User-generated content and relevance of sustainability dimensions in the wine market

Miguel-Ángel Gómez-Borja[1], Inmaculada Carrasco[1], and Juan Sebastián Castillo-Valero[1]

[1]Universidad de Castilla-La Mancha, IDR Instituto de Desarrollo Regional, Campus Universitario s/n, E02071, Spain

Abstract. This study explores the intersection of increasing environmental concerns in the wine industry and the influence of user-generated content (UGC) in shaping trends and attitudes towards sustainable practices. Consumers and businesses are highly interested in sustainable, eco-friendly, and organic wines and processes. With the exponential rise of UGC across various internet platforms, there lies an opportunity to analyze different behaviours and trends related to wine markets and consumption. This research aims to collect and scrutinize UGC on social media platforms, explicitly focusing on discussions related to sustainability in the wine industry. Furthermore, the study seeks to complement existing methodologies by analyzing these social conversations, providing valuable insights to wine producers on influential sustainability trends. In essence, this research leverages digital content ecosystems to understand better and predict future trends in wine markets.

1 Introduction

Consumer and business environmental concerns in the wine industry have grown recently. As a result, consumers are increasingly demanding more eco-friendly practices from wineries. Moreover, research has revealed that customers are willing to pay a premium for organic and biodynamic wines due to the perceived benefits of these products, which has contributed to the growth in their availability. Therefore, wineries that demonstrate their commitment to sustainability and environmental responsibility will be in the best position to meet consumer demand [1]. Tait and colleagues show how consumers are willing to pay for the presence of some wine-related sustainability attributes [2]. Nevertheless, literature has recently assessed the discrepancy between attitudes and behaviour regarding organic wine consumption [3].

User-generated content (UGC) has grown exponentially across Internet platforms and applications in recent years. Social networks and other specific Internet platforms and applications have allowed users to have spaces to post content in different formats and for different purposes [4]. Content also ranges from personal updates on social media, to reviews, opinions, and comments. Furthermore, content is materialized in text, ratings, photos, audio, and video. This digital content ecosystem can be an ideal context to explore and analyze the world, generate knowledge, and gain insights on different topics [5].

In this context, behaviours related to food and beverages are increasingly spreading on the main social networks and opinion spaces, so the amount of available information is increasing and of higher quality. In addition, dedicated social applications allow for more valuable information on specific sectors, such as the wine sector [6,7].

On the one hand, general social media analytics consists of collecting large amounts of data from social profiles to discover patterns and knowledge related to a specific topic and sector [5]. On the other hand, more specifically, spaces and particular applications associated with the world of wine (i.e., blogs, wine ratings, buying applications, wine communities, and tourism recommendation websites) can be used to evaluate specific opinions and ratings associated with wine products. Kotonya and colleagues, for example, carried out a detailed multi-market analysis of the data, opinions, and ratings of the Vivino application to detect trends, tastes, and preferences in different countries [6].

Therefore, the objective of this research will be based on the collection and analysis of user-generated content on social networks, such as Twitter or Instagram, news and blogs that are related to sustainability dimensions and attributes that consumers, businesses and other agents naturally talk about in social media and opinion forums. This analysis will allow us to complete the conclusions obtained through different methodologies

and offer a complementary tool to SMEs in the industry to decide which most relevant and activatable dimensions from the winemaker's marketing strategy perspective.

2 User-generated content

UGC, or user-generated content, encompasses all information and media shared online, particularly on social media platforms [4]. Content can include comments, reviews, posts, videos, and other forms of content created by users within digital ecosystems.

The variety of content and users means that UGC can be used to understand various economic and social phenomena and constitutes a comprehensive and high-quality information ecosystem for contrasting theories and understanding economic and business dynamics, especially detecting trends [5].

UGC analysis is particularly relevant today for several reasons. Firstly, we can find content in different formats. Although text is the most frequent format, users increasingly share images, photos, videos, audio, and other virtual formats. The particularities of each format and the integration of all of them allow a comprehensive analysis of many phenomena. Secondly, the concept of "user" is usually identified with a person but can also be associated with an organization or a brand, which implies analyzing the relationships in the market from the point of view of supply and demand. Thirdly, despite biases, they represent a real-time pulse of users' opinions and behaviours. Finally, and fourthly, new machine learning and artificial intelligence tools allow the processing of large volumes of different data formats and obtain relevant insights for decision-making.

In this research, we will consider content from different types of users and social media, although it will be limited to requests for information and analysis of text content.

3 Materials and methods

The analysis stems from a descriptive quantitative examination of messages shared on digital social media, which discuss sustainability within the viticulture sector. The observed information has undergone a content analysis, delineated from thirty categories that were deductively applied to the data (see Table 1). The scope of monitoring has been global. However, since it has been limited to two different languages (Spanish and English), the collected information presents a bias tending to give more relevance to countries where both languages are spoken. On the other hand, the monitoring has been applied from January 2019 to April 2023. Based on the 30 categories, different search queries associated with each term have been generated. In some cases, they are precise (e.g., Greenwashing: Ecolavado OR greenwashing), while in others, they are more general (e.g., Organic Farming / Organic: “vino orgánico” OR “vinos orgánicos” OR “organic wine” OR “organic wines” OR (vino NEAR/2 orgánico) OR (vinos NEAR/2 orgánicos) OR (wine NEAR/2 organic) OR (wines NEAR/2 organic) OR “agricultura orgánica” OR “organic agriculture”). The information was gathered using the BrandWatch platform, available at https://www.brandwatch.com/.

Table 1. Search queries in Spanish and English.

SPANISH. VINO (AND)
SOSTENIBILIDAD
ECOLÓGICO
NATURAL
VEGANO
SULFITOS
CONTAMINACIÓN AMBIENTAL
RESIDUOS, BASURA
HUELLA ANTROPOGÉNICA
HUELLA ECOLÓGICA
HUELLA DE CARBONO
IMPACTO AMBIENTAL
HUELLA HÍDRICA
CONSUMO DE AGUA
GASES DE EFECTO INVERNADERO
AGRICULTURA ORGÁNICA, ORGÁNICO
AGRICULTURA BIODINÁMICA, BIODINÁMICO
EFICIENCIA ENERGÉTICA
ENERGIAS RENOVABLES, ENERGIA SOLAR, ENERGÍA EÓLICA, BIOMASA
RECICLAJE
ECOLAVADO
DECLARACION AMBIENTAL DE PRODUCTO, DAP
PRODUCTOS DE PROXIMIDAD
CANALES CORTOS
ENVASE BIODEGRADABLE, ENVASE RECICLABLE, ENVASE ECOLÓGICO
ECOETIQUETA
IMPLICACION EN LA COMUNIDAD
DIVERSIDAD LABORAL, INCLUSION LABORAL
APOYO A PROYECTOS SOCIALES
OBJETIVOS DE DESARROLLO SOSTENIBLE, ODS
RESPONSABILIDAD SOCIAL, RESPONSABILIDAD SOCIAL CORPORATIVA, RSC
ENGLISH. WINE (AND)
SUSTAINABILITY
ECOLOGICAL, ECOLOGIC, ECO-FRIENDLY
NATURAL
VEGAN
SULPHITES
ENVIRONMENTAL POLUTION
WASTE
ANTHROPOGENIC FOOTPRINT
ECOLOGICAL FOOTPRINT
CARBON FOOTPRINT
ENVIRONMENTAL IMPACT
WATER FOOTPRINT
WATER CONSUMPTION
GREENHOUSE GASES
ORGANIC FARMING, ORGANIC AGRICULTURE, ORGANIC
BIODYNAMIC FARMING, BIODYNAMIC AGRICULTURE, BYODINAMIC
ENERGY EFFICIENCY
RENEWABLE ENERGIES, SOLAR ENERGY, WIND ENERGY, BIOMASS
RECYCLING
GREENWASHING

ENVIRONMENTAL PRODUCT DECLARATION, EPD
LOCALY SOURCE PRODUCTS
SHORT SUPPLY CHAIN
BIODEGRADABLE PACKAGING, RECYCLABLE PACKAGING, ECO-FRIENDLY PACKAGING
ECOLABEL
COMMUNITY INVOLVEMENT
WORKFORCE DIVERSITY, WORKPLACE INCLUSION
SUPPORT FOR SOCIAL PROJECTS
SUSTAINABLE DEVELOPMENT GOALS, SDG
SOCIAL RESPONSIBILITY, CORPORATE SOCIAL RESPONSIBILITY, CSR

Two specialists have overseen the whole process, both the generation of the query (set of keywords and logical operators) and the automatic categorization of the query results.

4 Results

To begin with, we present the total number of mentions (see Fig. 1).

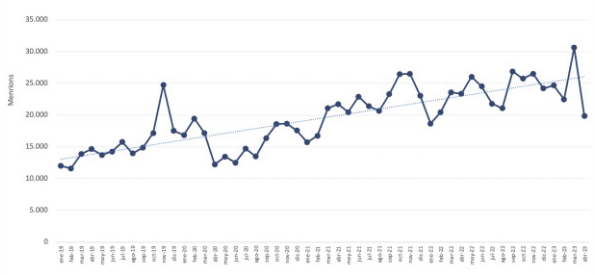

Figure 1. Total number of mentions.

Figure 1 illustrates that there has been a consistent increase in sustainability-related mentions over the past four years in all relevant markets analyzed.

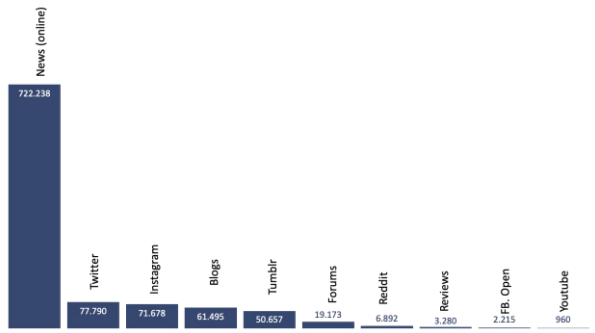

Figure 2. Mentions by media source.

Regarding the source of generated media on social platforms, Fig. 2 indicates that most of it originate from online news websites. However, social networks like Twitter and Instagram, blogs, microblogging sites like Tumblr, forums, and other social media platforms also contribute to the content.

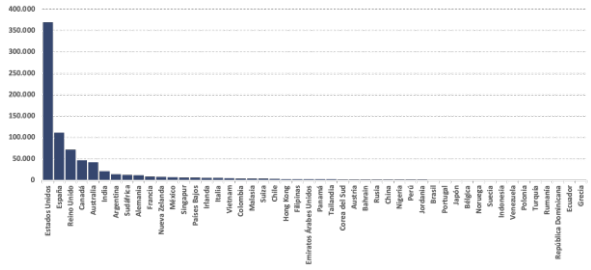

Figure 3. Mentions by country.

When looking at the countries mentioned in the searches conducted in Spanish and English, the highest number of mentions comes from the United States, followed by Spain, the United Kingdom, Canada, and Australia. Figure 3 illustrates how mentions related to sustainability are distributed across various countries. As these five countries hold significant importance in the sector, we will briefly analyse each separately.

Regarding the analysis of the mentions by category, it's worth mentioning that the sustainability category has been excluded from the results due to its broad nature, which resulted in over a million observations. Figure 4 illustrates the classes with the highest percentage of mentions.

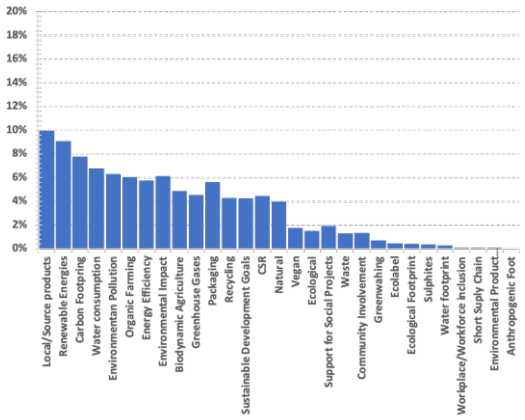

Figure 4. Percentage of mentions for different sustainability dimensions.

The findings indicate that the most discussed topics regarding sustainability are general concepts such as using locally sourced products, reducing carbon footprint, conserving water, combating environmental pollution and reducing greenhouse gases. Energy management, including using renewable energy and improving energy efficiency, is also essential. Discussions also centered around production models and methods, such as biodynamic farming, organic agriculture, packaging, and recycling. While SDGs and CSR are relevant, social sustainability aspects like community involvement, support for social projects, and workforce inclusion are less frequently mentioned. Specific product characteristics, such as being natural, vegan, carrying an

ecolabel, containing sulphites, or having an environmental product declaration, are of lesser importance in these discussions.

By analyzing geolocated mentions, one can explore conversation patterns related to sustainability across different countries and compare them to aggregated data. The figures below illustrate these comparisons for the United States, Spain, the United Kingdom, Canada, and Australia.

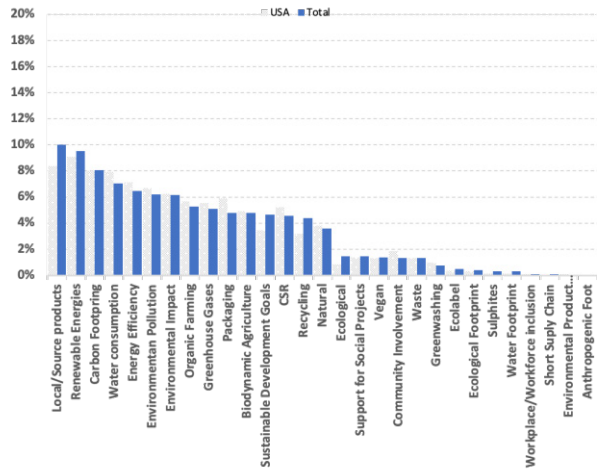

Figure 5. Percentage of mentions for the USA vs total.

In the case of the USA (Fig. 5), the pattern of mentions is more similar to the global pattern due to the weight that mentions from this country represent over the total. However, the lesser importance of concern for local/source products, renewable energies, SDGs or recycling categories could be highlighted. On the other hand, there are more mentions of CSR, packaging or water consumption.

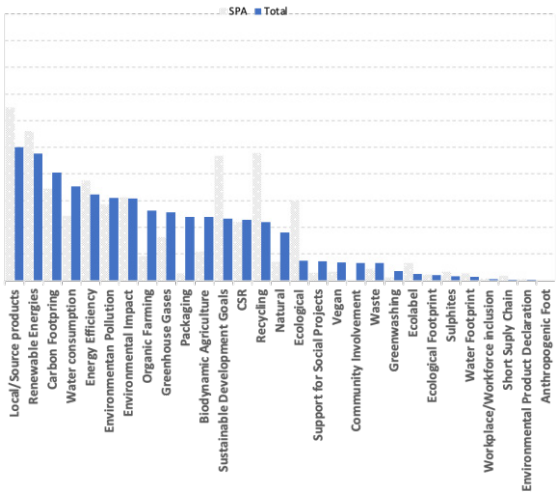

Figure 6. Percentage of mentions for Spain vs total.

In Spain (Fig. 6), the percentage of mentions of local/source products, renewable energies, energy efficiency, SDGs, recycling, ecolabel, or the general ecological category is higher than average. On the other hand, categories such as carbon footprint, water consumption, organic farming, packaging, biodynamic agriculture, natural, vegan, or greenwashing have a lower-than-average percentage of mentions.

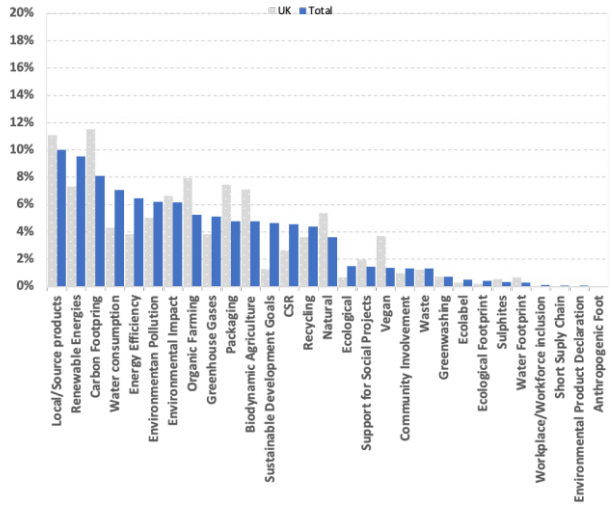

Figure 7. Percentage of mentions for the UK vs total.

In the UK (Fig. 7), the categories with a below-average percentage of mentions are related to renewable energies, water consumption, energy efficiency, environmental pollution, greenhouse gases, SDGs, CSR recycling or ecological. On the other hand, categories such as carbon footprint, organic farming, packaging, biodynamic, natural or vegan represent a higher-than-average percentage of mentions.

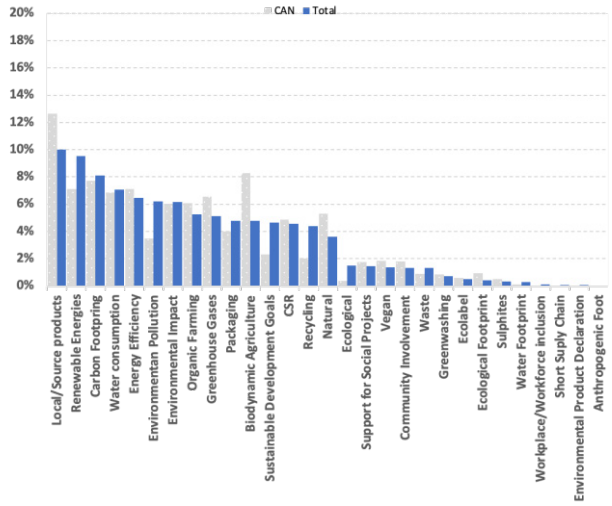

Figure 8. Percentage of mentions for Canada vs total.

The distribution of percentages of mentions in Canada (Fig. 8) shows how dimensions related to renewable energies, environmental pollution, SDGs, recycling or ecological represent categories with rates of mentions below the global average. On the other hand, categories such as local/source products, organic farming, greenhouse gases, biodynamic agriculture or natural show a higher-than-average percentage of mentions.

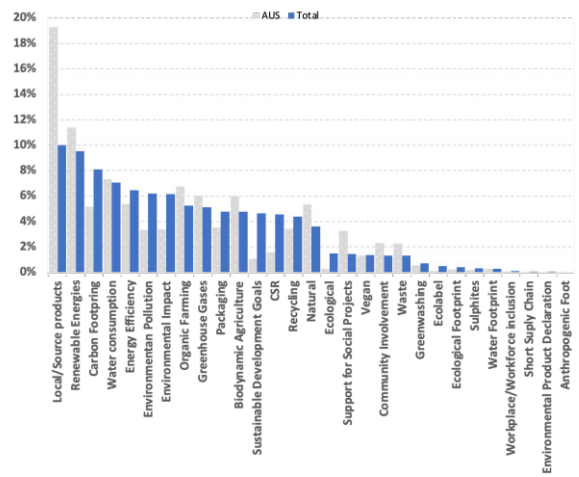

Figure 9. Percentage of mentions for Australia vs total.

Finally, in the case of Australia (Fig. 9), the themes with a higher-than-average percentage of mentions refer to local/source products, renewable energies, organic farming, greenhouse gases, biodynamic agriculture, natural, waste or social dimensions such as support for social projects or community involvement. On the other hand, issues such as carbon footprint, energy efficiency, environmental pollution, environmental impact, packaging, SDGs, CSR or ecological appear relatively less in the conversations.

5 Conclusions, limitations, and future research lines

The paper's results and research proposal represent an exploratory and general approach that constitutes a starting point for a promising area of research in the wine industry. Nevertheless, we can draw some conclusions, limitations, different proposals, and exciting research programs to be implemented in the wine sector.

5.1 Conclusions and Implications

The issue of sustainability in the wine sector value chain is a critical concern in the management of the organizations that integrate the sector and in the public regulation of the industry. Environmental aspects such as carbon footprint, energy efficiency, water footprint and consumption, greenhouse gas emissions, and social sustainability issues pose significant challenges for all actors in the wine industry.

Based on search queries concerning different environmental and social aspects of sustainability, we have conducted a content analysis of the information generated by users in different social media to detect the most relevant issues in the conversations among the actors in the digital ecosystem.

First, to highlight the relevance and usefulness of user-generated content analysis in understanding sustainability in the wine industry and wine markets. Thematic content generated by users and organizations is valuable for generating insights on current issues and market trends.

The importance of the different dimensions of sustainability is illustrated by considerable growth in mentions and conversations related to sustainability issues in the wine industry in recent years and in all countries with relevance to the wine business.

The importance of local/source products should be highlighted in the mentions relating to the different dimensions. The proximity of products is a significant issue in all markets, although a more detailed content analysis would be required to specify specifically which topics are included in this dimension. Secondly, aspects related to energy, resource efficiency and renewable energies are also an area of particular interest in the industry, particularly in an environment of rising energy prices and resource scarcity, such as water.

On the other hand, mentions of sustainability indicators also account for a significant amount of content. Thus, aspects such as carbon footprint, pollution, environmental impact, or greenhouse gases are highly relevant in the wine sector. Among the important topics, there are also mentions of packaging and recycling, highlighting the problem of the packaging's impact and the need for a recycling approach.

In a more specific way, and within the mentions with a certain relevance, the concepts of biodynamic agriculture and organic farming appear as concepts more associated with the sustainability of production.

Regarding social sustainability, the most relevant mentions are addressed to Sustainable Development Goals, which seem to have taken root in the sector, and Corporate Social Responsibility. In contrast, other aspects related to social sustainability have much less weight in the agents' conversations.

Regarding product-related dimensions, the most relevant concepts are those related to natural and, to a lesser extent, ecological or vegan, with the ecolabel or sulphites category being of practically residual importance.

Finally, mentions of categories such as greenwashing, short supply chain or Environmental Product Declaration are scarce, probably because these are very specific aspects for which users do not generate relevant content in social media.

A final observation concerns the differences in the relative importance of the issues for different countries. Although a more detailed analysis of the contents and their localization would be necessary, in the first instance, differences can be detected, which would be worth investigating further. In the case of the United States, the concern for proximity in products and the different dimensions of energy efficiency and environmental impact are similar or even below average. While the packaging concept has a greater presence, this is not the case for the recycling concept. Regarding social sustainability, the presence of CSR and, to a lesser extent, SDGs and other aspects should be highlighted. Product concepts such as natural, ecological, and vegan are similar to the average.

In the case of the UK, energy concerns (e.g. renewables, water, energy efficiency) are less relevant, and in the impact dimensions, carbon footprint and local products are relatively more important. The issue of biodynamic agriculture and organic farming is clearly above average, thus representing an essential element. As in the American market, the point of packaging is an aspect of interest to the British, although it is not as important as recycling. As far as product concepts are concerned, the idea of vegan and natural is more important than the concept of ecological. Finally, the UK shows the lowest mentions in terms of sustainability dimensions such as the SDGs or CSR.

In the case of the Canadian market, the importance of local products, biodynamic agriculture and organic farming, and the relatively more significant presence of concepts such as natural or vegan and, to a lesser extent, organic, should be highlighted. The fewer mentions of renewable energies, environmental pollution, SDGs, and recycling should be highlighted.

Finally, in the case of the Australian market, the substantial difference in mentions related to local products and renewable energies seems logical. Mentions of organic and biodynamic agriculture are also relevant. Although these aspects are of relatively low importance, it is worth highlighting the mentions of support for social projects or community involvement in the conversations. The concept of natural has more presence than organic or vegan. Finally, the aspects of energy efficiency or the dimensions of environmental impact seem to be less relevant.

5.2 Limitations

The shortcomings of the analysis have to do directly with the exploratory and descriptive nature of the study of the raw mentions and a need to go into more depth and detail on the specific contents of the mentions in the different categories and the sources from which they originate.

We have mixed the content generated by very different categories of users, including consumers, specialists and prescribers, and organizations and companies. As a result, the topics and interests of the conversations as well as the perspectives of each, are very different, creating the need for a segmented analysis.

Moreover, we have merged different sources or social spaces characterized by different conversation patterns. In this sense, most content comes from news sites, which may have a more technical or professional focus, while social networks such as Instagram generate more informal conversations.

In this first stage, we have worked with raw data. Although a system of queries has been set up to avoid entering irrelevant information, the filtering and structuring procedures must be further refined.

Finally, the aggregate analyses are purely descriptive. Although some interesting conclusions can be drawn, including more complex analysis tools incorporating more variables and indicators of the contents and mentions analysed is necessary.

5.3 Future research lines

Future research lines are linked to the limitations of the proposed research and, above all, to the possibilities offered by the analysis of UGC.

Firstly, and as a first aspect currently being worked on, is the specific analysis of the mentions' text contents through natural language processing tools that allow for identifying general and specific themes within the mentions. For example, Fig. 10 shows a preliminary analysis of the themes present in the global conversations resulting from all the queries. In this case, a thematic block related to restoration, tourism or travel is identified, which needs to be addressed in more detail.

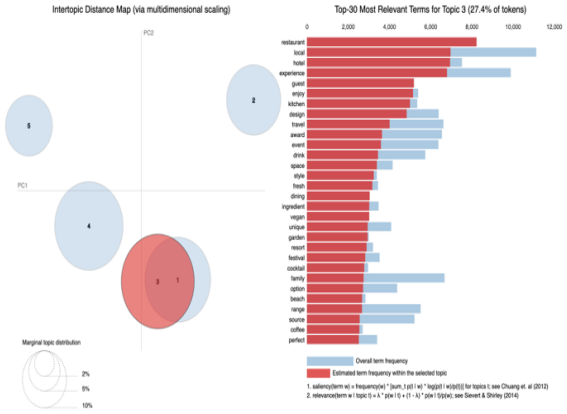

Figure 10. Topic identification using NLP (LDA).

The analysis carried out is purely descriptive and objective content analysis. However, it is possible to apply AI tools that allow a sentiment analysis to be conducted, which, in addition to the characteristics of the information, is also capable of providing an emotional assessment in the sense of whether it is positive, negative or neutral content.

As mentioned above, the analysis has been performed without segmenting the different categories of users from which the content originates. In this sense, it is essential to differentiate between organizations and individuals and, among the latter, between specialists or prescribers and other general users. The content and language in the conversations and the communication keys are different.

In addition to segmenting users, it is important to refine the approach by analyzing the different networks and media in which content is generated. For example, it would be interesting to analyze the information generated in the news separately or to consider niche blogs versus generalist social networks such as Twitter or aggregation spaces such as Reddit.

The possibility of geolocating content allows, in addition to the analysis of mentions by country, a specific content analysis by country or geographical area for those contents that can be geolocated in some way.

Considering broader temporal spaces, even the temporal span of this research, also makes it feasible to analyze the evolution of different topics or concepts over time, identifying trends in categories of mentions and

specific issues. This temporal analysis is of particular value in analyzing general trends and by country.

Finally, analyzing networks such as Twitter also enables the identification of influence networks using the users' reactions and the connections generated between them. This network analysis makes it possible to identify the most relevant nodes or agents creating trends in the sector.

Finally, as suggested in previous works, it is essential to approach the analysis of wine-specific networks and clubs in greater detail, focusing on the content analysis of reviews.

Research funding statement

This research is part of the project "Nuevos productos ambientalmente responsables derivados de estrategias de integración y valoración de innovaciones para la sostenibilidad integral en el sector vitivinícola y del aceite de oliva". (COOP-2021-CM-01) in the call of Dirección General de Agricultura y Ganadería, por la que se convocan, para el año 2021 las ayudas destinadas a promover la cooperación innovadora mediante la constitución de grupos operativos de innovación y la realización de proyectos piloto innovadores en la producción primaria agrícola y ganadera, en el marco del Programa de Desarrollo Rural de Castilla-La Mancha para el período 2014-2020. Extracto BDNS (Identif.): 562032, co-funded with FEADER/EAFRD funds.

References

1. Schaufele, U. Hamm, J. Clean. Prod. **147**, 379 (2017)
2. P. Tait, C. Saunders, P. Dalziel, P. Rutherford, T. Driver, M. Guenther, M., J. Clean. Prod. **233** 412 (2019)
3. E. Jorge, E. López-Valeiras, M.B. González-Sánchez, J. Clean. Prod. **257**, 120601 (2020)
4. M. Luca, Handbook of Media Economics (Vol. 1), 563 (2016)
5. J.R. Saura, D. Palacios-Marqués, D. Ribeiro-Soriano, Technovation, **january**, 102447 (2023)
6. N. Kotonya, N., P. De Cristofaro, E. De Cristofaro, *IEEE/ACM (ASONAM)* **august**, 387 (2018)
7. E. Mazzoli, I. Palumbo, L. *Available at SSRN 4114012* (2022)