This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 776503
- CIRC4Life-H2020-IND-CE-2016-2017/CIRC-2017/TwoStage

A circular economy approach for lifecycles of products and services

Development of the model for online mining consumer views for eco-products

Deliverable 3.3

PROJECT INFORMATION	
Type of Project	European Commission Horizon 2020
Topic	CIRC-01-2016-2017 Systemic, eco-innovative approaches for the circular economy: large-scale demonstration projects
Grant Agreement No.	776503
Project Duration	01/05/2018 – 30/04/2021 (36 months)
Project Coordinator	Nottingham Trent University (NTU)
Project Partners	Enviro Data (ENV), Jonathan Michael Smith (JS), Kosnic Lighting Limited (KOS), Centre of Research for Energy Resources and Consumption (CIR), European EPC Competence Center GmbH (EECC), The Institute for Ecology of Industrial Areas (IETU), SWEREA IVF AB (SWE), Make Mothers Matter (MMM), ONA PRODUCT (ONA), INDUMETAL Recycling (IND), GS1 Germany GMBH (GS1G), Laurea University of Applied Science (LAU), Centre for European Policy Studies (CEPS), Institute of Communication and Computer Systems (ICCS), Recyclia (REC), S.A.T. Alia (ALIA)

DOCUMENT INFORMATION	
Title	Development of the model for online mining consumer views for eco-product
Version	V1.0
Release Date (dd.mm.yy)	28.12.18
Work Package	WP3
Type	Other
Dissemination Level	PU

DOCUMENT AUTHORS AND AUTHORISATION	
Document Responsible	Zijian Chai (NTU), Wenjie Peng (NTU), You Wu (NTU)
Contributors	NTU
Reviewed by	ALIA, ICCS, LAU
Approved by	Daizhong Su

DOCUMENT HISTORY			
Version	Date (dd.mm.yy)	Description	Implemented by
V0.1	10.12.18	First draft is produced.	Zijian Chai, Wenjie Peng
V0.2	17.12.18	Several changes were made throughout the document based on the comments given by ICCS, ALIA and LAU	Wenjie Peng, Zijian Chai
V1.0	28.12.18	Up-date and final revision.	Wenjie Peng, Zijian Chai, You Wu

Executive Summary

The type of this deliverable is 'Other' *(European Commission, 2017)*, which is the key diagram shown in Fig. 1. The diagram describes the model developed in Task 3.5, supported by relevant graphs, images and text of explanation, which are detailed in the following sections.

This model is developed to collect consumer's review data from e-commerce websites, and further analyse the consumer's requirements on product features. As shown in Fig.1, the model consists of three modules: Data scraping, Data mining/processing, and Consumer requirement mapping.

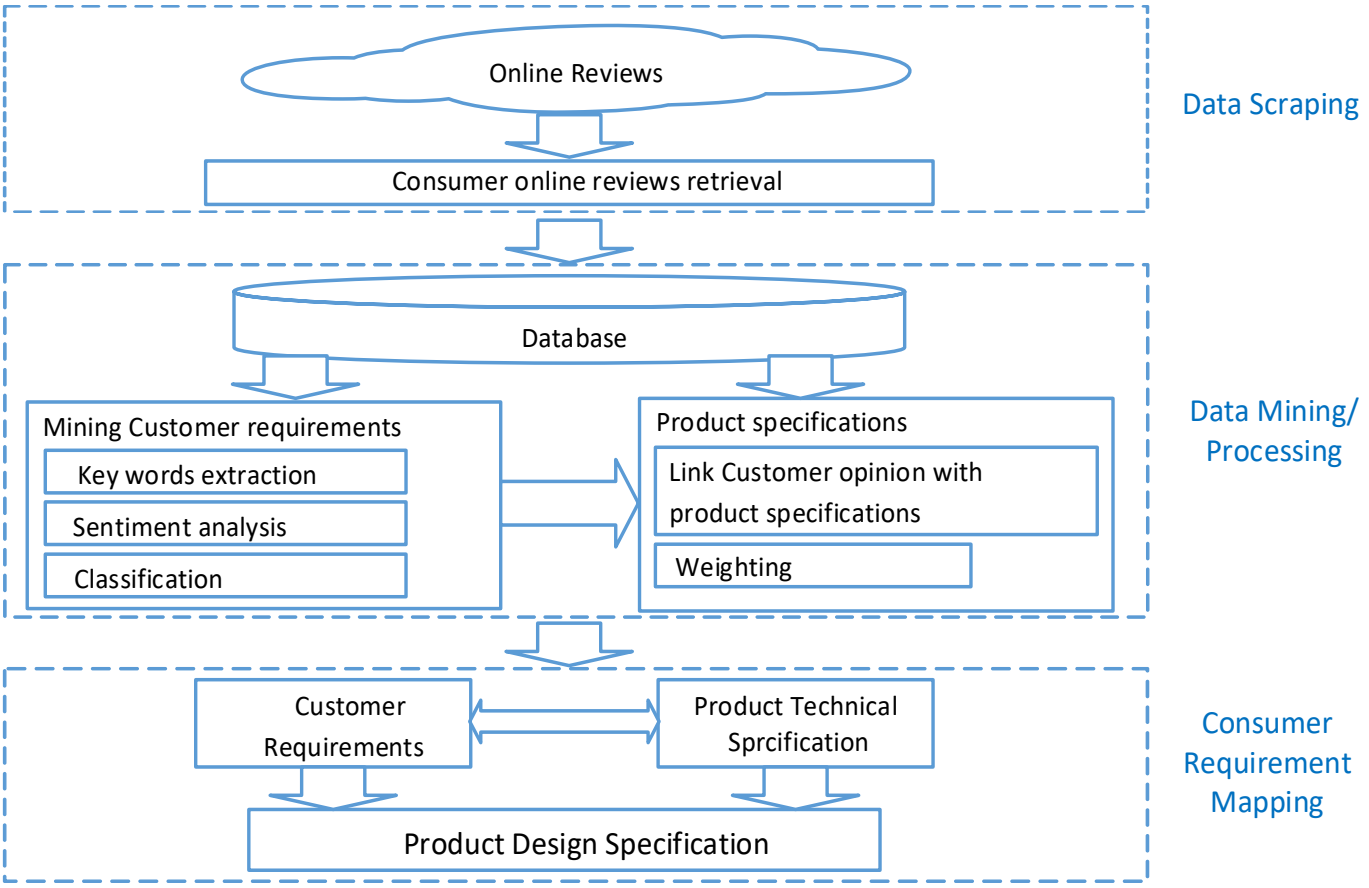


Fig. 1 Overview of online product data mining model

The **Data Scraping** module is to gather a large number of online customer reviews and save the data to the database.

In the **Data Mining / Processing** module, an analysis 'Task' is created to deal with the reviews. The keywords related to reviews are defined in the keyword library before data processing. Then the keywords defined are matched with the scraped reviews using data mining and natural language processing methods, in order to extract useful information in relation to customer requirements and product technical features. Then the classifier function and similarity training function are utilised to optimise the result and display in graphical outputs.

Finally, in the **Consumer Requirement Mapping** module**,** the customer requirements obtained from the above modules are combined with product technical requirements to produce Product Design Specification (PDS), which will help for the manufacturers to improve their products.

In this document, the methods to implement the above three modules are illustrated with relevant graphs, images and description. At the end of this document, the model developed is demonstrated with an example of online data mining on ALIA website, showing how the model works and providing positive supports for the manufacturer to confirm their PDS and further improvement.

The model developed will be utilised at the demonstration stage of this project from Month 19 to Month 33. At the demonstration stage, the online data mining from the ONA online shop will be conducted using this model for domestic lighting products. The ALIA example shown in this document is a prototype, which will be also further developed at the demonstration stage.

According to the data protection regulations, NTU worked with T3.5 team to contact the data owners of Websites, such as Amazon and John Lewis, to obtain permissions to access the review data that are publicly available in their Websites. The above process, as well as relevant work conducted by NTU team, are presented in Section 1 of this document.

Table of Contents

List of Tables

List of Figures

1. Progress

1.1 Permission for data access to e-commerce Websites

In order to develop an online product data mining system, consumers' review data are to be collected via e-commerce Websites. However, in order to do so, due to data protection regulation, permissions from owners of the e-commerce Websites have to be obtained. Below are examples of such permission requirements:

- Amazon *(Amazon, 2018)*
  https://www.amazon.co.uk/gp/help/customer/display.html?nodeId=1040616

  > COPYRIGHT, AUTHORS' RIGHTS AND DATABASE RIGHTS
  >
  > All content included in or made available through any Amazon Service, such as text, graphics, logos, button icons, images, audio clips, digital downloads and data compilations is the property of Amazon or its content suppliers and is protected by Luxembourg and international copyright, authors' rights and database right laws. The compilation of all content included in or made available through any Amazon Service is the exclusive property of Amazon and is protected by Luxembourg and international copyright and database right laws.
  >
  > You may not extract and/or re-utilise parts of the content of any Amazon Service without our express written consent. **In particular, you may not utilise any data mining, robots, or similar data gathering and extraction tools to extract (whether once or many times) for re-utilisation any substantial parts of the content of any Amazon Service, without our express written consent.** You may also not create and/or publish your own database that features substantial parts of any Amazon Service (e.g. our prices and product listings) without our express written consent.

- John Lewis (*John Lewis, 2018)*
  https://www.johnlewis.com/customer-services/shopping-with-us/terms-and-conditions

  > You acknowledge and agree that the material and content contained within the Website is made available for your personal non-commercial use only and that you may (if necessary to make a Purchase) download such material and content onto only one computer hard drive for such purpose. Any other use of the material and content of the Website is strictly prohibited. You agree not to (and agree not to assist or facilitate any third party to) copy, reproduce, transmit, publish, display, distribute, commercially exploit or create derivative works of such material and content.

  Based on the above reviewed results, the review data which are part of contents included within (made through) the Websites are not allowed to be used without the permission of the data owners.

To follow the data protection regulations, NTU team has worked with T3.5 team, making considerable efforts to contact the e-commerce Website owners in order to obtain the permissions, for example:

- NTU contacted Amazon via submitting a data-access request through the online service system (www.services.amazon.co.uk), and also made several telephone

conversions with Amazon Customer Service in UK to confirm if their data can be used. Requests in writing, similar to the one shown in **Appendix I** "Request of the access to Amazon data", were sent to several contacts in Amazon. However, no explicit confirmation has been received so far.

- NTU made several telephone conversations, as well as e-mail communications (see **Appendix 2** "Request of the access to John Lewis data (I)" and **Appendix 3** "Request of the access to John Lewis data (II)"), with several contact points of John Lewis, such as service centres and headquarter, requesting for permission of data access. However, so far, no explicit confirmation has been received from John Lewis
- Apart from NTU team, other CIRC4Life consortium partners of Task 3.5 also made efforts to contact relevant e-commerce Website owners, but all failed.

To overcome the above problem, it has been decided to use the e-commerce Websites of ALIA and ONA, who are the industrial partners of the CIRC4Life project. The progress regarding this is as follows:

- Permissions from ALIA and ONA have been obtained as shown in **Appendix 4** "Permission of the use of ALIA website review data" and **Appendix 5** "Permission of the use of ONA website review data".
- A case study is conducted using ALIA Website to illustrate how the model developed works, which is presented in Section 4.
- It has been planned that the data mining on ONA Website will be conducted at a late stage of the project, during the implementation of its demonstration.
- The amount of consumer review data in ALIA and ONA Websites will continuously increase, and, hence, at the late stage of the project when we conduct the demonstrations, there will be sufficient amount of data to produce more accurate results.

1.2 Development of online product data mining model

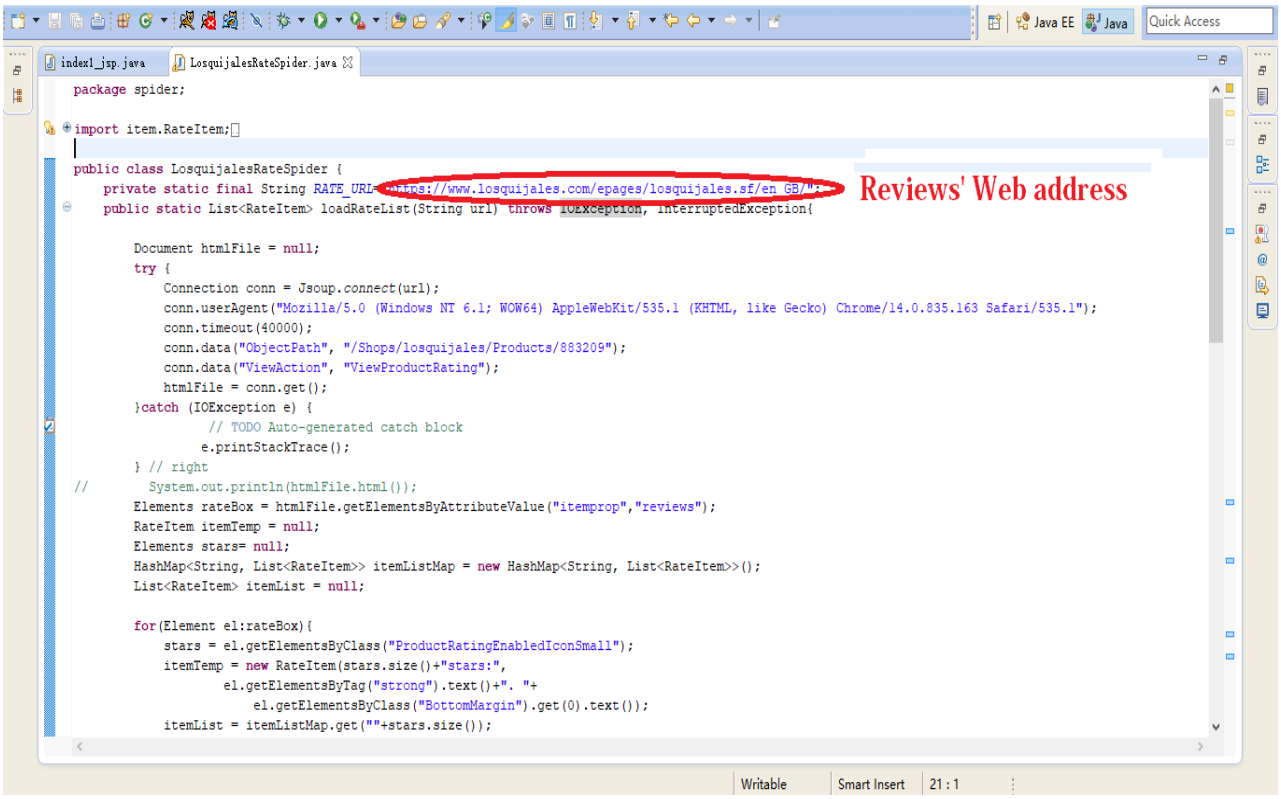
NTU has developed a model for online mining consumer opinions though e-commerce Websites, which aims to achieve the following functions:

- Online crawling of consumer reviews via commercial websites.
- Mining the consumer's opinions, with major concerns of products' sustainable features.
- Consumer's opinions are mapped to the Product Design Specification (PDS).
- A case study conducted using ALIA e-commerce Website proves that the model developed by this research provides positive support for the manufacture to assess and improve their product design specification.

2 Methods utilised in the development of the model

2.1 Data Scraping

Data scraping is applied to collect the consumer's reviews from the webpage and save them in a file in the location specified. As show in Fig. 2, when the target customer reviews' Web address was entered, the consumer reviews can be obtained immediately. Fig. 3 shows the data scraped from the meat product pages.




```
package spider;

import item.RateItem;

public class LosquijalesRateSpider {
    private static final String RATE_URL="https://www.losquijales.com/epages/losquijales.sf/en_GB/";
    public static List<RateItem> loadRateList(String url) throws IOException, InterruptedException{

        Document htmlFile = null;
        try {
            Connection conn = Jsoup.connect(url);
            conn.userAgent("Mozilla/5.0 (Windows NT 6.1; WOW64) AppleWebKit/535.1 (KHTML, like Gecko) Chrome/14.0.835.163 Safari/535.1");
            conn.timeout(40000);
            conn.data("ObjectPath", "/Shops/losquijales/Products/883209");
            conn.data("ViewAction", "ViewProductRating");
            htmlFile = conn.get();
        }catch (IOException e) {
                // TODO Auto-generated catch block
                e.printStackTrace();
        } // right
//        System.out.println(htmlFile.html());
        Elements rateBox = htmlFile.getElementsByAttributeValue("itemprop","reviews");
        RateItem itemTemp = null;
        Elements stars= null;
        HashMap<String, List<RateItem>> itemListMap = new HashMap<String, List<RateItem>>();
        List<RateItem> itemList = null;

        for(Element el:rateBox){
            stars = el.getElementsByClass("ProductRatingEnabledIconSmall");
            itemTemp = new RateItem(stars.size()+"stars:",
                    el.getElementsByTag("strong").text()+". "+
                        el.getElementsByClass("BottomMargin").get(0).text());
            itemList = itemListMap.get(""+stars.size());
```

Fig. 2 Data Scraping

```
2stars:  sodium nitrite contains. Very taste but unfortunately contains sodium nitrate Nitrite can damage cells and also morph into molecules that cause cancer
2stars:  way too fatty. I understand where people are coming from about the saltiness...but they are mortadella after all...so I'd thought I would still try them. being a coeliac
ot feel well. I tried, as I said, three times but no.
3stars:  Contains Sodium Nitrite. Sodium nitrite is linked with increased cancer risk. Delicious but not that good!
3stars:  highly spiced sausage. I personally found it too highly spiced - tasted to us more like a black version of a Haggis, rather than a black pudding.
3stars:  This is a really nice sausage but it didn't last that long in my fridge.. This is a really nice sausage but it didn't last that long in my fridge.
4stars:  Great service. I ordered some products and the delivery service was so fast. I find the reusable bags system a really good idea.
4stars:  Good mortadella. I think that this mortadella has a good flavour, but I would like it to be less greasy and to contain less chemical additives in order to make it healthi
e. I like this product, but I think it could be better to use more natural ingredients and not to use chemical preservatives. In my opinion, this would add value to the product.4st
I would suggest that people should probably get them if they not care about organic and free-range.
4stars:  excellent. Could not bring myself to buy proper mortadella. Theses were a brilliant organic alternative. Liked by children and adults
4stars:  Absolutely delicious!. We got this in error and liked it so much better. This product is a staple in our household now.
4stars:  Delicious moist garlicky Italian meat. Mmm! Really tasty cured meat, quite garlicky, perfect for antipasti and sandwiches.
4stars:  AMAZING. This is one of the best products I've tried, very flavoursome. this was a pleasant surprise and will buy again.
4stars:  Decent enough and good value for money. I'm surprised at the rave reviews as this is far from the best mortadella. It's nice enough though - good quality, decent fla
same flavour but without the excessive saltiness. Presumably the high levels of salt are used to preserve the products shelf life.
4stars:  Delivery cost. I bought this sausage every year for Christmas breakfast. The only downside is the delivery costs which make it more expensive than others which are s
about the kidness of the employee. He adapted to my availability since I was at work until the evening.
5stars:  Amazing. The best mortadella I've ever tasted. I recommend it to every one
5stars:  Really good!. Cheap, tasty and healthy. What else could we ask for?
5stars:  Love traditional products. I love traditional and local products. I think it is a very important thing to use local ingredients to prepare this mortadella. I suggest to take
5stars:  Good one. The delivery service is so eco-friendly as they reuse the plastic bags, I really like it.
5stars:  Perfect for a picnic. When I go to a picnic I always get one of this. It is so tasty and if you eat with a fresh tomato from the Region of Murcia you would enjoy it for s
to fry it a bit and to prepare with onion and pepper. That's delicious for me!
5stars:  One of the best products. I am a regular consumer of Los Quijales products and one that I like the most is this mortadella. I would like it to have a more eco-friendly
la from other brand, but once I tried it I realized that I was gonna continue buying this one. Excellent flavour and less greasy than the others.
5stars:  Great nutritional properties. It has excellent nutritional properties, I think it is a healthy mortadella.
5stars:  Excellent quality. The quality is amazing, I would recommend this mortadella to everyone.
5stars:  Great delivery service. The delivery service was great. When the employee arrived home I could checked that the van was a low emission vehicle, which I find really in
e waste. Great idea!
5stars:  Great traditional mortadella!. I love this mortadella because of its traditional recipe, with different products from the local area so it is better for the enviroment.5star
```

Fig. 3 Consumer review data shown in a special file

2.2 Data Mining / Processing

This module is developed to mine customer requirements, based on the Keywords method, by identifying product attributes via online reviews. To achieve this, the following analysis procedure is conducted, as shown in Fig. 4.

1) Frequently used words are identified as candidate attributes.
2) Meaningless words are pruned out of the frequently used words.
3) A list of words which related to their attributes is constructed.
4) The customer ratings for the attributes are then elicited/generated.

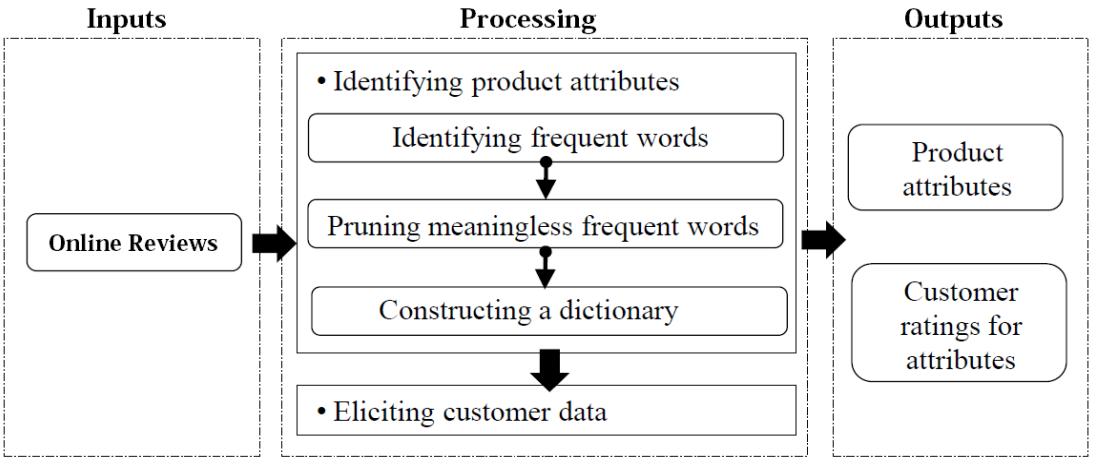


Fig. 4 Work Flow of Data Analysis

Fig. 5 shows the user interface, where are two options: the user could choose ‘Task list’ which lists exiting tasks, and ‘Create Task’ which allows the user to create new task. When choosing ‘Task list’ the user can view the existing tasks in the system, such as a task for meat product analysis; when choosing ‘Create Task’, the user is allowed to create a task to start a new analysis.

Fig. 6 shows the user interface of creating a new Task, where the user enters the necessary information, such as task title, target, etc. This task is used to record the consumer reviews which are scraped from the Website and analysis results (e.g. classification graphs of consumer opinions).

Fig. 5 Interface

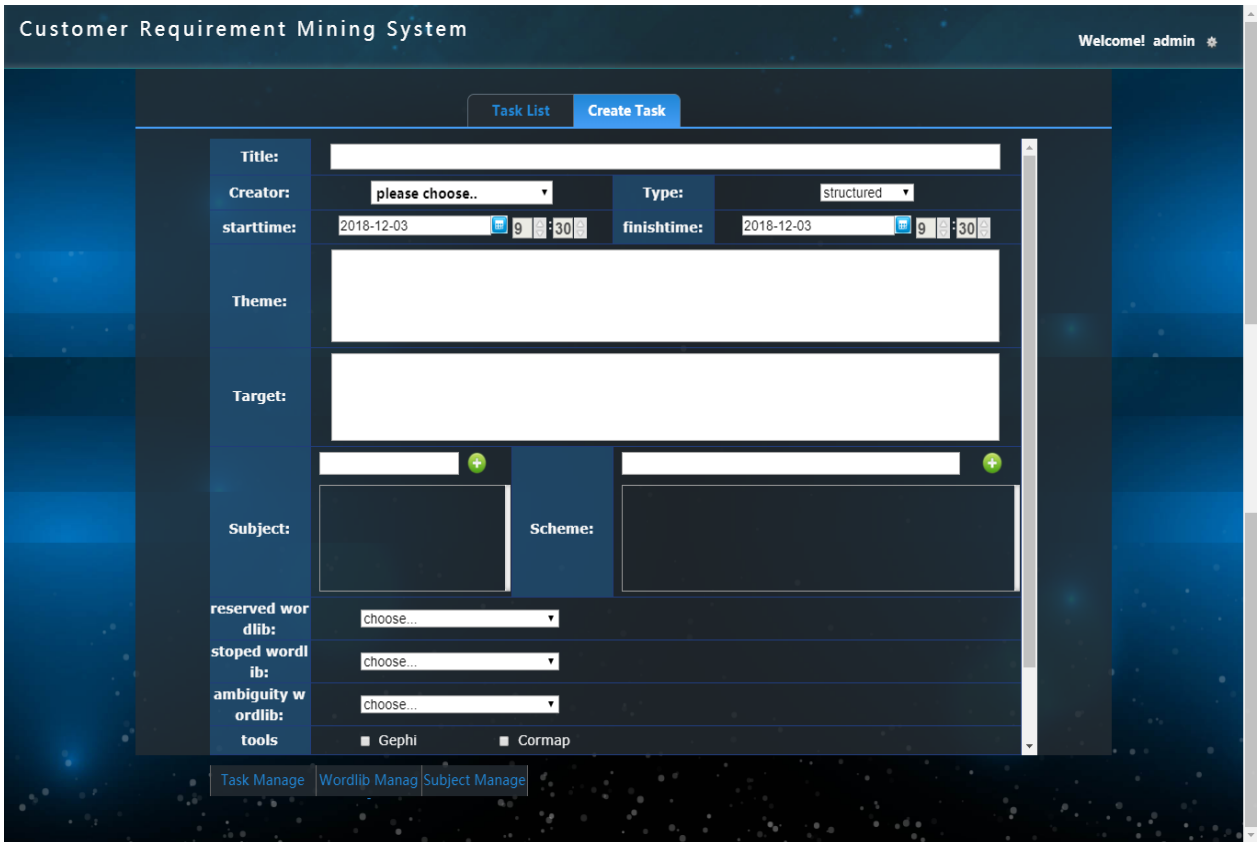


Fig. 6 Create new Task

A self-learning algorithm based on Keywords has been developed, which is used to analyse the correlation among the reviews and find out the internal patterns to achieve the classification of the consumer's opinions. The keywords defined are matched with the

scraped reviews to extract useful information in relation to customer requirements and product technical features. Then the classifier function and similarity training function are utilised to optimise the result and display in graphical outputs. Considering this task is intended to develop a method for online review mining, several review data have been used for training and validating the method developed.

Fig 7 shows the analysis results user-interface, which consists of the following three parts:

1) The top part of the image includes the consumer reviews on a meat product, which were scraped form ALIA's website.
2) The bottom-left part of the image shows the classification of consumer views. The system classifies the reviews into several categories marked with different colours (e.g. green, red, etc.) based on the Keywords of each review, and further, obtain consumers' opinions via identifying the internal relationship between those Keywords with a Network Analysis method. For more information please see Fig. 8.
3) The bottom-right part is a Cormap-based classification method, which reflects the consumer preferences according to 'Stars' rating scores given by consumers. For more information please see Fig 9 and Fig 10.

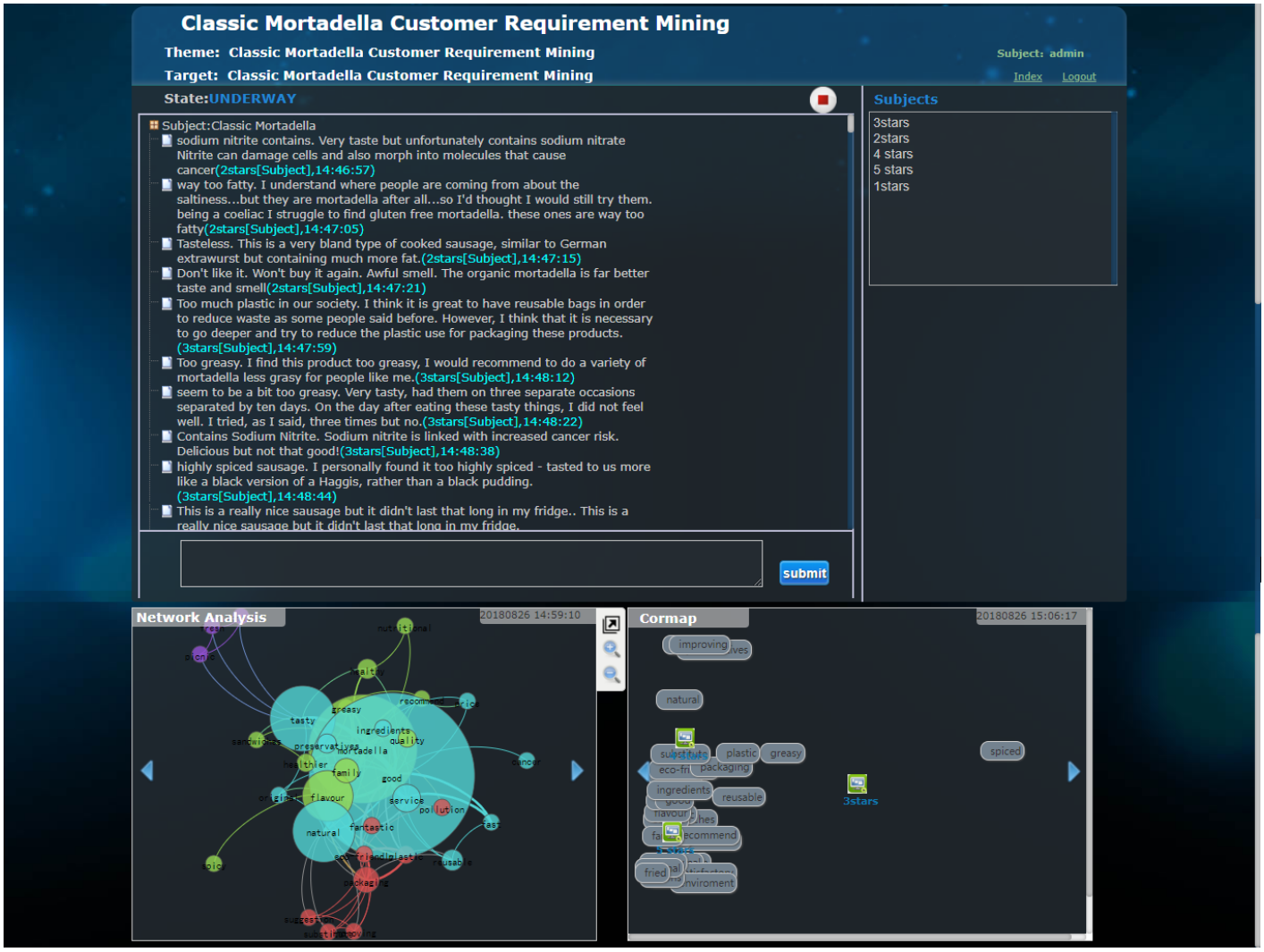


Fig. 7 Data processing results

Fig. 8 shows the online reviews analysis results based on Keywords. Each circle is related to a Keyword; the size of the circle indicates that the frequency (number of times) of the keyword used in the reviews. For instance, the keyword frequently appeared in the reviews is shown a large circle, such as ‘good’, ‘tasty’ and ‘natural’. The lines linked between the keywords indicate that the keywords are related, such as ‘packaging’ and ‘eco-friendly’ in red colour.

The system can track the analysis process. By clicking the arrow button in the left and right side, it displays the analysis results in the different amounts of reviews, e.g. 50, 100, 200 pieces of reviews.

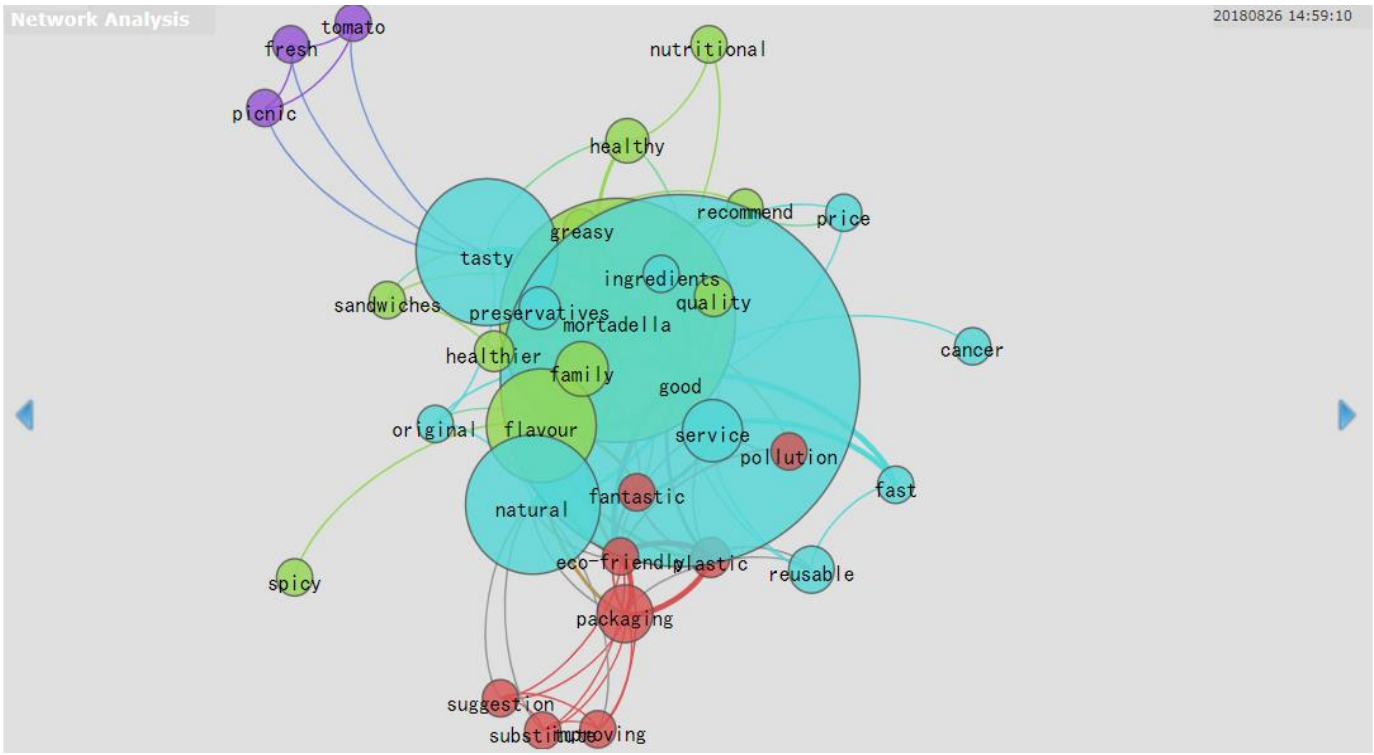


Fig. 8 Keyword-based network analysis results

Fig. 9 and Fig. 10 apply the Cormap method to classify consumer reviews. The Cormap analysis denotes a technology of exploratory analysis of textual data, which shows the group thinking structure from the large amount of reviews (Xijin Tang, 2016). In Fig 9, the reviews have been classified to different groups in terms of ‘Star’ ratings. When hovering a mouse pointer on a keyword (see Fig.10), the reviews related to this keyword can be shown in a pop-up user interface.

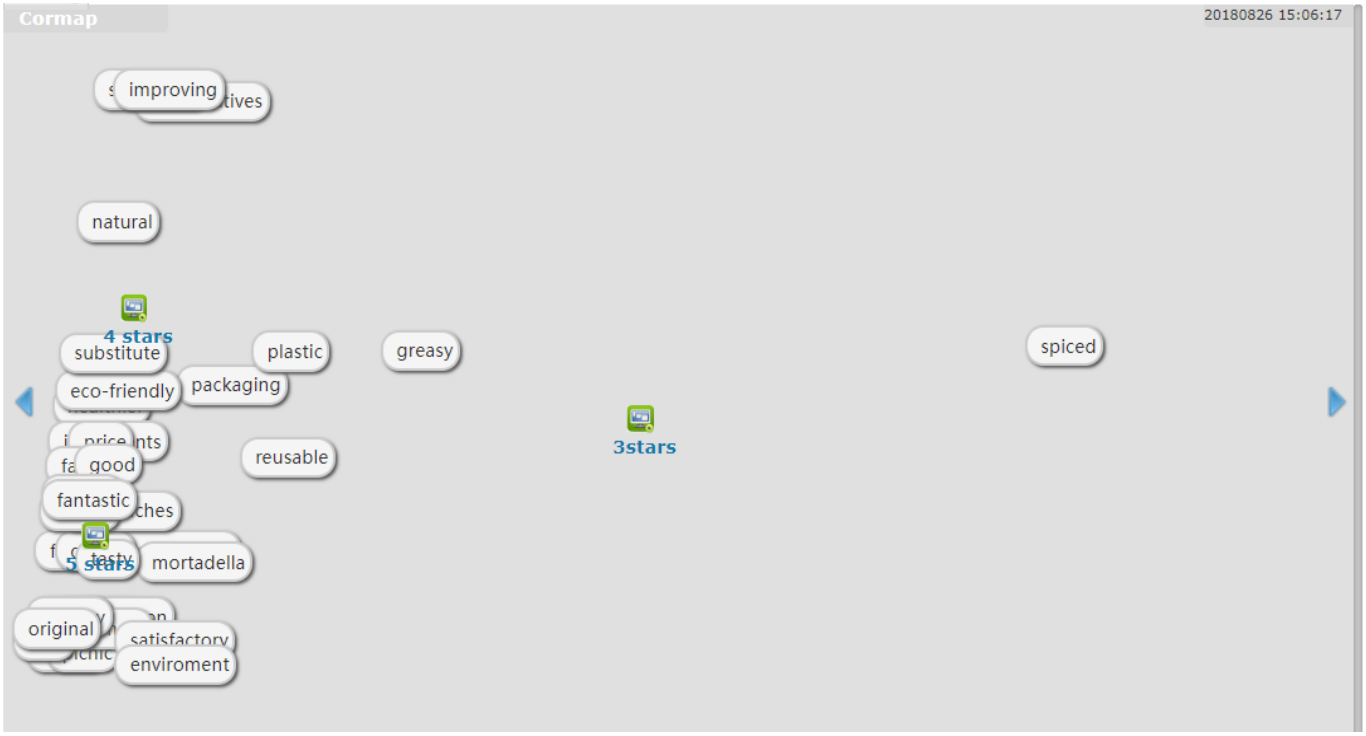


Fig. 9 Cormap calculation results 1

Fig. 10 Cormap calculation results 2

3. Utilisation of consumer reviews in PDS development

To capture the consumers' preference from their reviews on a particular product, a set of basic keywords are established first such as those shown in Fig.8 and Fig. 9, then the basic keywords are grouped under several top-level keywords. The top-level keywords are directly linked to the PDS items. The following example of a consumer review illustrates the basic and top-level keywords:

> 'I bought a *box* of the sausages for a party. It is a bit *salty* for me, but most other people attending the party like the *flavour*. The *cardboard* *box* used for the delivery is good because it is recyclable, but the *plastic* *bag*s inside the *box* could be replaced with recyclable *materials* such as *paper*.'

In the above consumer review, 'salty' and 'flavour' are basic keywords, which are grouped into the top-level keyword 'taste'; 'box', 'bag', 'cardboard', 'plastic', 'paper' and 'material' are basic keywords under the top-level keyword 'packaging'. With this model, the above consumer review was captured by the basic keywords and the information is then used to develop the PDS related to top-level keywords 'taste' and 'packaging'.

In the process to utilise the analysis results of the consumer reviews in the development of PDS, the following are considered:

- This task has to work closely with the WP1 tasks to ensure the successful development of the PDS model.
- The designers and producers involved in the value chain will be enabled to access the data/results from the repository related to PDS, in order to develop the eco-product reflecting consumer needs.
- The PDS repository should have the capability of automated update to ensure the access to the collection of the latest consumer views.

4. Example: PDS development incorporating with consumer requirements

This section presents an example how to utilise the consumer review online mining results in development of the PDS, which analyses the customer's reviews of an ALIA's meat product and then map their requirements to PDS. The customer reviews have been analysed from ALIA's website (www.losquijales.com) about the chosen meat product (CLASSIC MORTADELLA, a classic pork sausage).

4.1 The review results and mapping

Those reviews have been categorized into positive reviews, negative reviews, and both. In some customer reviews, the customer gave both positive and negative opinions in one review. The percentages of three types of review results are shown in Fig.11:

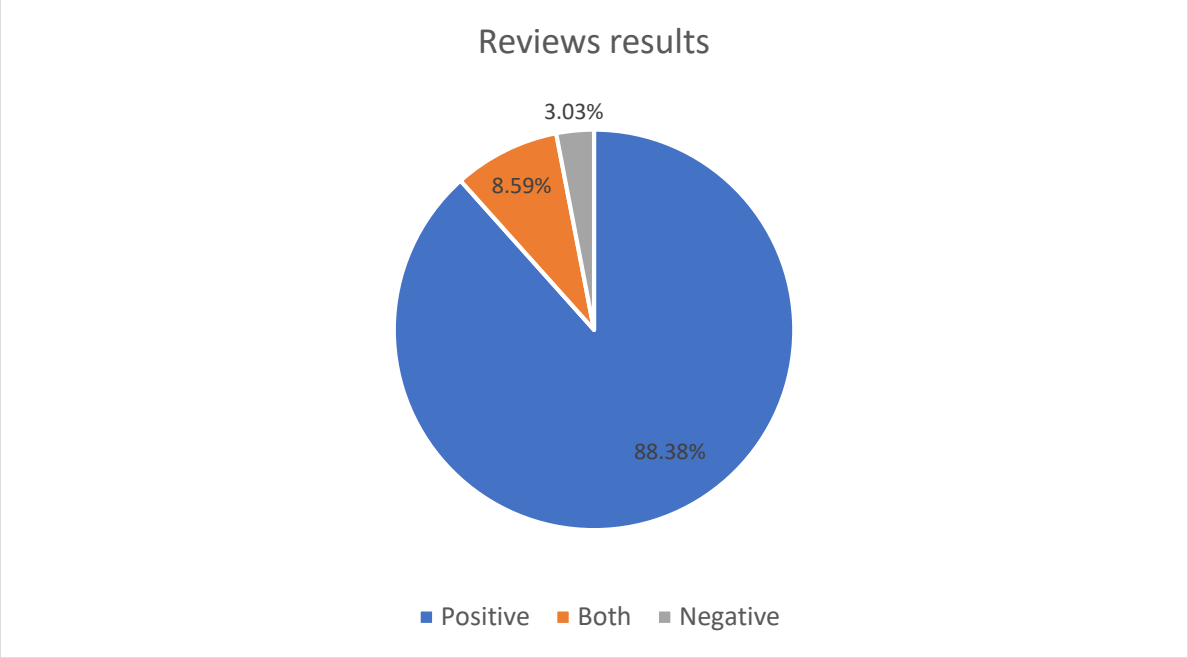


Fig. 11 Percentages of three types of review results

During the process of collecting and analysing the customer reviews, eight keywords are used, including price, taste, delivery, packaging, gluten free, greasy and chemical additives. The selection of the keywords is based on their frequencies of appearance in the consumer reviews and relevance to the PDS. The reviews related to the keywords are selected, which are further analysed and their distribution against the keywords is shown in Fig.12.

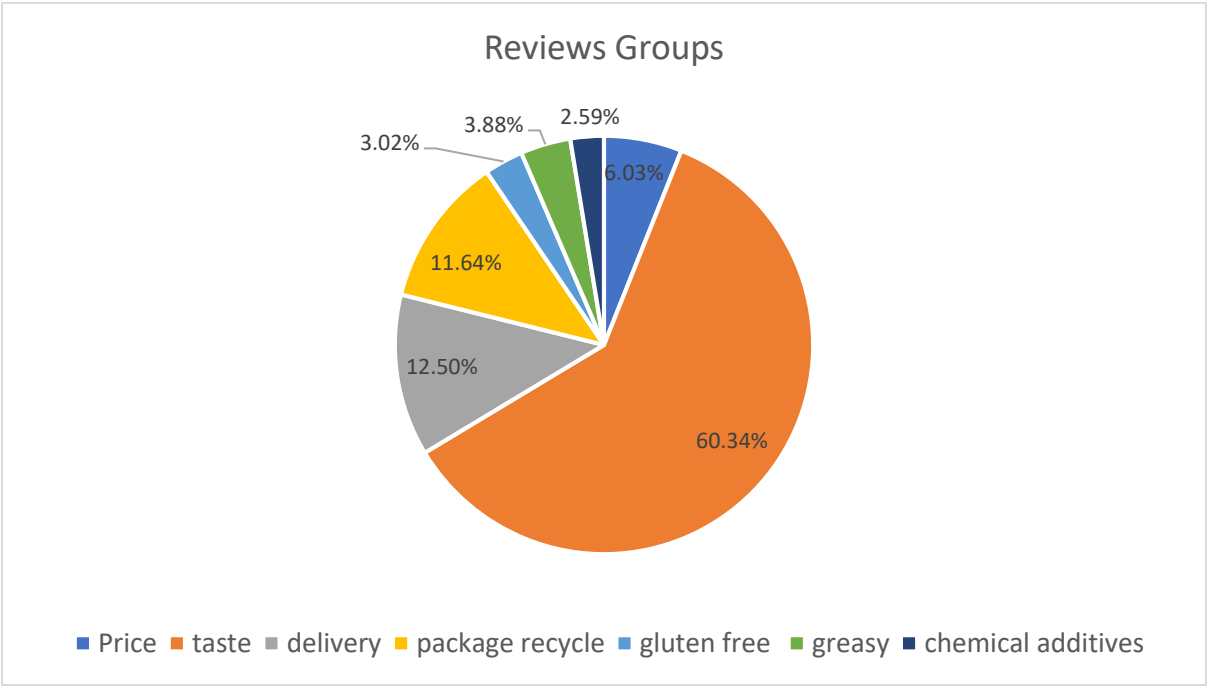


Fig. 12 Distribution of key review groups

The results shown in Fig. 12 include both positive and negative views, while Tables 1 and 2 present analysis results in two groups: positive and negative reviews.

Table 1 Positive review analysis results

Positive Groups	Times Mentioned (%)
Price	6.03%
Taste	59.05%
Delivery	12.50%
Delivery Package (Eco)	9.05%
Gluten free	3.02%

Table 2 Negative review analysis results

Negative Groups	Times Mentioned (%)
Salty	1.29%
Greasy	3.88%
Chemical Additives	2.59%
Product Package (Eco)	2.59%

The above results reveal the following:

1) 6.03% consumer reviews mentioned about the product price, and they were satisfied with the price.
2) 60.34% customer reviews mentioned about the taste of the current product, 59.05% customer were satisfied, but 1.29% customer thought it was too salty.
3) 12.5% customer reviews mentioned about the product delivery and they were satisfied
4) 9.05% customer reviews mentioned about the delivery packaging, and they were satisfied, and there were comments that the delivery package used recyclable materials, such as cardboard boxes, which is recyclable and hence is good for the environment.
5) 3.02% customer reviews mentioned the product is gluten free which they were satisfied with that.
6) There were 3.88% customer reviews commented that the product was too greasy, and 2.59% customer would like to use some natural ingredients to substitute the chemical additives in the current product.
7) There were 2.59% customer reviews mentioned that if change the product package material (plastic bags) to a recyclable material could give more benefit to the environment.

The above analysis results are then utilized to establish the PDS of the meat product. The meat product, i.e., the sausage, for which the online review data mining was conducted, is an existing product, and with the review analysis results shown above, the improved PDS list shown in Section 4 is derived which will help the manufacture to improve the product. Table 3 shows the relationship of the consumer review results with the PDS items.

Table 3 Mapping the consumer review with the PDS

Analysis results of the consumer reviews		**PDS items with the mapping**	
No.	Analysis results	No.	Relevant elements in the PDS
1	The price of the product is appropriate.	5	The price of the final meat product must be competitive in the market
2	Taste of the product meet the majority of the consumer preference.	14	To meet the consumer preference and demands

3	Satisfied with the delivery of the product	13	Delivery
4	Satisfied with the packaging. Recyclable materials are used in the delivery of the product.	10 13	Materials and reduce waste Packaging and delivery
5	The product with gluten free option meet the consumer needs.	14	To meet the consumer preference and demands
6	The product was too greasy. To use natural ingredients to substitute the chemical additives in the current product.	12	Food safety and health
7	change the product package material (plastic bags) to a recyclable material could give more benefit to the environment	10 13	Materials and reduce waste Packaging

As shown in Table 3, the review results cover part, not all, of the PDS items. This is because the other PDS items are more technical ones which consumers may unable to comment, and, also, the number of consumer reviews needs to be increased so that more accurate analysis results could be obtained.

4.2 Product Design Specification for meat product

This is an example of PDS for meat product, which will be provided to meat manufacturer for considerations in improving/developing their products.

1) Performance
The different stages of growth and fattening of the animals must be defined by the following periods of time:

i. Birth to end of weaning of piglets –21 days.
ii. End of weaning to fattening transition –42 days.
iii. Fattening –118 days

The weight of the pig at the end of the three must be between 110-115 kg.

2) Animal welfare
An occupation of animals in each cage allowing the free movement of animals must be allowed.
The temperature must be regulated so that it complies with the comfort conditions of the animals.
It should be chosen the specific breeding species in the farm according to the climatic conditions of different periods of the year and animal needs.

3) Prevention of pests and diseases
Preventive treatments by the use of natural products should be prioritized instead the use of antibiotics. This includes stocking rates, housing conditions, diet and probiotics.

4) Life in service (performance)
The animals must go to the slaughterhouse in a period of time between 5 and 6 months, and a weight of 110-115 kg.

5) Target Product Cost
The farm must not spend more than a certain amount of money 1, 05 €/kg of pig. The price of the final meat product must be competitive in the market.

6) Shipping
The transport must be carried out in an adapted truck for animal transport and must be as sustainable as possible in terms of CO2 emissions.

7) Quantity
A total of 25 batches will be produced (equivalent to 25 pigs).

8) Manufacturing Facility

- The energy sources should be as low carbon as possible and energy efficient.
- The tools used in the process must be reusable (silos, cages, water troughs)

9) Size and weight
The objective is that the animals go to slaughterhouse with an age 6 months and a weight 110-115 kg.

10) Materials

- The water used should be reused or recycled as much as possible.
- The animal feed must have sustainable characteristics (using other industries' by-products as ingredients, using formulations that reduce nitrogen and phosphorous content in animals' manure, including probiotics that reduce antibiotics need, favoring the selection ingredients with low environmental impact).
- The use of antibiotics should be minimized.
- Minimization of organic waste through systems such as hydrolysis of corpses, minimization of manures through diet, and management through agriculture use.
- Minimization of inorganic waste through use of bulk raw materials instead of packaged ones as much as possible (with silos storages).
- Increase in recycling rates and in the use of slurry for further uses in agriculture or energy recovery.
- Materials not related to the main activity (as office material) must be acquired with sustainability criteria.

11) Product life span
Breeding must be completed in 5-6 months.
Product shelf life time must be more than 6 months.

12) Food safety and health

- The Labeling information on product packaging should include the origin of the meat and record the meat processing history.
- Minimization of chemical additives used in the product.
- Increase the organic ingredients usage of the product.

13) Packaging and Delivery

a) Increase use of recyclable materials for packaging.
b) Include environmental criteria in the planning of delivery routes in order to reduce carbon emissions.

14) Customer
The elaboration of the products will be based on the preferences and criteria of clients

regarding environmental features (packaging, ingredients origin, etc.) through data analysis, and to meet the consumer preference and demands.

4.3 Conclusions

Online review mining of ALIA product was conducted to illustrate the development of PDS that addresses consumer requirements. The online review results provided the supports for meat manufacture to identify their PDS and to further improve their products.

This section incorporated the consumer requirements into the technical information provided by ALIA product to develop the product specification, which mainly includes the following features:

- Food safety and health will be improved (see item 12 of the PDS), for example, via including the source of meat on the product label, reducing chemical additives, and organic ingredient usage, based on the 2.59% of the consumers' comments on food health.
- Packaging will be improved with the use of recyclable materials (see item 13.a of the PDS), based on the 11.64% of the consumers' comments on eco-packaging.
- Environmental criteria will be included in the planning of delivery routes in order to reduce carbon emissions (see item 13.b of the PDS), based on the 12.5% of the consumers' comments given on the delivery.
- The price of product is satisfied by the consumers who commented on the price, which confirm that originally planned Target Product Cost is acceptable (see item 5 of the PDS).
- The taste of product is satisfied by 97.86% of the consumers who commented on the taste, which reflect that the materials used (e.g. ingredients) are reasonable (see item 10 of the PDS).

5. Concluding remarks

This deliverable developed a model to mine consumer's review data from e-commerce websites, and further analyse the consumer's requirements on product features. The model consists of three modules: Data scraping, Data mining/processing, and Consumer requirement mapping, which is illustrated with an example of online data mining on ALIA website, showing how the model works and providing positive supports for the manufacturer to confirm and further improve their PDS.

This model will be utilised at the demonstration stage of this project, from Month 19 to Month 33. At the demonstration stage, the online data mining will be conducted via e-commerce Websites through considerable online review data of products, such as domestic lighting products from the ONA online shop and meat products from ALIA e-store.

The ALIA example shown in Section 4 is a prototype, which will be also further developed at the demonstration stage, and with increased number of consumer reviews more accurate analysis results will be obtained.

References

Amazon (2018), Conditions of Use & Sale, 31 October, 2018. https://www.amazon.co.uk/gp/help/customer/display.html?nodeId=1040616

European Commission (2017), 'H2020 Programme Proposal template 2016-2017', The EU Framework Programme for Research and Innovation, Horizon 2020, Version 3.1, 11 January 2017

John Lewis (2018), General Terms, TERMS AND CONDITIONS, 2018, 31 October 2018 https://www.johnlewis.com/customer-services/shopping-with-us/terms-and-conditions

Xijin Tang (2016), 'Problem Structuring Process by Qualitative Meta-synthesis Technologies', Knowledge Synthesis: Western and Eastern Cultural Perspectives, Springer Tokyo, Heidelberg New York Dordrecht London, 2016, p.p.79-106.

Appendix 1: Request of the access to Amazon data

To whom it may concern at Amazon,

On behalf of the Advanced Design and Manufacturing Engineering Centre, at the Nottingham Trent University, we contact you with the aim of seeking advice regarding the permission of access to the public data available in the Amazon web site.

Currently, our department is leading the EU H2020 CIRC4Life project (Grant No. 776503) which is focused on the development and implementation of a circular economy approach for sustainable products and services through their value and supply chains. The consortium of the project is completed by 17 European partners who will participate in the execution of the project activities.

Among the project tasks, the CIRC4Life consortium is developing an approach to form product design specifications (PDS) by mining consumer opinions through product's online review data of the e-commerce Web site (e.g. Amazon), in order to largely reflect the consumer's preferences into the product development. To achieve this, it is significant to access the Amazon's public Web-pages for gathering the consumer's online review data. From a technical point of view, the Web pages related to the comments on the specific product (e.g. domestic lighting products, electronic tablets) would be downloaded. Then all the review/comments would be extracted from the pages and finally passed to the software tool for data analysis (e.g. categorisation, semantics). Following this procedure, we would be able to complete the PDS reflecting consumers' preference.

After this short explanation of our targets and as it has been mentioned at the beginning, the aim of this letter is to request your authorisation for the use of the review public data available in Amazon Webpages for the purpose stated above. For that, the CIRC4Life Consortium explicitly confirms that:

- The CIRC4Life Consortium will use only the product's review/comment data published by the consumers via the Amazon Website.
- The CIRC4Life Consortium won't use the consumer's personal data, such as their user name and password.
- The CIRC4Life has an ethical board that oversees all activities related to the use of data and how to protect relevant information for agents/stakeholders.
- All data will be employed in the CIRC4Life framework and under no circumstances for commercial use.
- The CIRC4Life Consortium will show always the source/origin of data in all the project publications (reports, conference/journal papers, e-news, press release, etc). I believe that this would be a good publicity for Amazon.

In the event that Amazon considers essential the inclusion of additional agreements, please let us know.

Finally, if you need further information about the CIRC4Life Project, please click on the link below

CIRC4Life	"CIRC4Lfife, A circular economy approach for lifecyles of products & services" https://cordis.europa.eu/project/rcn/214414_en.html

Looking forward to your reply. Thanks in advance.

Kind regards,

Wenjie Peng

PhD, MSc,
Research Fellow in Product Design
Advanced Design and Manufacturing Engineering Centre (ADMEC),
School of Architecture, Design and the Built Environment,
Nottingham Trent University

Appendix 2: Request of the access to John Lewis data (I)

School of Architecture, Design and the Built Environment
50 Shakespeare Street, Nottingham NG1 4FQ. Tel. +44 (0)115 941 8418 www.ntu.ac.uk/adbe

Professor Daizhong Su
Head, Advanced Design & Manufacturing Engineering Centre
Room 206, Maudslay Building
Burton Street, Nottingham, NG1 4BU, UK
Tel: +44 (0)115 848 2306, E-mail: daizhong.su@ntu.ac.uk,

20th July 2018

To whom it may concern at John Lewis

I am a regular John Lewis customer with a My John Lewis card and a John Lewis Partnership card, and also a Professor at Nottingham Trent University. I am writing to seek John Lewis' support for our research.

I am currently leading the following big project supported by the European Commission, with 7.3 million Euros budget and 17 partners across 8 EU countries.

CIRC4Life	"CIRC4Lfife - A circular economy approach for lifecycles of products and services", EU H2020 grant agreement No. 776503 https://cordis.europa.eu/project/rcn/214414_en.html

As a task of the project, we are developing an approach to form product design specifications (PDS) by mining consumer opinions through product's online review data of e-commerce Websites, in order to largely reflect consumers' preference in product design. We would like to use John Lewis' Website for this purpose, because not only it suits for this research, but also I used it a lot and like it very much. Although the consumer reviews on the Website are open to public, I am writing to ask for John Lewis' kind permission to use the review data, according to the EU General Data Protection Regulation.

In the research, the consumer comments on specific products (e.g. lighting products, electronic tablets) on the Web pages would be downloaded and passed to a software tool for data analysis. Following this procedure, we would be able to complete the PDS reflecting consumers' preference.

On behalf the CIRC4Life Consortium, I explicitly confirm that:

- The CIRC4Life Consortium will use only the product's review/comment data published by the consumers via the John Lewis Website, but not use any consumer's personal data such as their user names and passwords.
- The CIRC4Life has an ethical board that oversees all activities related to the use of data and how to protect relevant information for agents/stakeholders.
- All data will be used for research purpose and under no circumstances for commercial use.
- The CIRC4Life Consortium will indicate the source/origin of the data and acknowledge John Lewis' support in all publications (reports, conference/journal papers, e-news, press release, etc) where the data will be used. I believe that this would be also a good publicity for John Lewis.

In the event that John Lewis considers essential the inclusion of additional agreements, please let us know.

Looking forward to your reply. Thanks in advance.

Professor Daizhong Su

The Queen's Anniversary Prizes For Higher and Further Education 2015

NOTTINGHAM TRENT UNIVERSITY

Appendix 3: Request of the access to John Lewis data (II)

Peng, Wenjie 02 <wenjie.peng@ntu.ac.uk>

Wed 06/06/2018 12:48

To: technicalsupport@johnlewis.com <technicalsupport@johnlewis.com>;

Cc: Su, Daizhong <daizhong.su@ntu.ac.uk>;

To whom it may concern at John Lewis,

My name is Wenjie Peng, research fellow in sustainable product development, Nottingham Trent University. I am seeking advice from you related to the permission of access to the data which are publicly available in the John Lewis Website.

I am currently working on the EU H2020 CIRC4Life project (Grant No. 776503), which will develop an approach to form product design specifications (PDS) by mining consumer opinions through product's online review data of the e-commerce Web site (e.g. John Lewis), in order to largely reflect the consumer's preferences into the product development. To do so, it is vital to access the John Lewis's public Web-pages for gathering the consumer's online review data. Technically, the Web pages related to the comments on the specific product (e.g. domestic lighting products) will be downloaded. Then large amount of review/comments will be extracted from the pages, which are subsequently passed to the software tool for data analysis (e.g. categorisation, semantics). This way will enable us to make the PDS best reflect consumers' preference. This is for completing the research project, not for commercial use.

I am writing to apply for your permission for us to use the review data publically available in John Lewis Webpages for the purpose stated above.

We shall acknowledge John Lewis's permission in all our publications related to the use of John Lewis data. Those publications could include reports, conference/journal papers, e-news, press release, etc. I believe that this would also be a good publicity for John Lewis.

The following is the brief information of the CIRC4Life project:
This project aims to develop and implement a circular economy approach for sustainable products and services through their value and supply chains. Three new circular economy business models will be developed including (i) co-creation of products and services, (ii) sustainable consumption, and (iii) collaborative recycling and reuse. For more information about this project, please click on the link below:

https://cordis.europa.eu/project/rcn/214414_en.html

I am looking forward to your reply. Thanks in advance.

Kind regards,

Wenjie

PhD, MSc,
Research Fellow in Product Design
Advanced Design and Manufacturing Engineering Centre (ADMEC),
School of Architecture, Design and the Built Environment,
Nottingham Trent University

Appendix 4: Permission for the use of ALIA website review data

SOCIEDAD AGRARIA DE TRANSFORMACION 2439 (ALIA),
S/N - PARAJE DEL DUENDE DE LORCA
DIPUTACION DE LA HOYA,
MURCIA 30816, Spain

15th June 2018

Dear Professor Daizhong Su,

Thank you for enquiring the permission for access to the ALIA's Website (https://www.losquijales.com). I am writing to confirm that Nottingham Trent University (NTU) can use the product's review/comment data published by the consumers via the ALIA's Webiste. All of data will be used for research purpose of the European Commission Horizon 2020 CIRC4Life project.

Best regards,

(signature)

SOCIEDAD AGRARIA DE TRANSFORMACION 2439 (ALIA)

Research and Innovation
European Commission H2020
EU Proyect ID 776503-2
CIRC4Life

Sociedad Agraria de Transformación Nº2.439, Inscrita en el Registro General de Sociedades Agrarias de Transformación, Folio 1, Hoja 2.439, Folio 29 del Tomo 13, Asiento 1.ª con C.I.F. F-30530753

SAT. Nº2.439
Dip. La Hoya, s/n
Tel. 968 48 67 00 [CENTRALITA]
Fax 968 48 68 28
Apartado 272
30816 LORCA [MURCIA]
http:www.tas.es

Appendix 5: Permission for the use of ONA website review data

Ona Product, SL B96830732 C/San Bartolome, 28 46115 Alfara del Patriarca

18th June 2018

To:
Professor Daizhong Su
Coordinator of the EU H2020 CIRC4Life project
Head of Advanced Design and Manufacturing Engineering Centre
Nottingham Trent University,
50 Shakespeare Street, Nottingham, NG1 4FQ, UK

Dear Professor Su,

Re: Permission for mining consumer views from ONA online shop for CIRC4Life project.

Further to your request, I am writing to give permission for the EU H2020 CIRC4Life project to use the product review/comment data published by consumers on our ONA online shop. This permission is for the research to be conducted by the CIRC4Life project only.

Yours scenically

ona.es
ONA PRODUCT,S.A.
C/ SAN BARTOLOME, 28
ALFARA DEL PATRIARCA

(Signature)

Juan Costa
Director