



amazon discourses

An analysis of 'greenness' in product descriptions of household products and their online marketing

Group project presentation

Šimon Trlifaj, Barbora Bromová, Markéta Ovečková

30/03/2023

Table of contents

01

Introduction & Hypothesis

02

Dataset & Method

Web scraping of product descriptions from Amazon

03

Exploratory Analysis

Average frequency of greenwashing word combinations

04

Word Embeddings

'Greenness' in latent vector space

05

Topic Modelling

Latent Dirichlet Allocation (LDA)

06

Potential continuation & main findings

Image & marketing analysis



01

Introduction & Hypothesis

Introduction & Hypothesis

- Assumption that “green products” (eg. products whose word embeddings are closer to a ‘green ideal’) do not need an extensive description to persuade consumers
- Certain products, e.g. detergents, are traditionally perceived as harmful for the environment
- To what extent do Amazon suppliers attempt to compensate for *a priori* non-eco-friendly products association in product descriptions?

02

**Dataset
& Method**

Data Collection

- Amazon has API, but we did not get access
- We web scraped 4,014 product pages in 71 categories from Amazon
- For each product, we have its name, description, search keyword, link, and webpage

Bleach tablets
Kitchen cleaner
Dishwasher detergent
Window spray
Stain eliminator
Machine descaler
Mouthwash
Contact Lens Solutions
Prescription Medications
Fabric Softener
Razor
Deodorant
...

Web Scraping

Step 1

amazon Deliver to Simon Wien 1100 All Kitchen cleaner EN Hello, Simon Account & Lists Returns & Orders Cart

All Today's Deals Customer Service Buy Again Simon's Amazon.com Gift Cards Registry Browsing History Sell

1-48 of over 5,000 results for "Kitchen cleaner" Sort by: Featured

More-sustainable Products
☐ Climate Pledge Friendly

Department
Household Cleaning
All-Purpose Household Cleaners
Household Degreasers
Disinfectant Sprays & Liquids
Dishwasher & Garbage Disposal Cleaners
Commercial All-Purpose Cleaners
Household Cleaning Tools
Household Surfaces Scouring Pads & Sticks

Customer Reviews
★★★★☆ & Up
★★★★☆ & Up
★★★★☆ & Up
★★★★☆ & Up

Brands
☐ Lysol
☐ MRS. MEYER'S CLEAN DAY
☐ Easy Off
☐ Mr. Clean
☐ Clorox
☐ Method
☐ Fantastik
See more

Price
... ..

Check each product page for other buying options. Price and other details may vary based on product size and color.

Mrs. Meyer's All-Purpose Cleaner
Spray, Lemon Verbena, 16 fl. oz - Pack of 3
Liquid - 16 Fl Oz (Pack of 3)
4.8 ★★★★★ ~ (36,261)
\$14.97 (\$0.31/Fl Oz)
Save 50% on 1 when you buy 2
Delivery Thu, Apr 13
Ships to Austria
Climate Pledge Friendly

Foaming Heavy Oil Stain Cleaner, All Purpose Bubble Cleaner
Kitchen Deep Cleaning Spray, All...
Foam - 3.52 Fl Oz (Pack of 2)
2.2 ★☆☆☆☆ ~ (547)
\$23.99 (\$0.06/Ounce)
Delivery Fri, Apr 14
Ships to Austria
Only 10 left in stock - order soon.
Options: 2 sizes

Easy Off Specialty Kitchen Degreaser Cleaner, Clear, Lemon, 16 Fl Oz
Aerosol - 16 Fl Oz (Pack of 1)
4.4 ★★★★★ ~ (13,793)
\$7.09 (\$0.44/Fl Oz)
Save 5% on any 4 or more
Ships to Austria
More Buying Choices
\$5.39 (4 new offers)

Step 2

amazon Deliver to Simon Wien 1100 All Kitchen cleaner EN Hello, Simon Account & Lists Returns & Orders Cart

All Today's Deals Customer Service Buy Again Simon's Amazon.com Gift Cards Registry Browsing History Sell

Back to results

Lysol Pro Kitchen Spray Cleaner and Degreaser, Antibacterial All Purpose Cleaning Spray for Kitchens, Countertops, Ovens, and Appliances, Citrus Scent, 22oz (Pack of 2)
Visit the Lysol Store
★★★★★ 4,607 ratings | 26 answered questions

List Price: \$8.40 Details
Price: **\$7.94** (\$0.18 / fl oz)
You Save: \$0.46 (5%)
No Import Fees Deposit & \$15.50 Shipping to Austria Details
Save 5% on any 4 or more Shop qualifying items

Brand Lysol
Item Form Liquid
Scent Citrus
Material Feature Spray
Item Volume 22 Fluid Ounces

About this item

- ELIMINATES 99.9% OF KITCHEN VIRUSES AND BACTERIA INCLUDING SALMONELLA AND E.COLI: Lysol Pro Kitchen Cleaner is tested and proven to eliminate 99.9% of viruses and bacteria* (when used as directed). EPA Reg #777-91.
- CUTS THROUGH TOUGH, BAKED ON GREASE: Easily penetrates grease and grime to clean even the toughest kitchen messes, leaving nothing but a fresh citrus scent.
- NO HARSH CHEMICAL RESIDUE: Ideal for counter tops, tables, oven exteriors, highchairs, and kitchen exterior appliances* (when used as directed, rinse food contact surfaces with potable water).
- EASY TO USE KITCHEN SPRAY: Spray on your greasy stains and

One-time purchase:
\$7.94 (\$0.18 / fl oz)
No Import Fees Deposit & \$15.50 Shipping to Austria Details
Delivery April 21 - 27
Or fastest delivery April 14 - 26
Deliver to Simon - Wien 1100

Qty: 1

Add to Cart

Buy Now

Secure transaction

Ships from Amazon.com
Sold by Amazon.com

Details

Return policy: Eligible for Return, Refund or Replacement within 30 days of receipt.
☐ Add a gift receipt for easy returns

Subscribe & Save: 5% / 15%
\$7.54 (\$0.17 / fl oz)
First delivery on Apr 21
Ships from: Amazon.com
Sold by: Amazon.com

Add to List

03

Exploratory Analysis

Average frequency of green description expressions

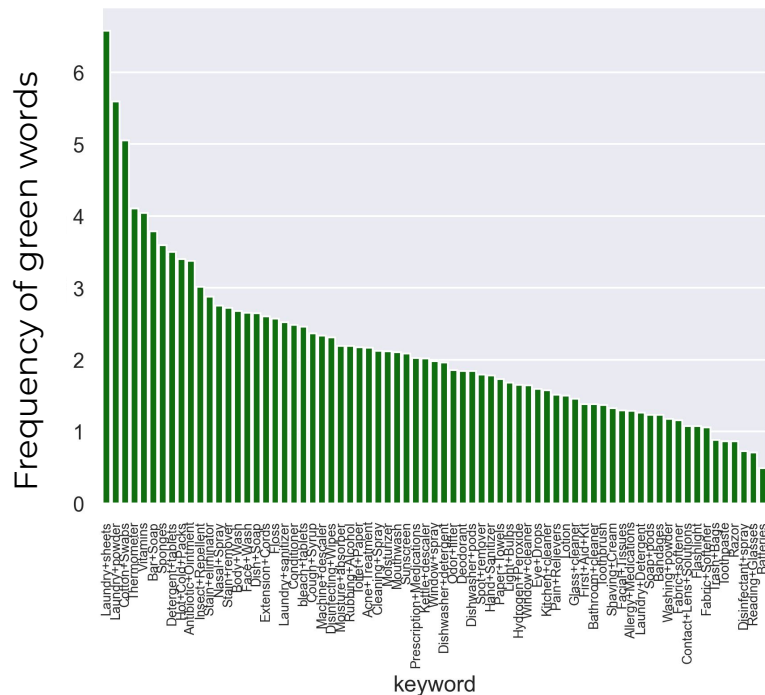
We counted the frequency of occurrence of adjectives that may be used for greenwashing

eco friendly
earth friendly
environmentally friendly
eco
bio
not tested on animals
reusable

gentle
probiotic
botanical
natural
plant based
plant-based
responsible

responsible
responsibility
certified
durable
no chemicals
non gmo
...

Average frequency of green description expressions



keyword

Laundry+sheets 6.578947

Laundry+powder 5.592593

Detergent+tablets 3.500000

Stain+eliminator 2.875000

Stain+remover 2.722222

Laundry+sanitizer 2.520833

bleach+tablets 2.458333

Machine+descaler 2.333333

Moisture+absorber 2.192308

Kettle+descaler 2.020833

04

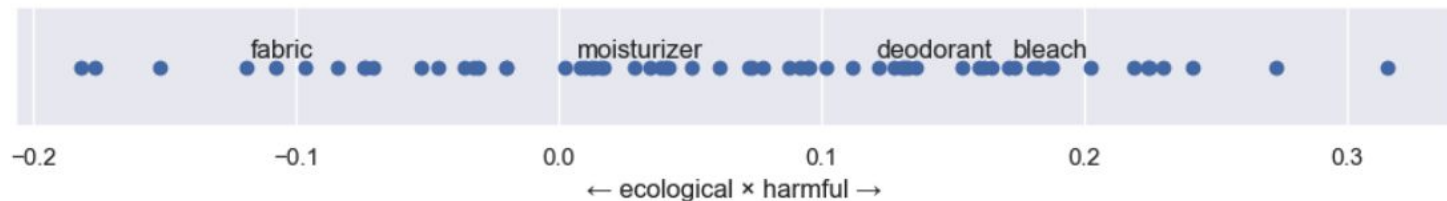
Word Embeddings

Variable 1: Negative associations with products

We ranked the perception of a product being ecological on a vector in a word embedding space

ecology
environment
sustainable
renewable
nature

pollution
emission
chemical
deforestation
smog



Variable 2: Frequency of positive adjectives

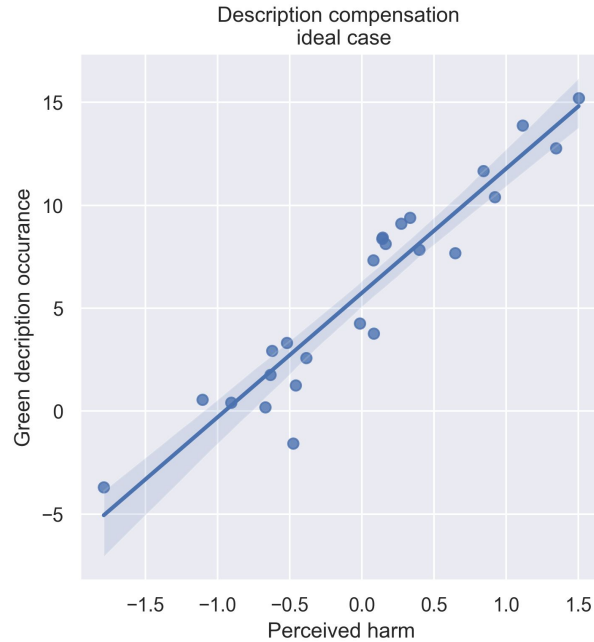
We counted the frequency of occurrence of adjectives that may be used for greenwashing

eco friendly
earth friendly
environmentally friendly
eco
bio
not tested on animals
reusable

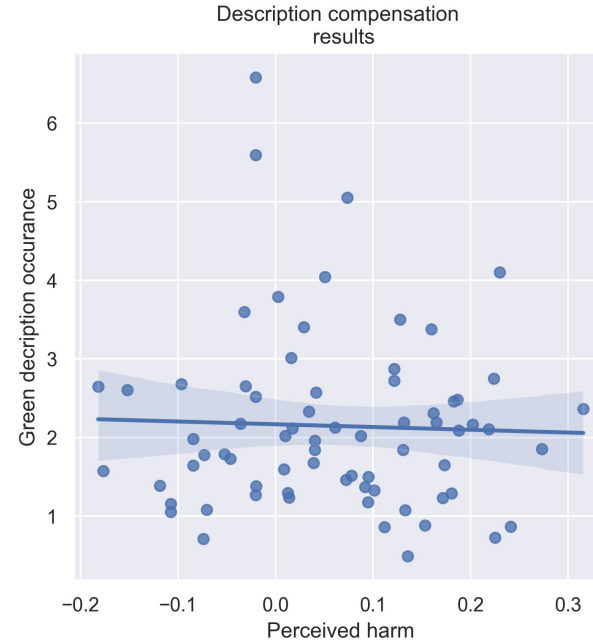
gentle
probiotic
botanical
natural
plant based
plant-based
responsible

responsible
responsibility
certified
durable
no chemicals
non gmo
...

What we hoped for



what did we get



05

**Topic
Modelling**

Latent Dirichlet Allocation (LDA)

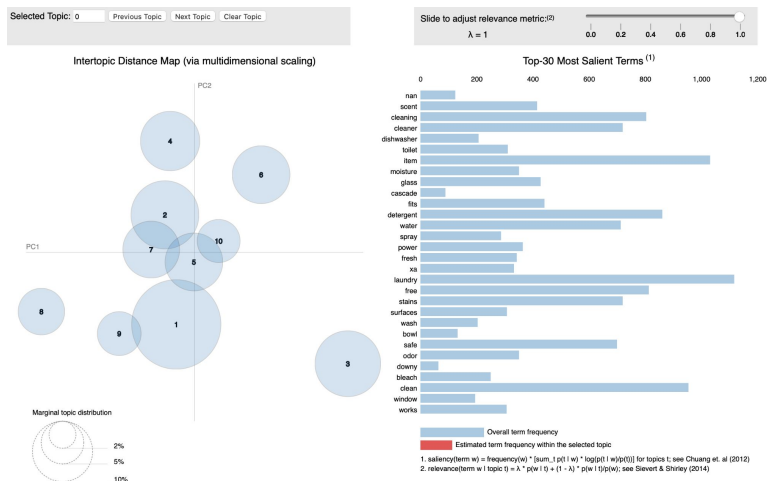
- Generative (unsupervised) topic model: grouping data by 'unobserved groups' or latent topics by classifying tokens.
 - Topics themselves can be explored and visualised (eg. *pyLDAvis*)
 - Documents can be assigned to topics
 - Synthetic documents can be created (reflecting statistical characteristics of original corpus)
- Is there an identifiable 'environmental' topic in our data?

Exploratory analysis - WordCloud

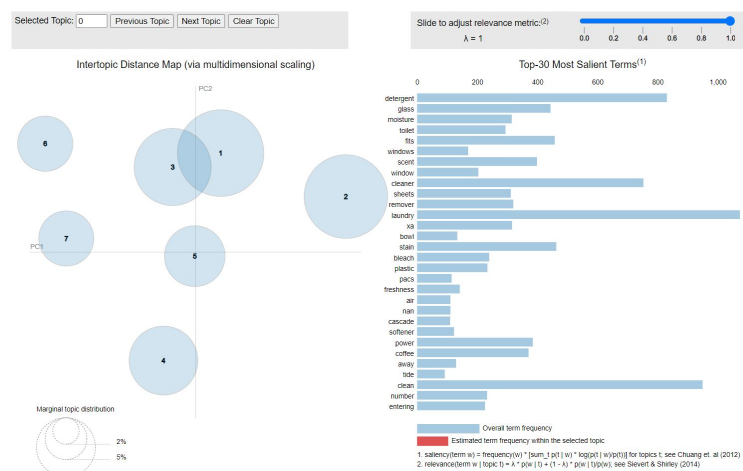


Helpful for adding corpus-specific stopwords.

Initial configuration



Optimised topics



Finding Topics & Intertopic Distance

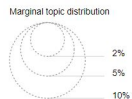
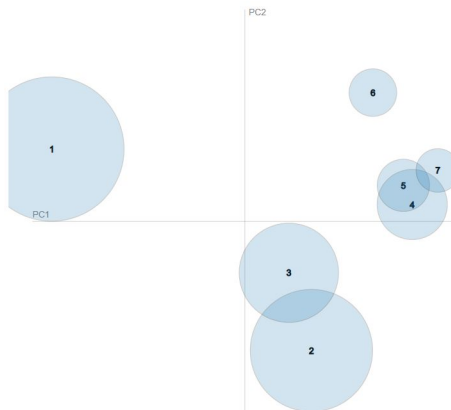
- Initial coherence score:
34.3%

- Improved coherence:
59.04%

- 7 topics
- $\alpha = 0.31$
- $\beta = 0.91$

Selected Topic: 0 Previous Topic Next Topic Clear Topic

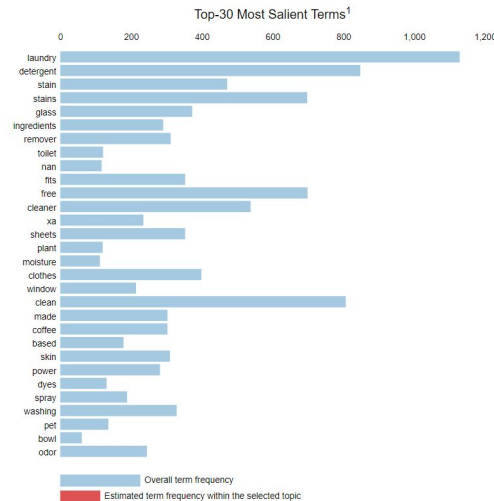
Intertopic Distance Map (via multidimensional scaling)



Slide to adjust relevance metric:⁽²⁾

$\lambda = 1$

0.0 0.2 0.4 0.6 0.8 1.0



1. saliency(term w) = frequency(w) * $\left[\sum_t p(t | w) * \log(p(t | w) / p(t)) \right]$ for topics t; see Chuang et. al (2012)
2. relevance(term w | topic t) = $\lambda * p(w | t) + (1 - \lambda) * p(w)$; see Sievert & Shirley (2014)

Cluster 1

Cleanliness and safety (eg. 'clean', 'water', or 'laundry').

- Prominent keyword = 'scent'.

Cluster 2

Chemical and infrastructural topic (eg. 'stain', 'glass', 'odor').

- Prominent keyword = 'odor'.

Cluster 3

Harmful topic (eg. 'bacteria', 'kill', 'toilet', 'excess').

- Prominent keyword = 'air', also 'moisture'.

06

Potential
future
steps

Image & marketing analysis

- Types of colours and design used on packaging, scents, visuals



Conclusion

What did we learn?

- Our 'green compensation' assumption doesn't appear in the data: compensatory words are distributed all over
 - Not an overall trend, but rather brand- or product-specific? Some occurrence of 'green' outliers.
- Distinct topics exist, but they seem to be more centred around health of users (rather than health of environment)
- Could qualitative analysis give us different insights?

“You may think you're washing green,
but maybe you're just being
greenwashed.”

