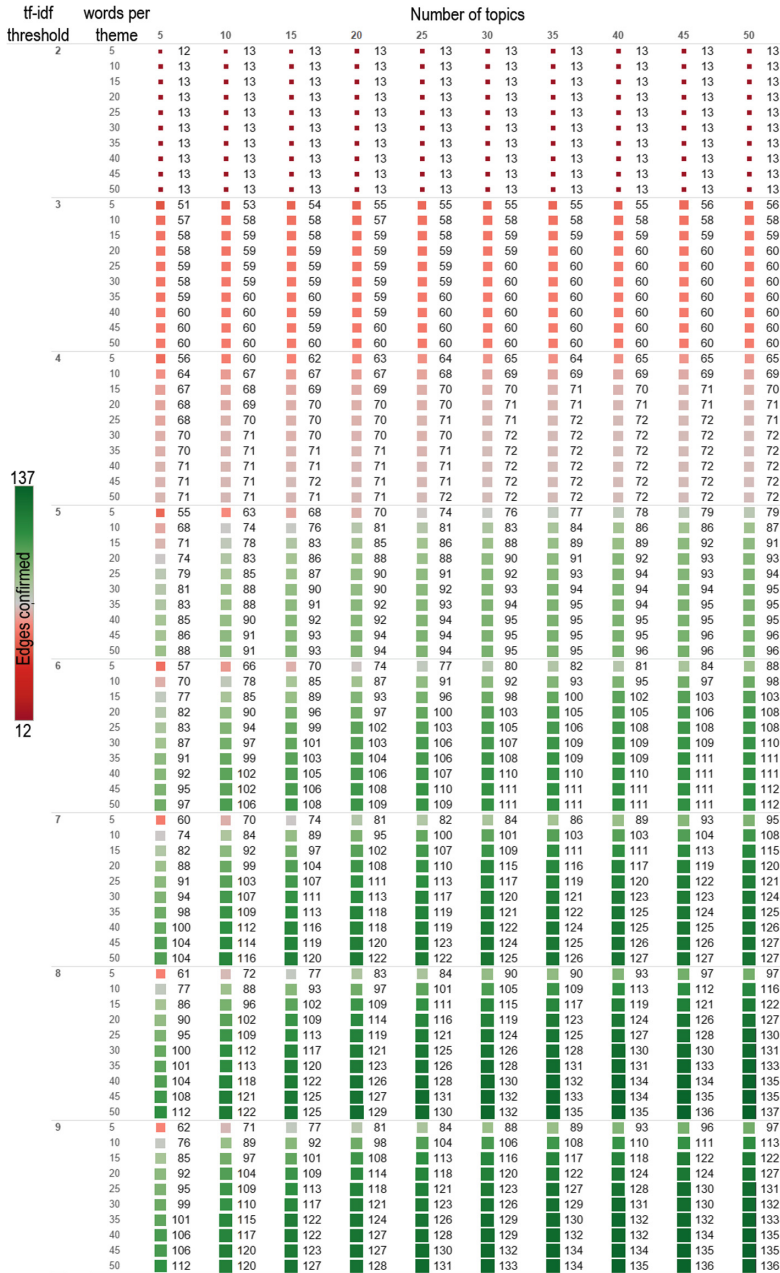


Fig. 6. Average number of edges confirmed (out of 177 in the PHSA map) for each combination of parameter values over ten experiments.

are published in [24]). We combined the three reports with the PyPDF2 library, leading to 310 pages, and we kept 247 pages after removing those that were either blank or only contained images. Pages were then transformed into raw text using the pdftotext library and divided into 4,302 sentences using the full point ('.'). Pre-processing was finally applied, using the same script as for tweets while noting that several options such as removing emojis would not be triggered. The resulting dataset had 3447 sentences.

4.2 Validating the Model for Each Dataset

The methods introduced in Sect. 3 are implemented in Python, relying on libraries as listed in Table 1. While our implementation was able to cope with millions of tweets, we note that a larger volume of data may also require a distributed database architecture and an efficient search engine such as Elastic-search [85].

Table 1. Libraries used in each step (Sect. 3) of our experiments.

Step	Library	Used for
	NetworkX	Conceptual model (accessing node labels and edges)
1a	Stanford coreNLP	Lemmatization
1b	WordNet	Derivationally related forms
3	[86]	Parallel, multi-core Latent Dirichlet Allocation (LDA) model for big data
5	Google Cloud Natural language API	Entity identification
6	scikit-learn (CountVectorizer, TfidfTransformer)	Sorting words by tf-idf

Our approach has three parameters: number of themes, number of words per theme, and tf-idf threshold to eliminate noise. Hyperparameter optimization was thus necessary to use each dataset most efficiently, and fairly compare their potential in validating a model. To optimize performances with expert reports, we performed a grid search by varying the number of topics and words per topic from 5 to 50 in increments of 5, and we varied the tf-idf from 2 to 9 by increments of 1. This resulted in 800 combinations of parameter values. As there is randomness in the LDA model, we performed ten experiments per combination of parameter values, leading to a total of 8,000 experiments. At most, our process validated an average of 136.5 edges (77.11% of the map) using 50 topics, 50 words per topic, and a td-idf threshold of 8 (Fig. 6).

A grid search was also performed on the Twitter dataset. However, our current implementation takes approximately five days to compute the results for

one combination of parameter values (single experiment), using a server-grade workstation (Dual Xeon Gold 6140). Given this limitation, we used single experiments and a coarser grid. At most, our process validated 101 edges (57.06%) using 50 topics, 50 words per topic, and a tf-idf threshold of 9.

5 Discussion

A focus group with a few participants may only discuss some of the interrelationships at work in overweight and obesity, and may avoid sharing opinions that are potentially disapproved by others. In contrast, social media such as Twitter provide access to a massive number of participants who can use conditions of anonymity to share opinions more freely. Social web mining applied to Twitter thus comes with the potential to explore many interrelationships in an unobtrusive fashion. In particular, crowdsourcing over Twitter holds the promise of easily building large conceptual models, under the assumption that at least some groups of users will touch on each part of the model. Our study questions this potential and promises by analyzing whether millions of tweets are more useful to develop a conceptual model of obesity than a handful of reports.

Although conceptual models can be automatically compared [78], developing a model from each dataset (tweets vs. reports) and comparing them would not be able to tell us which one is ‘better’. Our study question thus requires a referential. We use a previously developed conceptual model of obesity and well-being to serve as referential, and we establish how much of this model would have been obtained if we used either tweets or reports. In other words, we measured the percentage of the model’s structure that is confirmed with each dataset.

While both datasets were able to cover over half of the model, we note that it only took three expert reports compared to using millions of tweets. In addition, despite the abundance of tweets, the three expert reports touched on more relationships. Within our application context, these results suggest that an exclusive reliance on social media may result in oversimplifying a complex system, thus limiting the potential to automatically develop models using such a source. We note that a comprehensive analysis across subjects and using a variety of maps would be needed to assess whether our results produced on one model (the Provincial Health Services Authority map) and one application subject (obesity) can be generalized to other models and subjects.

There are several limitations to this study, which we intend to address in our future research. First, one of the premises of big data research is that a large volume may compensate for many imperfections in the individual data points. Although we used a similar number of tweets to other studies at the interface of natural language processing and obesity research [38–40], it is possible that some of the interrelationships of the model we seek to validate are rare and thus only detectable in even larger datasets. Repeating this study with significantly larger datasets could elucidate this question. However, we then run into the second issue: our process to validate a causal map against textual data is very computational intensive. The search space to optimize the result is defined

by three parameters which involve randomness, thus requiring several experiments for each combination of parameter values. On a server-grade workstation, a single combination with a CPU-based implementation requires in the order of days. Optimizing results and using larger datasets will thus require implementations that scale, with a particularly promising option consisting of a GPU-based implementation. Alternatively, we may reduce the search space if we can better characterize the impact that parameters generally have on the results and then devise more computationally efficient processes. For instance, the tf-idf threshold plays an essential role in driving performances (Fig. 6) but may be replaced by additional pre-processing steps preventing the inclusion of noise, such as classifiers removing unwanted documents [87].

6 Conclusion

Both social media data and expert reports may be used to take into account popular perspectives and expert opinions when creating large conceptual models. In the case of obesity, we found that three expert reports discussed 77% of all possibilities while millions of tweets on obesity and its cognates covered fewer interrelationships. Creating models using social media only may thus result in an oversimplification of complex problems.

Acknowledgments. The authors are indebted to Mitacs Canada for providing the financial support which allowed MS to perform this research at Furman University, while mentored by PJG (local advisor) and VKM (home advisor). Publication costs are supported by an NSERC Discovery Grant for VKM. We thank Chetan Harichandra Mendhe for gathering the tweets under supervision of VKM.

Contributions. MS wrote the scripts to generate the results and analyzed them. PJG wrote the manuscript and designed the methods. MS was advised by PJG and VKM, who jointly initiated the study. All authors read and approved of this manuscript.

References

1. Centers for Disease Control and Prevention (CDC): Selected health conditions and risk factors, by age: United states, selected years 1988–1994 through 2015–2016
2. Peralta, M., et al.: Prevalence and trends of overweight and obesity in older adults from 10 European Countries from 2005 to 2013. *Scand. J. Public Health* **46**, 522–529 (2018). <https://doi.org/10.1177/1403494818764810>
3. Lubbe, J.: Obesity and metabolic surgery in South Africa. *S. Afr. Gastroenterology Rev.* **16**(1), 23–28 (2018)
4. Wang, Y., Wang, L., Qu, W.: New national data show alarming increase in obesity and noncommunicable chronic diseases in China. *Eur. J. Clin. Nutr.* **71**(1), 149 (2017)
5. Ng, M., et al.: Global, regional, and national prevalence of overweight and obesity in children and adults during 1980–2013: a systematic analysis for the global burden of disease study 2013. *Lancet* **384**(9945), 766–781 (2014)

6. Skinner, A.C., Perrin, E.M., Skelton, J.A.: Prevalence of obesity and severe obesity in US children, 1999-2014. *Obesity* **24**(5), 1116–1123 (2016)
7. Bleich, S.N., et al.: Interventions to prevent global childhood overweight and obesity: a systematic review. *Lancet Diabetes Endocrinol.* **6**(4), 332–346 (2018)
8. Hutchesson, M., et al.: eH health interventions for the prevention and treatment of overweight and obesity in adults: a systematic review with meta-analysis. *Obes. Rev.* **16**(5), 376–392 (2015)
9. Rajjo, T., et al.: Treatment of pediatric obesity: an umbrella systematic review. *J. Clin. Endocrinol. Metab.* **102**(3), 763–775 (2017)
10. Teixeira, P.J., et al.: Successful behavior change in obesity interventions in adults: a systematic review of self-regulation mediators. *BMC Med.* **13**(1), 84 (2015)
11. National Institute for Health and Care Excellence: Managing overweight and obesity in adults-lifestyle weight management services. NICE Public Health Guideline, **53** (2014)
12. Blackburn, G.: Effect of degree of weight loss on health benefits. *Obes. Res.* **3**(S2), 211s–216s (1995)
13. Fink, D.S., Keyes, K.M.: Wrong answers: when simple interpretations create complex problems. In: *Systems Science and Population Health*, pp. 25–36 (2017)
14. Froot, S., et al.: Obesity, complexity, and the role of the health system. *Curr. Obes. Rep.* **2**(4), 320–326 (2013)
15. Finegood, D.T.: The complex systems science of obesity. In: *The Oxford Handbook of the Social Science of Obesity* (2011)
16. Rutter, H., et al.: The need for a complex systems model of evidence for public health. *Lancet* **390**(10112), 2602–2604 (2017)
17. Giabbanelli, P.J.: Analyzing the complexity of behavioural factors influencing weight in adults. In: Giabbanelli, P.J., Mago, V.K., Papageorgiou, E.I. (eds.) *Advanced Data Analytics in Health*. SIST, vol. 93, pp. 163–181. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-77911-9_10
18. Deck, P., Giabbanelli, P., Finegood, D.T.: Exploring the heterogeneity of factors associated with weight management in young adults. *Can. J. Diabetes* **37**, S269–S270 (2013)
19. Giabbanelli, P.J., Torsney-Weir, T., Mago, V.K.: A fuzzy cognitive map of the psychosocial determinants of obesity. *Appl. Soft Comput.* **12**(12), 3711–3724 (2012)
20. Jebb, S., Kopelman, P., Butland, B.: Executive summary: foresight ‘tackling obesity: future choices’ project. *Obes. Rev.* **8**, vi–ix (2007)
21. Xue, H., et al.: Applications of systems modelling in obesity research. *Obes. Rev.* **19**(9), 1293–1308 (2018)
22. Frerichs, L., et al.: Mind maps and network analysis to evaluate conceptualization of complex issues: a case example evaluating systems science workshops for childhood obesity prevention. *Eval. Program Plan.* **68**, 135–147 (2018)
23. Johnston, L.M., Matteson, C.L., Finegood, D.T.: Systems science and obesity policy: a novel framework for analyzing and rethinking population-level planning. *Am. J. Public Health* **104**(7), 1270–1278 (2014)
24. Drasic, L., Giabbanelli, P.J.: Exploring the interactions between physical well-being, and obesity. *Can. J. Diabetes* **39**, S12–S13 (2015)
25. Dubé, L., Du, P., McRae, C., Sharma, N., Jayaraman, S., Nie, J.-Y.: Convergent innovation in food through big data and artificial intelligence for societal-scale inclusive growth. *Technol. Innov. Manag. Rev.* **8**, 49–65 (2018)
26. Jha, S.K., Gold, R., Dube, L.: Convergent innovation platform to address complex social problems: a tiered governance model. In: *Academy of Management Proceedings*, Volume 2016, Academy of Management Briarcliff Manor, NY 10510 (2016)

27. Finegood, D.T., Merth, T.D., Rutter, H.: Implications of the foresight obesity system map for solutions to childhood obesity. *Obesity* **18**(S1), S13–S16 (2010)
28. Giabbanelli, P.J., Baniukiewicz, M.: Navigating complex systems for policymaking using simple software tools. In: Giabbanelli, P.J., Mago, V.K., Papageorgiou, E.I. (eds.) *Advanced Data Analytics in Health*. SIST, vol. 93, pp. 21–40. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-77911-9_2
29. Giabbanelli, P., et al.: developing technology to support policymakers in taking a systems science approach to obesity and well-being. *Obes. Rev.* **17**, 194–195 (2016)
30. Owen, B., et al.: Understanding a successful obesity prevention initiative in children under 5 from a systems perspective. *PloS one* **13**(3), e0195141 (2018)
31. McGlashan, J., et al.: Quantifying a systems map: network analysis of a childhood obesity causal loop diagram. *PloS one* **11**(10), e0165459 (2016)
32. McGlashan, J., et al.: Comparing complex perspectives on obesity drivers: action-driven communities and evidence-oriented experts. *Obes. Sci. Pract.* **4**, 575–581 (2018)
33. Allender, S., et al.: A community based systems diagram of obesity causes. *PLoS One* **10**(7), e0129683 (2015)
34. Giles, B.G., et al.: Integrating conventional science and aboriginal perspectives on diabetes using fuzzy cognitive maps. *Soc. Sci. Med.* **64**(3), 562–576 (2007)
35. Voinov, A., et al.: Tools and methods in participatory modeling: selecting the right tool for the job. *Environ. Model. Softw.* **109**, 232–255 (2018)
36. Reddy, T., Giabbanelli, P.J., Mago, V.K.: The artificial facilitator: guiding participants in developing causal maps using voice-activated technologies. In: *International Conference on Augmented Cognition* (2019)
37. So, J., et al.: What do people like to “share” about obesity? A content analysis of frequent retweets about obesity on twitter. *Health Commun.* **31**(2), 193–206 (2016)
38. Chou, W.Y.S., Prestin, A., Kunath, S.: Obesity in social media: a mixed methods analysis. *Transl. Behav. Med.* **4**(3), 314–323 (2014)
39. Shaw Jr., G., Karami, A.: Computational content analysis of negative tweets for obesity, diet, diabetes, and exercise. *Proc. Assoc. Inf. Sci. Technol.* **54**(1), 357–365 (2017)
40. Karami, A., et al.: Characterizing diabetes, diet, exercise, and obesity comments on twitter. *Int. J. Inf. Manag.* **38**(1), 1–6 (2018)
41. Giabbanelli, P.J., Adams, J., Pillutla, V.S.: Feasibility and framing of interventions based on public support: leveraging text analytics for policymakers. In: Meiselwitz, G. (ed.) *SCSM 2016*. LNCS, vol. 9742, pp. 188–200. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-39910-2_18
42. Harris, J.K., et al.: Communication about childhood obesity on twitter. *Am. J. Public Health* **104**(7), e62–e69 (2014)
43. Lydecker, J.A., et al.: Does this tweet make me look fat? A content analysis of weight stigma on twitter. *Eat. Weight. Disord.-Stud. Anorex. Bulim. Obes.* **21**(2), 229–235 (2016)
44. Lee, J.L., et al.: What are health-related users tweeting? A qualitative content analysis of health-related users and their messages on twitter. *J. Med. Internet Res.* **16**(10), e237 (2014)
45. Alnemer, K.A., et al.: Are health-related tweets evidence based? Review and analysis of health-related tweets on twitter. *J. Med. Internet Res.* **17**(10), e246 (2015)
46. De Gagne, J.C., et al.: Uncovering cyberincivility among nurses and nursing students on twitter: a data mining study. *Int. J. Nurs. Stud.* **89**, 24–31 (2019)

47. May, C.N., et al.: Weight loss support seeking on twitter: the impact of weight on follow back rates and interactions. *Transl. Behav. Med.* **7**(1), 84–91 (2016)
48. Turner-McGrievy, G.M., Beets, M.W.: Tweet for health: using an online social network to examine temporal trends in weight loss-related posts. *Transl. Behav. Med.* **5**(2), 160–166 (2015)
49. Sui, Z., et al.: Recent trends in intensive treatments of obesity: is academic research matching public interest? *Surg. Obes. Relat. Dis.* (2019). <https://www.sciencedirect.com/science/article/pii/S1550728918311948>
50. O’Leary, D.E.: Twitter mining for discovery, prediction and causality: applications and methodologies. *Intell. Syst. Account. Financ. Manag.* **22**(3), 227–247 (2015)
51. Boulos, M.N.K., et al.: Social web mining and exploitation for serious applications: technosocial predictive analytics and related technologies for public health, environmental and national security surveillance. *Comput. Methods Programs Biomed.* **100**(1), 16–23 (2010)
52. Paul, M.J., Dredze, M.: You are what you tweet: analyzing twitter for public health. *Icwsm* **20**, 265–272 (2011)
53. Eichstaedt, J.C., et al.: Psychological language on twitter predicts county-level heart disease mortality. *Psychol. Sci.* **26**(2), 159–169 (2015)
54. Ediger, D., et al.: Massive social network analysis: mining twitter for social good. In: *2010 39th International Conference on Parallel Processing*, pp. 583–593. IEEE (2010)
55. Tiggemann, M., et al.: Tweeting weight loss: a comparison of# thinspiration and# fitspiration communities on twitter. *Body Image* **25**, 133–138 (2018)
56. Culotta, A.: Estimating county health statistics with twitter. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1335–1344. ACM (2014)
57. Abbar, S., Mejova, Y., Weber, I.: You tweet what you eat: studying food consumption through twitter. In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pp. 3197–3206. ACM (2015)
58. Alajajian, S.E., et al.: The lexicocalorimeter: gauging public health through caloric input and output on social media. *PloS One* **12**(2), e0168893 (2017)
59. Nguyen, Q.C., et al.: Building a national neighborhood dataset from geotagged twitter datafor indicators of happiness, diet, and physical activity. *JMIR Public Health Surveill.* **2**(2), e158 (2016)
60. Eke, P.I.: Using social media for research and public health surveillance. *J. Dent. Res.* **90**(9), 1045 (2011)
61. Patel, R., et al.: Social media use in chronic disease: a systematic review and novel taxonomy. *Am. J. Med.* **128**(12), 1335–1350 (2015)
62. Charles-Smith, L.E., et al.: Using social media for actionable disease surveillance and outbreak management: a systematic literature review. *PloS One* **10**(10), e0139701 (2015)
63. Waring, M.E., et al.: Social media and obesity in adults: a review of recent research and future directions. *Curr. Diabetes Rep.* **18**(6), 34 (2018)
64. Penn, A.: Moving from overwhelming to actionable complexity in population health policy: Can alife help? (2018)
65. Silverman, E.: Bringing alife and complex systems science to population health research. *Artif. Life* **24**(3), 220–223 (2018)
66. Giabbanelli, P.J., Crutzen, R.: Using agent-based models to develop public policy about food behaviours: future directions and recommendations. *Comput. Math. Methods Med.* (2017). <https://www.hindawi.com/journals/cmmm/2017/5742629/abs/>

67. Giabbanelli, P., Crutzen, R.: An agent-based social network model of binge drinking among Dutch adults. *J. Artif. Soc. Soc. Simul.* **16**(2), 10 (2013)
68. Khademi, A., Zhang, D., Giabbanelli, P.J., Timmons, S., Luo, C., Shi, L.: An agent-based model of healthy eating with applications to hypertension. In: Giabbanelli, P.J., Mago, V.K., Papageorgiou, E.I. (eds.) *Advanced Data Analytics in Health. SIST*, vol. 93, pp. 43–58. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-77911-9_3
69. Zhang, D., et al.: Impact of different policies on unhealthy dietary behaviors in an urban adult population: an agent-based simulation model. *Am. J. Public Health* **104**(7), 1217–1222 (2014)
70. Giabbanelli, P.J., et al.: Modeling the influence of social networks and environment on energy balance and obesity. *J. Comput. Sci.* **3**(1–2), 17–27 (2012)
71. Verigin, T., Giabbanelli, P.J., Davidsen, P.I.: Supporting a systems approach to healthy weight interventions in British Columbia by modeling weight and well-being. In: *Proceedings of the 49th Annual Simulation Symposium*, Society for Computer Simulation International, p. 9 (2016)
72. Fallah-Fini, S., et al.: Modeling us adult obesity trends: a system dynamics model for estimating energy imbalance gap. *Am. J. Public Health* **104**(7), 1230–1239 (2014)
73. Mago, V.K., et al.: Fuzzy cognitive maps and cellular automata: an evolutionary approach for social systems modelling. *Appl. Soft Comput.* **12**(12), 3771–3784 (2012)
74. Giabbanelli, P.J., Jackson, P.J., Finegood, D.T.: Modelling the joint effect of social determinants and peers on obesity among Canadian adults. In: Dabbaghian, V., Mago, V. (eds.) *Theories and simulations of complex social systems. ISRL*, vol. 52, pp. 145–160. Springer, Heidelberg (2014). https://doi.org/10.1007/978-3-642-39149-1_10
75. Giabbanelli, P.J., Crutzen, R.: Creating groups with similar expected behavioural response in randomized controlled trials: a fuzzy cognitive map approach. *BMC Med. Res. Methodol.* **14**(1), 130 (2014)
76. Pillutla, V.S., Giabbanelli, P.J.: Iterative generation of insight from text collections through mutually reinforcing visualizations and fuzzy cognitive maps. *Appl. Soft Comput.* **76**, 459–472 (2019)
77. Giabbanelli, P.J., Jackson, P.J.: Using visual analytics to support the integration of expert knowledge in the design of medical models and simulations. *Procedia Comput. Sci.* **51**, 755–764 (2015)
78. Giabbanelli, P.J., Tawfik, A.A., Gupta, V.K.: Learning analytics to support teachers' assessment of problem solving: a novel application for machine learning and graph algorithms. In: Ifenthaler, D., Mah, D.-K., Yau, J.Y.-K. (eds.) *Utilizing Learning Analytics to Support Study Success*, pp. 175–199. Springer, Cham (2019). https://doi.org/10.1007/978-3-319-64792-0_11
79. Jianqiang, Z., Xiaolin, G.: Comparison research on text pre-processing methods on twitter sentiment analysis. *IEEE Access* **5**, 2870–2879 (2017)
80. Singh, T., Kumari, M.: Role of text pre-processing in twitter sentiment analysis. *Procedia Comput. Sci.* **89**, 549–554 (2016)
81. Symeonidis, S., Effrosynidis, D., Arampatzis, A.: A comparative evaluation of pre-processing techniques and their interactions for twitter sentiment analysis. *Expert. Syst. Appl.* **110**, 298–310 (2018)

82. Keerthi Kumar, H.M., Harish, B.S.: Classification of short text using various pre-processing techniques: an empirical evaluation. In: Sa, P.K., Bakshi, S., Hatzilygeroudis, I.K., Sahoo, M.N. (eds.) *Recent Findings in Intelligent Computing Techniques*. AISC, vol. 709, pp. 19–30. Springer, Singapore (2018). https://doi.org/10.1007/978-981-10-8633-5_3
83. Barnes, M.: Solving the problem of childhood obesity within a generation. White House Task Force on Childhood Obesity Report to the President, Washington, DC (2010)
84. Daghofer, D.: From weight to well-being: time for shift in paradigms. Technical report, a discussion paper on the inter-relationships among obesity, overweight ... (2013)
85. Shah, N., Willick, D., Mago, V.: A framework for social media data analytics using Elasticsearch and Kibana. *Wirel. Netw.*, 1–9 (2009)
86. Rehurek, R., Sojka, P.: Software framework for topic modelling with large corpora. In: *In Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*. Citeseer (2010)
87. Robinson, K., Mago, V.: Birds of prey: identifying lexical irregularities in spam on twitter. *Wirel. Netw.* 1–8 (2018). <https://doi.org/10.1007/s11276-018-01900-9>