GoVegan: Exploring Motives and Opinions from Tweets

Phoey Lee Teh[1]([✉]) [ID] and Wei Li Yap[2]

[1] Department of Computing and Information Systems, Subang Jaya, Malaysia
phoeyleet@sunway.edu.my
[2] School of Science and Technology, Subang Jaya, Malaysia

Abstract. This report is suggesting the beneficial effect of clustering micro bloggers tweets from 60 hash tags relating to the issue of Veganism. Going Vegan is a wellknown effect on health. We aimed to analyze tweets coming from casual Twitter users and twitter accounts representing the veganism society and industry. We cluster the group of discourse that coming from 60 and more hashtags. These tags include tweets that have tagged with #plantbaseddiet, #vegan food, #vegetarian, etc. We collected n = 50,634 tweets and analyzed n = 25,639 processed tweets. The result shows that sampled tweets, which includes 1) concerns about animal welfare; 2) sustainability (environment) 3) ways to live a healthier lifestyle (Health), and 4) methods and options for Vegan (recipe). Although with 60 + hash tags, this grouping practice allows decision making processes more manageable. This work not only demonstrates the application of a clustering algorithm to collate micro blogs with different hash tags into groups of similar topics but also shown that it is possible to develop a platform for automatically assembling information on the same subject from a range of different micro blogs. The application can significantly assist others, including academic researchers, or businesses, to quickly and effectively find information and knowledge from these sources. This application is possible for society looking for a healthy life.

Keywords: Knowledge management · Clustering · Text analysis · Opinion mining · Tweets

1 Introduction

Vegan, Pronounced as "Veegan", was coined in 1944 by a carpenter Donald Watson from a suggestion by early members Mr George A [1]. Henderson and his wife mark the beginning and end of vegetarianism [2]. Strict vegetarian and vegan households around the world are tabulated to acknowledge the importance of veganism diet for global environmental change [2]. There are more than 2.5% of U.S. adults that have adopted a strict vegan lifestyle in the year 2018. Within six months (January to June) in 2019, a total of 250,000 people in the U.K. signed up the Campaign to go Vegan [3]. With the number of people growing, knowledge-sharing hence become critical. Through social media, society can quickly share strategies and ways to support a vegan lifestyle. For instance, Twitter can be a place for connection and information sharing. It allows

plenty of information to share and connecting subscribers. Taken these advantages, politicians, agencies, and environmental activist has, in this even, also use this channel to disseminate information online. Studies have revealed that twitter opinion leadership makes a significant contribution to individuals involvement, often, contribute to the area of the political process and decision making happen on social media [4]. In many ways, social media has proven useful on various kinds of matters [5–7], in others words, to gather opinions on what public talk about going Vegan, tweets are the media selected for the purpose.

However, Tweets have a limit of 144 characters per post. Multiple hash tags are user generated to indicate the topic or topics of a tweet. For example, about vegan tweets, there is a total of 60 + hash tags has been used [8]. These number of hash tags keeps growing every day. It has created much effort to the user who just wishes to source for one piece of information. This work demonstrates that using clustering can sort data tidily. This work also argues that knowledge from tweets is useful if presented in a more organized manner.

2   Literature Review

Social big data are a valuable source for value creation [9, 10]. Research related to Technology–Organization–Environment and Diffusion of Innovations are the most popular theoretical models for big data [9]. That said, the social network (such as Twitter) is currently the primary means of information sources or communication around the world [11, 12]. Clustering the opinion from this big set of data to explore insights about Veganism is a value-creation technique and open to a new opportunity. Value creation is helpful in decision making, such as looking for patterns or trending practice [13]. Let get back to the topic, about Veganism. Is the only value being all about health?

Vegan that concern of animal welfare would relate meat with disgust and emotional anguish [14]. Their moral concern leads them to believe that all animal has a right to life and freedom [15]. A business that has a similar arm of code would tag along to influence society as a motive. Some vegan goes on social media with the motive to seek diet opinion. For instance, there are four types of vegan diet [16], 1) Lacto-vegetarian diet, 2) Lacto-Ovo vegetarian diets, 3) Pesco-vegetarian diets and 4) semi-vegetarian diets. The complexity, such as vegetarians consume animal-derived products while vegans strictly refrain from consuming animal-derived products [17]. Or some vegans only consume raw vegetables, nuts, fruits, seeds, sprouted grains and legumes [14]. Further, Lacto-Vegetarian consumes dairy and milk products, but not egg, red meat and fish [14, 18]. Lacto-Ovo Vegetarian consume dairy products and eggs but no fish or meat (e.g., red meat and poultry) [14, 16]. Pesco-vegetarian consumes fish and dairy products, but no red meat or poultry [19–21]. Semi-vegetarian (flexitarian) consume meat occasionally [22]. With all those different types of diets, information seeker that intended to ask for a single piece of knowledge, such as on special diets plan or opinion from tweets could become complicated.

Studies have associated the motives to go Vegan for several reasons. For instance, animal welfare [23, 24]; health-conscious [16, 20, 25, 26]; to sustain the planet [27–31]; to obtain a sense of belonging in social media or gaining social identity [32, 33];

or to endorse celebrities/society [14, 34, 35]. Are these opinions discussed on Twitter? The primary objective of this study is to explore the motives and opinions of Veganism in tweets. And the secondary aim is to demonstrate via clustering the opinions, the complexity to get a piece of single information is reduced, despite many hashtags.

3 Methods

3.1 Data Collection

A total of 50634 veganism related tweets posted between April 2019 to July 2019 are collected. Sixty-one vegan-related hash tags tweets include #vegan, #veganism, #plant-baseddiet, #plantbased, #govegan, etc. are collected. During the pre-processing, text is cleaned and filtered. Frequency of text was quantified. We listed words such as food (6797), animal (5352), vegetarian (4321), eat (4528) and vegan food (3330) as top 5 after the list of hashtags. We identified the parent node. We placed terms with synonym with parent name, for instance, terms such as'delicious' as a parent to replace terms such as 'delicious', 'yummyy', 'delicious', 'yuumy'. These are all done using SAS Enterprise Miner.

3.2 Clustering

We used Expectation-Maximization (E.M.) algorithm to group tweets into a disjoint cluster of documents and to provide the descriptive terms for each theme. This algorithm assumes that a mixture model approximates the data distribution by fitting k cluster density functions fh (h = 1, …, k), to a data set with d variables [36]. First, it finds initial parameter estimates. Then, it applies the standard or scaled version of the E.M. algorithm to identify primary clusters and to update parameter estimates. The algorithm stops when two consecutive log-likelihood values differ by a specific amount or when reach five iterations. This step aims to explore the connecting terms, to cluster into a central theme. The reason to select this algorithm is that this approach is different from clustering because clustering allocates each document to a single group while Text Topic node assigns a score for each material and term to each topic. Thresholds is used to identify if the association is strong enough to consider that the term belongs to the subject. Therefore, documents (tweets) may belong to more than one subject or none.

3.3 Interrater Reliability

We carried out Human Interrater Reliability test to find the percentage of agreement of association topics clustered. Human Coders is used in making value judgements. Several researchers have performed a similar analysis in their studies. For instance, to evaluate a performance of a system comparing the result with human coders [37]; to assess the irony value in Verbal Irony Procedure (VIP) [38]; to assess the degree of agreement on the hotel reviews from Trip Advisor [39]; and also to rate the sophisticated sarcasm term that was processed by sarcasm detection tools [40]. We have recruited three strict vegans with at least three years of experience to perform read across the text to verify

the reliability of the machine clustered set of discourse. This for a reason to ensure that the coder is conversant or experienced with the phenomenon of Veganism.

Based on the 90% of confidence level and margin error of 10, a total number of 265 tweets were sampled from the results to obtain the percentage of agreement between the association of Veganism discussed relevant [41]. Coders with knowledge of Veganism can increase the accuracy of the ratings. We invited three coders named as T, W, and Y with the condition that they have been committed to a strict vegan lifestyle for at least six years and 7 years. Coders T and W stated that the reason to become a vegan spark after the watch of vegan documentaries. Coder Y noted the main reason was a health concern. A a total of 265 tweets are given to each individual to ensure not influenced by one another. They have rated Score = 1 for Yes or score = 0 for No to indicate the agreement of tweet association with the vegan topic, respectively according to human judgement. Cohen's kappa coefficient (K) is used to measure inter-rater reliability [44]. The percent agreement is summed up and divides by the total number of observations [44]; we repeat this procedure for all the different I.D. of a topic.

4   Results

Concept links assist in understanding the association between the terms based on the co-occurrence of terms in the document. The thickness of the line shows the strength of the association between the terms. A thicker link between 2 terms demonstrates a stronger association than a thinner link. Each associated term can be expanded to observe its sub-linked terms to have a better idea of the association between the two significant terms.
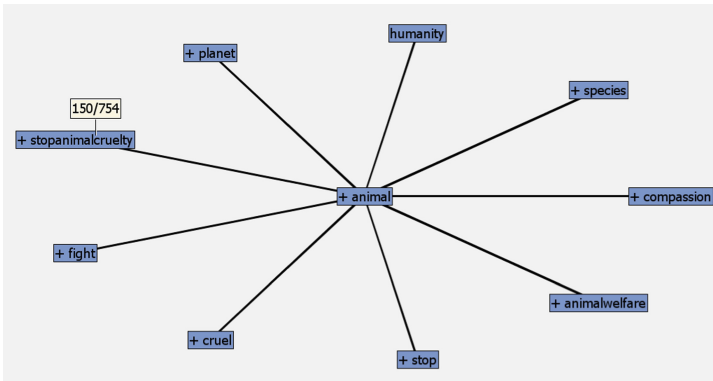

Fig. 1. An example of concept link for the term "animal."

Figure 1 shows an example of the term animal that appears 4989 times. Other terms that are highly related to this term are compassion, "stop animal cruelty", and cruel, indicating that vegans in twitter think that being vegan and can stop animal cruelty by not consuming or using any animal products. The term such as "compassion" is strongly associated with the term "animal" which indicates that a vegan diet increases compassion

for the animal. A total of 150 tweets contain both terms "stop animal cruelty" and "animal", and 754 tweets includes the term "stop animal cruelty". SAS enterprise miner allows all the terms to be explored individually. To simplify the results presentation, Table 1 presented the four sets of highly distributed clusters.

Table 1. Results from clustering and human coders

Cluster-ID	Terms	Percentage	Description	Agreement of human coders (%)
1	Animal, meat, consume, animal welfare, love, dairy, cruelty-free, stop animal cruelty, go cruelty-free, eat plant-based, ditch dairy, milk, product	19.3%	Related to compassion for animals, ditching dairy and adhering plant-based diet can protect animal rights and stop animal cruelty	87%
2	People, time, knowledge, life, today, want, year, live, first, look, world, thing, help, week	22.8%	Related to the lifestyle or duration/period/time of adopting a vegan diet	89%
3	Vegan food, vegetarian, delicious, gluten-free, healthy, vegans of instagram, organic, vegetable, what vegans eat, vegan life, vegan food share, healthy food, salad, foodie	32.3%	Related to what they eat and healthy nutritious discussion	83%
4	Good, food, day, consume, recipe, diet, plant, burger, vegetarian, health, delicious, healthy, free, base, plant-base	25.5%	Related to vegan recipes and option. The users mentioned that vegan food is deliciously healthy	87%

We gave each of the clusters an I.D., we explained in the following paragraph with some sample of exhibits of tweets that we coded to the specific class.

Cluster-ID 1, "Animal Welfare" is assigned for this discovery. The discussion is about going Veganism with the opinion that animals have a right to life and freedom [43]. A total of 87% agreement is reached on this association to this cluster.

Sample exhibit of tweets selected from Cluster-ID 1:

- 300,000 animals are killed for food every minute!
- Do you think that animals don't feel pain?
- Some of the horrible methods that are used to torture animals, so that people can eat them! Please Go Vegan

Cluster-ID 2, "Environment" is assigned for this discovery. The discussion mainly about stops consuming animal products. Adhere to the plant-based diet, which can reduce the devastating impact of the dairy industry on the environment. Livestock rearing results in global warming through the methane gas the animals produce [44]. There are 89% of the agreement reached on the association between Veganism by human raters. Sample exhibit of tweetsselected from Cluster-ID 2:

- We all love animals. Why do we call some 'pets' and others' dinner?
- I made the most yumtastic maple and sriracha roasted chickpeas omg soooooo delicious. Great for snacking, salads or rice bowls.

Cluster-ID 3, "Health" is assigned for this discovery. The keywords indicate that there is a discussion about following a vegan diet can improve body wellness, aid weight loss and diabetes. A total of 83% of the agreement is reached on the association with health. Sample exhibit tweets selected from Cluster-ID 3

- The health benefits of a vegan diet are life-changing. And to know the food on your plate has not killed innocent animals is satisfying. Coming this April inspired by the power of the plate and heart.
- Eating A Good Is Great For Your Health I'm Just Going All Out Before I Go Fully Vegan

Cluster-ID 4, "Recipe" is assigned for this discovery. A total of 87% of the text were agreed with association adhering Veganism with a recipe or other vegan option. Association between Veganism and vegan option is high because vegan options are ubiquitous nowadays [45]. There was very little vegan food to choose during the 1990s [45]. However, things have changed over the past 20 years due to the increase of people who adopt plant-based diets; restaurants have been evolved to satisfy customers' demands and appetites. Some tweets also mentioned that "Every takeaway place now does some vegan food … every pub has a vegan menu now … every main street has a restaurant where you can be guaranteed to eat vegan food'. In 2018, restaurants and fast-food chains that added vegan options to their menus were Shake Shack, McDonald's, Ikea, Impossible Burger, A&W and so-on [46].

Sample exhibit of tweets selected from Cluster-ID 4:

- Delicious Vegan Recipes That Are Affordable
- Had an Impossible Burger today! Soooo good Can't wait until this comes to Canada.

Finally, in this study, there is no I.D. clustered that have any relation to endorsing the celebrities/society as the motive to become vegan. Besides that, assembling this topic not only informing us about the motives of becoming vegan but also exploring the opinions of what people talk about Vegan in tweets, in this case for that specific period in our data collection.

5   Conclusion

This article aims to demonstrate the concept of ability to sort information from tweets in a better-organized manner. It can display organized information aggregated from different postings around the world using the relevant hashtag.

A further application for it can be various. For instance, tweets concerning animal welfare with photos and videos link can be themed and organized into a more structured and up-to-date education online e-book. Cruelty-free products companies can extend their service to introduce non-animal-based products in the same platform online. A data scientist can refer to this platform to identify research opportunity, such as if nonanimal products may contain harmful ingredients. Brands like The Body Shop, Lush, Aveda, Urban Decay and others may use this platform as the channel to advertise and share their products information. Non-profit organizations such as AHSM (Association for Healthcare Social Media), can come in to promote Veganism for health. World Wildlife Fund (WWF) can seek supports and solutions to protect forests and secure water and to discuss environmental awareness [47] through becoming a vegan. A restaurant such as Impossible Burger, could offer more vegan options and announce in this platform. All with the same goals, to drive the whole nation about Veganism in one platform. United Nation of Global Sustainable Development Goal (About the Sustainable Development Goals - United Nations Sustainable Development, n.d.), can use this as an indicator to connect to their activities and grant opportunity. Government to share policy and guideline into this platform. In terms of practical implication, the influencer has many advantages to business and services in promoting their products and services. The implementation of this clustered I.D. can provide products influencers and idea to share opinions using similar I.D., this can shorten the information acquisition steps of opinion seekers. For this case, brands such as The Body Shop, Lush, Aveda, Urban Decay and others may use this information as the channel to communicate their products and services.

The limitation of this study is that only one technique of clustering is being used to present the overall opinions. Future research can use others type of clustering technique and provide a new set of representation, another limitation of our study is that the number of tweets is only collected from this specific period of time, if the technique can be continuous using machine language to find any difference of the result from this set of I.D., it would be beneficial.

Finally, similarly used of this concept could be replicated on other issues or another movement, such as tweets that are discussing COVID-19, treatments, methods to avoid and adhere, products to use and policy and news. Real-time posting of tweets into one category, and reorganized into an individual perspective using artificial intelligence, provide a proper referencing platform to drive or tackle a particular movement.

References

1. Suddath, C.: A brief history of veganism. Time Magazin (2008)
2. Leahy, E., Lyons, S., Tol, R.S.J.: An Estimate of the Number of Vegetarians in the World. An Estimate of the Number of Vegetarians in the World, vol. 340, pp. 1–44 (2010)
3. Vegan Variety: Plant-based Foods Trend. https://www.preparedfoods.com/articles/119792-vegan-variety-plant-based-foods-trend. Accessed 08 May 2018
4. Shane-Simpson, C., Manago, A., Gaggi, N., Gillespie-Lynch, K.: Why do college students prefer Facebook, Twitter, or Instagram? Site affordances, tensions between privacy and selfexpression, and implications for social capital. Comput. Hum. Behav. **86**, 276–288 (2018)
5. Kim, J., Brossard, D., Scheufele, D.A., Xenos, M.: Shared information in the age of big data: exploring sentiment expression related to nuclear energy on twitter. J. Mass Commun. Q. **93**(2), 430–445 (2016)
6. Sander, T., Teh, P.L., Sloka, B.: Your social network profile reveals you. Int. J. Web Inf. Syst. **13**(1), 14–24 (2017)
7. Teh, P.L., Huah, L.P., Si, Y.W.: The intention to share and re-shared among the young adults towards a posting at social networking sites. Adv. Intell. Syst. Comput. **275**, 13–21 (2014)
8. Hashtags for #vegans on Instagram, Twitter, Facebook, Tumblr| best-hashtags.com. http://best-hashtags.com/hashtag/vegans. Accessed 01 April 2020
9. Baig, M.I., Shuib, L., Yadegaridehkordi, E.: Big data adoption: state of the art and research challenges. Inform. Process. Manag. **56**(6), 102095 (2019)
10. Del Vecchio, P., Mele, G., Ndou, V., Secundo, G.: Creating value from social big data: implications for smart tourism destinations. Inf. Process. Manage. **54**(5), 847–860 (2018)
11. Ficamos, P., Liu, Y.: A topic based approach for sentiment analysis on twitter data. Int. J. Adv. Comput. Sci. Appl. **7**(12), 201–205 (2016)
12. Priya, S., Sequeira, R., Chandra, J., Dandapat, S.K.: Where should one get news updates: Twitter or Reddit. Online Soc. Netw. Media **9**, 17–29 (2019)
13. Saggi, M.K., Jain, S.A.: Survey towards an integration of big data analytics to big insights for value-creation. Inf. Process. Manage. **54**(5), 758–790 (2018)
14. Phua, J., Jin, S.V., Kim, J.: The roles of celebrity endorsers' and consumers' vegan identity in marketing communication about Veganism. Journal of Marketing Communications, pp. 1–23 (2019)
15. Compassion for animals, The Vegan Society: Between the Species. https://www.vegansociety.com/sites/default/files/CompassionForAnimals.pdf. Accessed 10 April 2019
16. Tonstad, S., Butler, T., Yan, R., Fraser, G.E.: Type of vegetarian diet, body weight, and prevalence of type 2 diabetes. Diab. Care **32**(5), 791–796 (2009)
17. Waters, J.: A model of the dynamics of household vegetarian and vegan rates in the U.K. Appetite **127**, 364–372 (2018)
18. Praharaj, A.B., Goenka, R.K., Dixit, S., Gupta, M.K., Kar, S.K., Negi, S.: Lacto vegetarian diet and correlation of fasting blood sugar with lipids in population practicing sedentary lifestyle. Ecol. Food Nutr. **56**(5), 351–363 (2017)
19. Clarys, P., Deliens, T., Huybrechts, I., Deriemaeker, P., Vanaelst, B., De Keyzer, W., Hebbelinck, M., Mullie, P.: Comparison of nutritional quality of the vegan, vegetarian, semi-vegetarian, pesco-vegetarian and omnivorous diet. Nutrients **6**(3), 1318–1332 (2014)
20. Tonstad, S., Nathan, E., Oda, K., Fraser, G.E.: Prevalence of hyperthyroidism according to type of vegetarian diet. Publ. Health Nutr. **18**(8), 1482–1487 (2015)
21. Mihrshahi, S., Ding, D., Gale, J., Allman-Farinelli, M., Banks, E., Bauman, A.E.: Vegetarian diet and all-cause mortality: evidence from a large population-based australian cohort - the 45 and up study. Prev. Med. **97**, 1–7 (2017)

22. Forestell, C.A.: Flexitarian diet and weight control: healthy or risky eating behavior? Front. Nutr. **5**, 1–6 (2018)
23. Radnitz, C., Beezhold, B., DiMatteo, J.: Investigation of lifestyle choices of individuals following a vegan diet for health and ethical reasons. Appetite **90**, 31–36 (2015)
24. Edmundson, W.A.: Do animals need citizenship? Int. J. Const. Law **13**(3), 749–765 (2015)
25. Appleby, P.N., Davey, G.K., Key, T.J.: Hypertension and blood pressure among meat eaters, fish eaters, vegetarians and vegans in EPIC–Oxford. Publ. Health Nutr. **5**(05), 645–654 (2003)
26. Tantamango-Bartley, Y., Knutsen, S.F., Knutsen, R., Jacobsen, B.K., Fan, J., Lawrence Beeson, W., Sabate, J., Hadley, D., Jaceldo-Siegl, K., Penniecook, J., Herring, P., Butler, T., Bennett, H., Fraser, G.: Are strict vegetarians protected against prostate cancer? Am. J. Clin. Nutr. **103**(1), 153–160 (2016)
27. Henning, B.: Standing in livestock's "long shadow": the ethics of eating meat on a small planet. Ethics Environ. **16**(2), 63 (2011)
28. Doyle, J.: Celebrity vegans and the life styling of ethical consumption. Environ. Commun. **10**(6), 777–790 (2016)
29. Rosi, A., Mena, P., Pellegrini, N., Turroni, S., Neviani, E., Ferrocino, I., Di Cagno, R., Ruini, L., Ciati, R., Angelino, D., Maddock, J., Gobbetti, M., Brighenti, F., Del Rio, D., Scazzina, F.: Environmental impact of omnivorous, ovo-lacto-vegetarian, and vegan diet. Sci. Rep. **7**(1), 1–9 (2017)
30. Baroni, L., Filippin, D., Goggi, S.: Helping the planet with healthy eating habits. Open Inform. Sci. **2**(1), 156–167 (2018)
31. Chai, B.C., van der Voort, J.R., Grofelnik, K., Eliasdottir, H.G., Klöss, I., Perez-Cueto, F.J.A.: Which diet has the least environmental impact on our planet? A systematic review of vegan, vegetarian and omnivorous diets. Sustain. (Switz.) **11**(15), 4110 (2019)
32. Yang, C.-C., Holden, S.M., Carter, M.D.K.: Social media social comparison of ability (but not opinion) predicts lower identity clarity: identity processing style as a mediator. J. Youth Adolesc. **47**(10), 2114–2128 (2018)
33. Davis, J.L., Love, T.P., Fares, P.: Collective social identity: synthesizing identity theory and social identity theory using digital data. Soc. Psychol. Q. **82**(3), 254–273 (2019)
34. Bratu, S.: The phenomenon of image manipulation in advertising. Econ. Manag. Financ. Markets **5**(2), 333–338 (2010)
35. Lundahl, O.: From a moral consumption ethos to an apolitical consumption trend: the role of media and celebrities in structuring the rise of Veganism. University of Vaasa (2017)
36. Getting Started with SAS® Text Miner 12.1. SAS Institute Inc (2008)
37. Lee, K.J., Choi, Y.S., Kim, J.E.: Building an automated English sentence evaluation system for students learning English as a second language. Comput. Speech Lang. **25**(2), 246–260 (2011)
38. Burgers, C., van Mulken, M., Schellens, P.J.: Finding irony: an introduction of the verbal irony procedure (VIP). Metaphor Symbol **26**(3), 186–205 (2011)
39. Khoo, F.S., Teh, P.L., Ooi, P.B.: Consistency of online consumer's perceptions of posted comments: an analysis of trip advisor reviews. J. Inform. Commun. Technol. **2**(2), 374–393 (2017)
40. Teh, P.L., Ooi, P.B., Chan, N.N., Chuah, Y.K.: A comparative study of the effectiveness of sentiment tools and human coding in sarcasm detection. J. Syst. Inf. Technol. **20**(3), 358–374 (2018)
41. Sample Size Calculator. In SurveyMonkey. https://www.surveymonkey.com/mp/samplesize-calculator/
42. McHugh, M.L.: Interrater reliability: the kappa statistic. Biochemia Med. **22**(3), 276–282 (2012)
43. Compassion for animals, The Vegan Society (2018). https://www.vegansociety.com/sites/default/files/CompassionForAnimals.pdf. Accessed 08 April 2020

44. Plant-based diet can fight climate change - U.N. BBC News, August. https://www.bbc.com/news/science-environment-49238749. Accessed 08 April 2020
45. Mawunyo, G.: Vegan options are everywhere nowadays as restaurants evolve to meat demand. ABC Radio Sydney. https://www.abc.net.au/news/2019-05-18/restaurants-takenote-as-more-people-go-vegan/11119160. Accessed 08 April 2020
46. Sharon, V.: 15 Vegan Options Added to Restaurants and Fast Food Chains in 2018. ABC News. https://www.onegreenplanet.org/vegan-food/15-vegan-options-added-to-restaurantsand-fast-food-chains-in-2018/. Accessed 08 April 2020
47. Ganga, S.D.: Social media and the rise of Vegan. Experiencing Public Relations (2018)