

# yoo\_2022\_exploring\_the\_nexus\_between\_food\_and\_vegan\_lifestyle\_via\_text\_mining\_based\_online\_community\_analytics

## Year

2022

## Author(s)

Ra Yoo and Seo-Young Kim and Do-Hee Kim and Jiyoung Kim and Ye Ji Jeon and Jung Han Yoon Park and Ki Won Lee and Hee Yan

## Title

Exploring the nexus between food and veg\*n lifestyle via text mining-based online community analytics

## Venue

Food Quality and Preference

---

## Topic labeling

Manual

## Focus

Secondary

## Type of contribution

Established approach

## Underlying technique

Manual labeling

## Topic labeling parameters

\

## Label generation

In order to establish the labels of topics, four researchers reviewed the labels generated based on the keywords of each topic and the final labels were determined.

The nine subcategories of the Food cluster were labeled with themes that best describe the characteristics: Topic 1 as Social conflicts associated with dietary restriction (24.33%), Topic 2 as Transition tips for diet shifts (22.02%), Topic 3 as Concerns related to eating out (11.34%), Topic 4 as Maintaining a balanced diet (8.92%), Topic 5 as Ambiguity of food labels (8.33%), Topic 6 as Shopping tips for veg\*n friendly items (8.09%), Topic 7 as Options of dairy alternatives (6.38%), Topic 8 as Easy meal ideas (5.87%), and Topic 9 as Dilemmas associated with meat analogues (4.73%).

## Motivation

\

---

## Topic modeling

LDA

## Topic modeling parameters

Performed on cluster 1 (Food)

Nr of topics: 9 (and neighbouring values)

## Nr. of topics

9

---

## Label

Manually assigned single or multi word labels

## Label selection

\

## Label quality evaluation

\

## Assessors

\

---

## Domain

Paper: Nutrition

Dataset: Social media (Reddit)

## Problem statement

This study investigated two vegan subreddits, r/Vegan and r/Vegetarian, using multiple text mining techniques to classify users' interests and preferences.

## Corpus

Origin: Reddit

Nr. of documents: 109,274

Details:

- Posts from r/Vegan and r/Vegetarian uploaded between January 1, 2016, and December 30,

## Document

Reddit post

## Pre-processing

- eliminating the posts with no subtitles or bodies and deleted post
  - convert the posts into lowercase English and to remove special characters (e.g. @,!), and URLs.
  - tokenization
  - lemmatisation
  - Stop words were eliminated
  - tokens with a length of 3 characters or less were removed
- 
- Clustering of documents using k-means and tf-idf into: Food, Perception, Health, Altruism, Emotion, and Situation

---

```
@article{yoo_2022_exploring_the_nexus_between_food_and_vegan_lifestyle_via_text_mining_based_online_community_analytics,
```

```
    abstract = {Attention towards veg*nism is increasing as the impact of food choices on health and sustainability as well as ethical concerns regarding animal welfare emerge. Although online user analysis is an effective tool to obtain practical insights without geographical constraints, implementation on a large veg*n population has been carried out within a limited scope. This study investigated two veg*n subreddits, r/Vegan and r/Vegetarian, using multiple text mining techniques to classify users' interests and preferences. Based on K-means and term frequency-inverse document frequency, six clusters were identified: Food, Perception, Health, Altruism, Emotion, and Situation. The proportion of each cluster and keywords representing the clusters were obtained. Being a major sector, further assessment of the Food cluster was conducted using Latent Dirichlet Allocation topic modeling technique. Confusion was observed in relation to being pressured with sudden changes in dietary patterns, including meal composition, preparation, and shopping routines. The results also revealed barriers to transition for individuals who have recently started veg*n diets, and those wishing to switch to stricter dietary patterns. In addition to difficulties relating to economic and social aspects, our findings suggest that the establishment of detailed guidelines may help accommodate the various
```

dietary compositions across the veg\*n spectrum, and clear information relating to veg\*n food products is needed from manufacturers.},

author = {Ra Yoo and Seo-Young Kim and Do-Hee Kim and Jiyoung Kim and Ye Ji Jeon and Jung Han Yoon Park and Ki Won Lee and Hee Yang},

date-added = {2023-04-20 22:01:25 +0200},

date-modified = {2023-04-20 22:01:25 +0200},

doi = {https://doi.org/10.1016/j.foodqual.2022.104714},

issn = {0950-3293},

journal = {Food Quality and Preference},

keywords = {Vegan, Vegetarian, Online communities, Consumer, Perception, Text mining},

pages = {104714},

title = {Exploring the nexus between food and veg\*n lifestyle via text mining-based online community analytics},

url = {https://www.sciencedirect.com/science/article/pii/S0950329322001896},

volume = {104},

year = {2022}}

#Thesis/Papers/FS