Article

The Value of Web Data Scraping: An Application to TripAdvisor

Gianluca Barbera [1], Luiz Araujo [2] and Silvia Fernandes [2,3,*]

1   School of Political Sciences "Cesare Alfieri", University of Florence, 50127 Florence, Italy;
     gianluca.barbera@stud.unifi.it
2   Faculty of Economics, University of Algarve, 8005-139 Faro, Portugal; a78318@ualg.pt
3   CinTurs—Research Centre for Tourism, Sustainability and Well-being, University of Algarve,
     8005-139 Faro, Portugal
*   Correspondence: sfernan@ualg.pt

Abstract: Social Media Analytics (SMA) is more and more relevant in today's market dynamics. However, it is necessary to use it wisely, either in promoting any kind of product/brand, or interacting with customers. This requires its effective understanding and monitoring. One way is through web data scraping (WDS) tools that allow to select sites and platforms to compare them in their performances. They can optimize extraction of big data published on social media. Due to current challenges, a sector that can particularly take advantage of this source is tourism (and its related sectors). This year has the hope of tourism's revival after a pandemic whose impacts are still affecting several activities. Many traders and entrepreneurs have already used these versatile tools. However, do they really know their potential? The present study highlights the use of WDS to collect data from TripAdvisor's social pages. Besides comparing competitors' performance, companies also gain new knowledge of unnoticed preferences/habits. This contributes to more interesting innovations and results for them and for their customers. The approach used here is based on a project for smart tourism consultancy, from the identification of a gap in our region, to aid tourism organizations to enhance their digital presence and business model. Many things can be detected in this big source of unstructured data very quickly and easily without programming. Moreover, exploring code, either to refine the web scraper or connect it with other platforms/apps, can be an object of future research to leverage consumer behavior prediction for more advanced interactions.

Keywords: social media; data scraping; tourism; smart consultancy; cognitive system

1. Introduction

Social networks/channels are the internet tools most increasingly used. Their influence is seen in many sectors, but it is necessary to use them wisely. They are necessary to promote products, brands, or ideas; to consolidate one's own company or launch a new one; or to build loyalty with customers and seek new ones [1].

This year seems to be prepared for a social and economic revival after a 2022 that saw the world emerge from a heavy situation due to the pandemic. Many companies have been able to proceed through their own boldness or conscious use of online platforms, with emphasis on social media. These have many times been revealed as the best quality/price promotion tools. The results obtained also stem from a conscious investment in them. Those who invested wisely now see a return and will certainly acknowledge that one can no longer do without using social media and related apps.

One of the most affected sectors, given the movements dictated by government decrees to limit human mobility and circulation, has been tourism. This is a key sector for many regions and countries that claim to depend heavily on it [2]. Tourism is a sector that can also affect several other sectors, such as hospitality and traveling. Besides these, it can influence

handicraft, environmental, gastronomic, and artistic–cultural sectors. It is therefore decisive for the economic health of many people.

Hence, this present study is based on a project for smart tourism consultancy. It aims at analyzing, on one hand, the potential of web data scraping (WDS), and on the other hand the application of a web scraper to TripAdvisor's social pages. The period chosen is the whole year of 2022 due to its post-pandemic conditions with signs of revival. Moreover, specific related objectives deal with discerning unanswered trends and data-scraping levels. The web scraper used is Fanpage Karma (FpK), which will be compared with other similar tools to point out key features and choice impact. This approach elected TripAdvisor (TA) since it is an important travel site that provides information and reviews of tourism-related user-generated content. Finally, the results obtained are discussed in order to discern lessons for tourism companies and users.

This paper is thus organized as follows: in Section 2, the potential of WDS and initial view of the web scraper to discern relevant metrics and KPI; in Section 3, the application to TA's main social pages to discern main trends, issues, and lessons for better tourism management and forecasting; in Section 4, a comparison between WDS tools to discuss the importance of choosing the right solution by looking at the goals of three organizations in Algarve (interested in the original project), along with a holistic view toward advanced cognitive tools; and in Section 5 the conclusions, other trends, implications, and considerations for future research are provided.

2. Materials and Methods

The referred goal of a project for smart tourism consultancy was very welcomed by important organizations in Algarve (a very touristic region in Portugal) such as Dengun (one of the most successful digital marketing companies); CRIA (the innovation accelerator of the University of Algarve); and Tilia (an innovative hostel located in the capital's center). Their managers/owners are even interested in investing in the idea as there is nothing similar in the region.

2.1. WDS: A Big Data Source

As mentioned, TA is one of the largest websites for hotel, bed-and-breakfast, and restaurant reviews, accommodation bookings, and other travel-related content. It also includes interactive travel forums. This site was an early adopter of user-generated content and is supported by an advertising business model. As tourism is a key sector for many economies, it is important that such a platform be monitored in order to better predict and decide upon disruption events. Comparing their content quickly and easily, the level of engagement and other metrics can guide strategies and feedbacks on a timely, regular basis. For example, current concerns in the area are related to safety, sustainability, and well-being.

No company or entrepreneur can think of not being active on social networks today, since this medium is one of the largest data sources most widely used in the world [1]. Consequently, businesses and brands have been progressively investing in them as relevant communication channels [3]. Therefore, a practice in which many are investing to enhance performance through the analysis of desired markets/competition is WDS. Also known as web extraction, this technique extracts data from the internet and saves them to a file system or a database for analysis [4]. It is also widely acknowledged as a powerful technique for collecting big data [5,6]. Thus, web scraping is a practice that captures a large amount of data on the internet [7] to obtain knowledge about competing firms or websites.

Besides web scrapers, one of the most widely used resources for doing WDS is Python language. It is particularly used for the analysis of Twitter pages, through a dedicated library—Tweepy [8,9]. Another way of applying Python is through the Text Blob Tool which analyzes the sentiments of tweets [10]. Other tools in Python enable web-scraping activities such as Scrapy and Pandas [11], SpaCy [12], and Python's Natural Language Toolkit [13]. However, to extract information from a website by only having knowledge of

Python is not sufficient. It is necessary to know HTML because web scraping is a task that needs to be divided into subtasks. Once this marking language is known, which describes the elements of a page, Python libraries can be used more easily (such as Beautiful Soup, Scrapy, or Selenium) [14,15]. However, a software tool such as FpK is much more intuitive and easier to use by the end-user as it does not require knowing the programming code.

Nevertheless, programming language is crucial to create functions that SMA (Social Media Analytics) tools may lack, such as personalizing the scraper or connecting it with other tools or platforms for more advanced goals. An example of this kind of connection is sentiment analysis [16]. Figure 1 illustrates an example of text networks where, from the most prominent terms and their relations, firms might discern unanswered trends and prepare more assertive and timely strategies to cope with them.
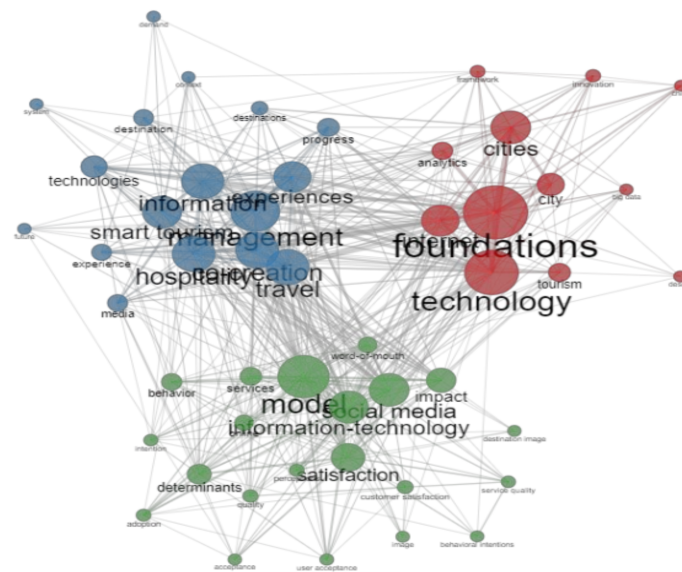

Figure 1. Example of text networks that can be used for sentiment analysis (own elaboration).

Other possibilities can be envisioned by integrating WDS with other tools or systems, which can be the object of future research on predicting consumer's behaviors. These discoveries can revolutionize tourism patterns and services [17]. Other aspects related to time and space can be captured for more data granularity (e.g., discover what happens in a certain specific place, etc.). This is relevant in today's context of these critical dimensions. These experiments can contribute to developing AI (artificial intelligence) or cognitive-driven decision-making to help firms be better equipped to fight disasters.

2.2. FpK: An Ease of Use Tool

FpK was launched in 2012 by two online marketing experts from Berlin. Their main motivation was to interpret and analyze Facebook pages, or other social media, in order to optimize them. It helps to manage social media presence and performance. It enables organizations to collect web data in order to compare their own social pages with the competitors' pages. Data about prices practiced, follower growth, and best time to post can be easily obtained. Key features are its ease-of-use and variety of functions related with web data extraction to produce insights. Besides calculating metrics, it has other actions such as benchmarking, tagging, history, live (see data in real-time), etc. (Figure 2). For example, it can check hotel room prices and other information, collect tweets, monitor reputation and SEO (Search Engine Optimization), and extract emails and addresses from map and business websites.
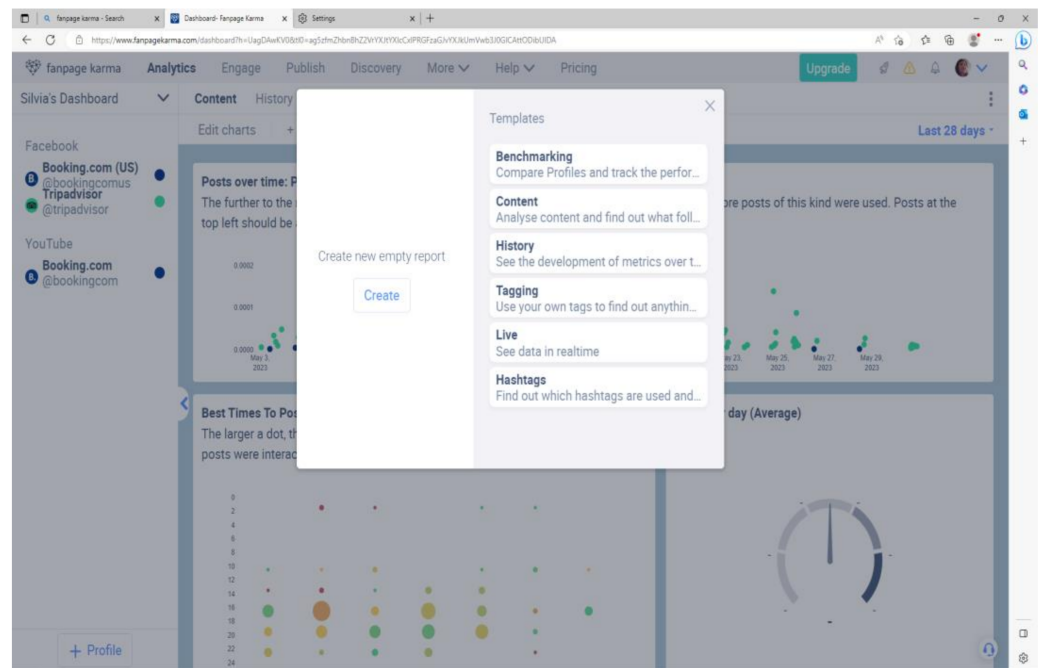
Figure 2. Creating a report of the analyses made (own elaboration).

This tool also allows visualizing several metrics such as number of followers in a competitor's page, its growth rate, engagement, response time, influencers, etc. Since it returns the results in a friendly design (including charts), their comparison is much easier. This is very interesting because the information about the competition's strategic options can be more easily discerned and understood. Moreover, this big source of unstructured data can then be accessed and used for deeper analyses of behavioral patterns and trends that went unnoticed before [18]. Currently in FpK there are more than 350 different metrics to choose and arrange in a dashboard, including the websites that can be added to it (Figure 3).
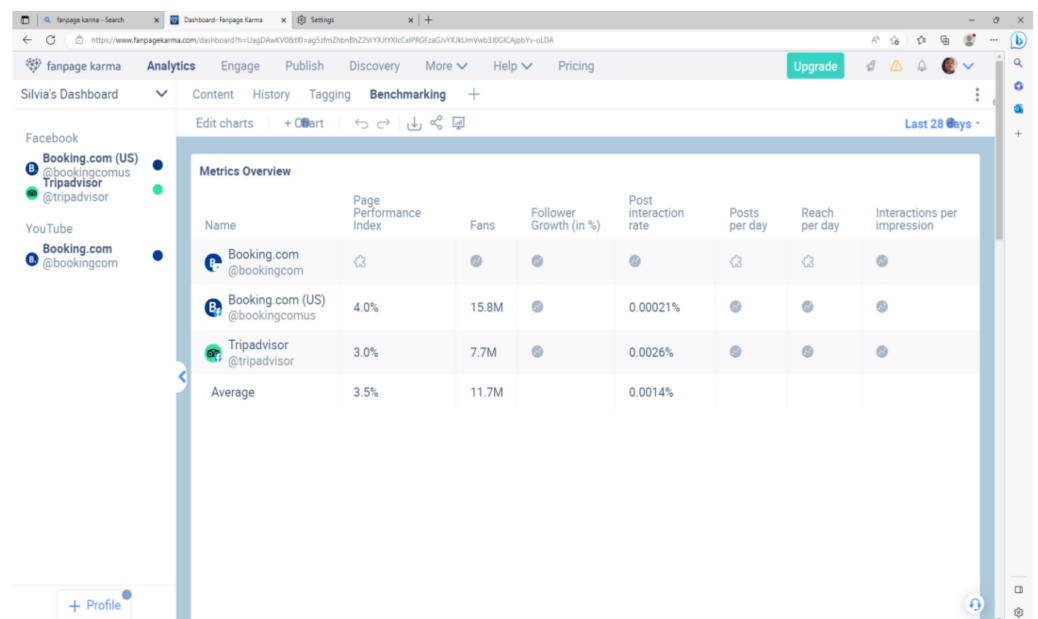

Figure 3. Function Analytics –> Benchmarking (own elaboration).

The dashboard is like a matrix of results that are the "raw data" for a graphical customized report. This report can help think of other metrics and combine them to obtain key performance indicators (KPI) that best fit the firm's goals. In addition, the function Analytics can create the dashboard for the last month (or year), or for comparing several months (or entire years) within the period needed.

For instance, two main KPIs of SMA are "follower growth" and "fan interactions" (engagement). The resulting chart (Figure 4) shows which site is the best. In this case, it is Booking.com (upper right corner) compared with TA (lower left corner). This is important to know in order to identify the factors behind it and better plan the structure and strategy of social media portfolios. However, these two social channels have different goals: TripAdvisor has more information because, besides hosting data, it shows reviews and recommendations for various places and activities, such as sights and restaurants.
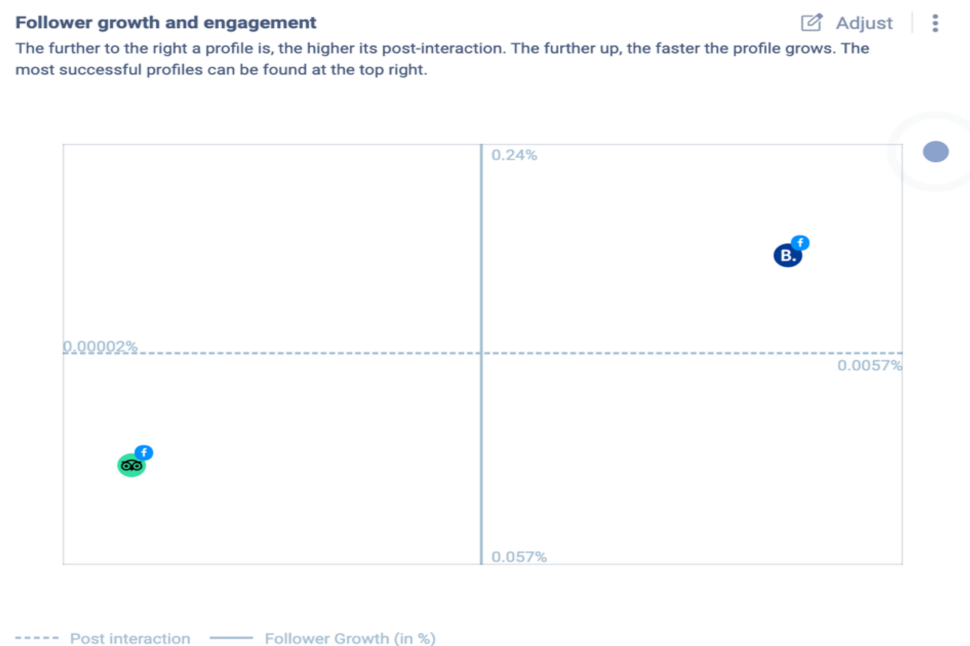

Figure 4. Main social performance metrics (own elaboration).

Effectively the web scraper FpK has been applied in several studies to investigate a wide variety of topics, which include political communication [19–21]; public communication [22]; healthcare [23]; customer behavior [24–26]; and tourism [27–30].

3. Results

Considering the main aim of this work—analyzing the potential of web data scraping (WDS) applied to TA's social pages—some trends of post-pandemic tourism can be discerned. As a 'sample', we have searched them among the main social channels of TA. This famous platform allows tourists to describe their experiences with hotels, restaurants, and services, among other aspects. Regarding the impact of the pandemic on people's travel habits, it is important to understand how this source of user-generated content has dealt with the pandemic and discern lessons for future tourism management.

Thus, this study has an exploratory nature and follows a comparative approach. It can then contribute to the tourism knowledge base to endure as one of the most important sectors worldwide. Through FpK, we collected data from TA's most active social networks in 2022 (Facebook, Instagram, Twitter, Pinterest, and TikTok). Then, we analyzed the resulting tables and charts. Several metrics were considered, such as page performance index, fans, follower growth, post interaction, and posts per day (Table 1).

Table 1. Performance of Tripadvisor's social pages (own elaboration). Period: 1 January to 31 December 2022.

TA Profiles	Page Performance Index	Fans	Follower Growth (%)	Post Interaction	Posts Per Day
TikTok	10.0%	1127	1843.1%	1.31%	0.052
Pinterest	10.0%	228,048	8.6%	0.0006%	1.038
Twitter	8.0%	3,424,208	0.97%	0.0005%	6.917
Instagram	14.0%	2,769,720	12.38%	0.06%	1.057
Facebook	9.0%	7,626,233	0.94%	0.0008%	4.608
Average	10.2%	2,809,867.2	373.2%	0.27%	2.734

Looking at the page performance index, which combines fans' engagement with page growth, we can see that Instagram had the highest value (14%). This is a platform that had around 1.28 billion users in 2022 (Statista, 2022). However, the figure that stands out the most is the % of follower growth on TikTok. However, TA only has had its TikTok page since the beginning of 2022 (having started posting only in October). It is followed by Instagram with 12.38% follower growth. Regarding post interaction, TikTok was also the best, but this is a biased metric for the same reason mentioned above. In conclusion, Instagram is confirmed as the social network where followers most interacted with the posts published. Finally, the social network that posts the most was Twitter, but it still has the least interaction level.

Therefore, this first scraping query shows at least two things: that more metrics are needed to understand these issues and that tourists/visitors engage more with visual social media (rich in pictures, short videos, and storytelling). This trend is augmenting what requires more experience to be prepared for future VR (virtual reality) or AI-based tourism/hospitality projects. Then, we pursued the analysis to discover which specific content got the most engagement from users. This information is important in order to change or improve campaigns and product/service innovation. Figure 5 shows which post features were published the most (left side of chart) and which ones had the most interactions (right side of chart) according to the following categories: links, pictures, status, carousels, reels, and videos.
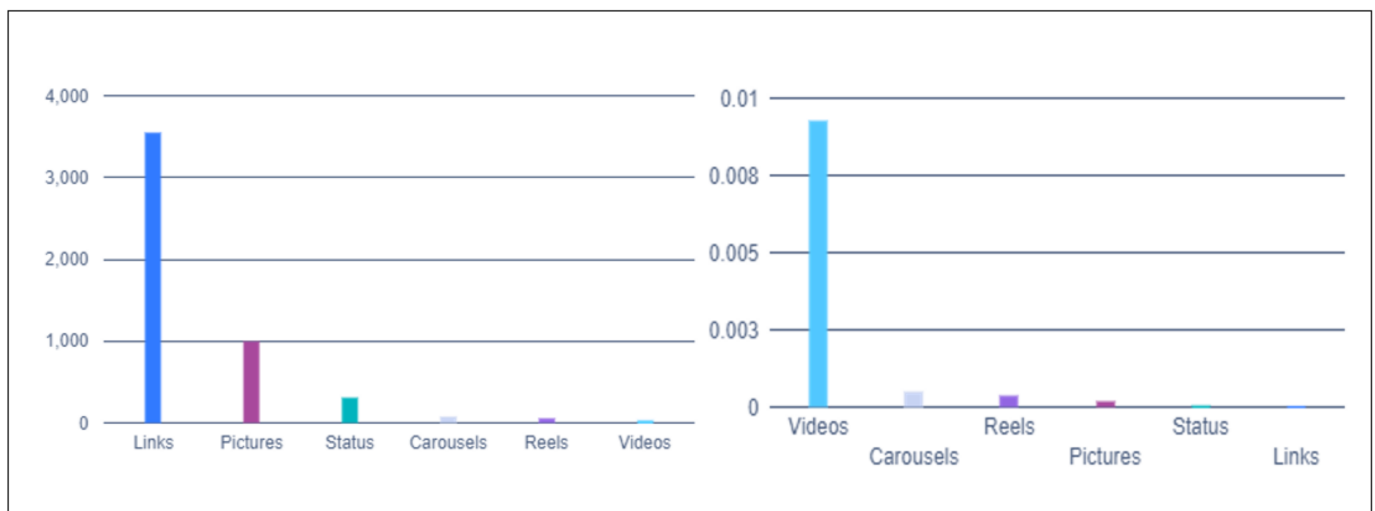
Figure 5. Number of posts and post-interaction in TA's social media (own elaboration).

Effectively, video posts had the highest level of interaction against pictures, which had the lowest level. The low results in the other categories (reels: full-screen vertical videos; carousels: ads that combine videos and images; status; and links) suggest (re)thinking the contents and the modern ways of presenting them. For instance, a study on

these issues [31] revealed an increasing relevance of interacting with influencers and of co-creating (i.e., including users' ideas in product innovation). Travel, hospitality, and tourism organizations should consider data from WDS tools along with case studies in this area to manage their social media more smartly. Their digital marketing plans must break into these trends to keep up with, or anticipate, the digital transformation.

4. Discussion

Regarding the results obtained about the potential of WDS for smart tourism consultancy, through an application to the TA case, we can acknowledge that more metrics of SMA are necessary. To release managers (and other users in the company) from the need to program them, a way of exploring either more metrics or more functions is by comparing web scrapers such as the nine best ones referred in the work of [32] (Table 2).

Table 2. The 9 best WDS tools (own elaboration, based on [32]).

Web Scraper	Key Features
Octoparse	Anonymous web data scraping behind login forms
ParseHub	Efficiency of data extraction from complex web pages
Mozenda	Job sequencers to collect web data in real-time; highly scalable
Webhose.io	Fast content indexing; get machine-readable data sets
Content Grabber	Allows to build web apps and offers a wide variety of formats
Common Crawl	Support for non-code usage; has resources to teach data analysis
Scrapy	Open-source tool and easily extensible. Middleware modules available for integrating useful tools
Scraperapi	Easy to integrate; allows price scraping, search engine scraping
Scrape-it.Cloud	Easy integration in other systems; scrapers for popular sites

In fact, this study uses FpK as an example of this type of specialized software tool. These are non-code web scrapers that enable users to extract data without the knowledge of HTML structures and elements [32]. However, each one has limitations and choosing the right one is a crucial task. For example, the previously mentioned organizations interested in this project—Dengun, CRIA, and Tilia Hostel—have different goals in mind. Dengun aims to use WDS insights to accelerate potential business in the phases of customer validation and proof of concept; CRIA has both academic and commercial interests: academic in terms of being able to absorb this type of information and transfer it to other sectors, and commercial as an enterprise incubator; and Tilia's owner envisions anticipating the demands of potential customers, maintaining the maximum capacity of its rooms, understanding new ways to approach customers at the right time, and realizing new marketing actions to interact with its publics. The other goals these three interveners mentioned are more assertiveness of investment in new businesses; hiring staff in advance of the high season; forecasting tourists with potential consumption in the region; forecasting segments to be explored for meeting new demands; and perceiving more assertively the desires of consumers.

4.1. Towards Cognitive Scraping

This work has shown that many things can be discovered through this big data source. Nevertheless, deeper aspects can be detected, especially by programming, in order to develop connections either within WDS tools or between them and other platforms. This issue raises the discussion around discerning unanswered trends and data-scraping levels.

Currently, 80% of available data are unstructured [33] such as those on the web. This wealth of data contains valuable information about customer habits, preferences, dislikes, intentions, and much more [34]. Well-extracted and analyzed, they may help identify innovation opportunities. Using the right tool(s), firms can start capturing value from this big data source [35]. The knowledge absorbed can give them a competitive advantage.

Considerable amounts of unstructured data are generated every minute [36]. Applying advanced cognitive computing to them will create more complete models of customers' behavior with the relationships and environments in which they act (Figure 6). In turn, the massive volume of these actions will enable better prediction of consumer behaviors [35,37].

For instance, a company can provide branded content to a customer on a social network platform, then track the person/group with whom he shares that content and how they respond. This can help to understand how consumers feel about its brands to find gaps in the marketplace and develop brand positions.

Analyzing data can be a time-consuming practice, but good cognitive analyses can leverage this big amount of data [38]. As this problem can become unsustainable to process, there is a need for better computer tools and developing languages [39]. The developments involved in this holistic view (Figure 6) should be further explored in tourism and hospitality since they have been focused on areas such as education, healthcare, commerce, and human augmentation [40].
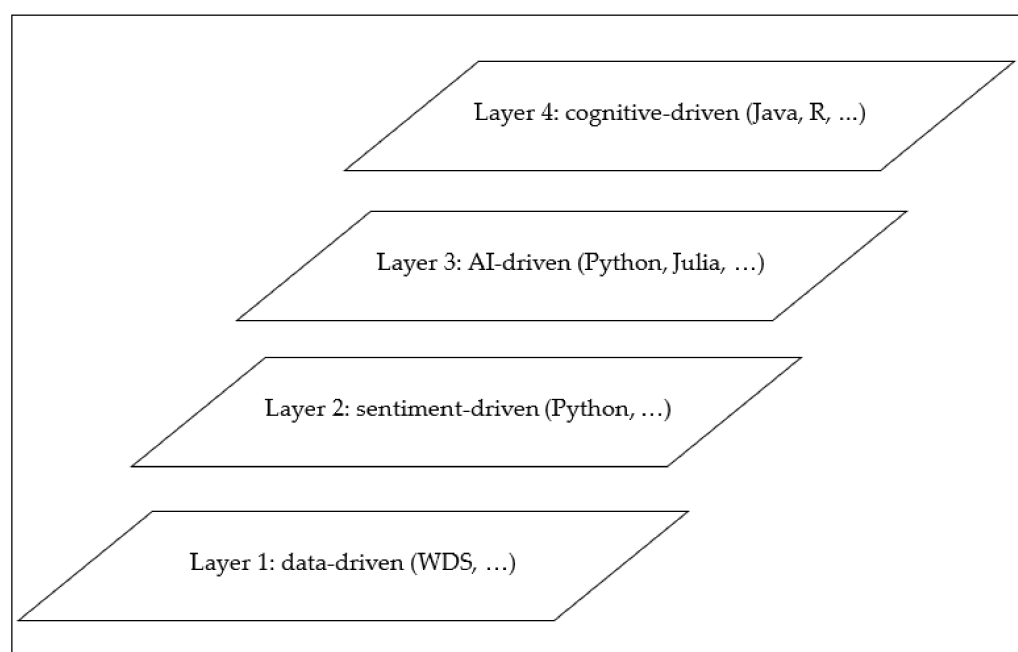

Figure 6. Holistic view toward advanced cognitive tools (own elaboration, based on [40,41]).

5. Conclusions

This study can contribute, through its innovative approach, to the discussion of web scraping's potential for scalable big data cognitive analytics. WDS tools can search for web data in a very fast and intuitive way, without running into legal issues, such as lack of consent to access private information, as well as wasting software/hardware resources. These tools can aid consultancy in tourism, among other segments, as they can optimize the extraction of competitor's data and other relevant information. This knowledge and its real-time monitoring can give companies a competitive advantage. They provide data quickly and clearly on the performance of any web page. As a case, we decided to explore which of TA's social pages performed best in 2022, the most recent post-pandemic year for tourism as a key sector for our region/country.

We saw that Instagram is the best performing social network, even though it does not have the highest number of followers. Its modern structure and interactive content certainly facilitate user engagement. Moreover, it is precisely through the right interactions (shares, comments, sales, etc.) that the firms obtain value and returns. Another interesting social service is TikTok, whose adoption is augmenting especially among young people. TA has noticed this trend and launched its own page in 2022. Regarding Twitter, the social

network that posted the most, it still has the least interaction level. It is mainly used to answer individual questions from users, and often reposts what has been published on other social media.

Given the pace of current challenges, web scraping can be a differential and timesaving resource. WDS tools can easily capture and compare this big data source on a regular basis. There are different ways of exploring their potential, some more dependent on programming and others more independent and intuitive. Several studies made use of both ways to extract and compare data from web and/or social media pages. However, we emphasize that web scrapers usually do not require knowledge of programming languages, which makes them easier to use/explore by managers and marketers. Many metrics can then be obtained and combined in versatile KPI with minimal effort.

5.1. Other Trends

A groundbreaking technology that has already started to influence the internet is augmented reality (AR). It may have a strong impact on social media apps in the coming years. It lies between the virtual and real world and is gradually gaining popularity as the virtual elements brought into the real world provide an immersive experience for the user. With the help of AR, 3D content will replace 2D content on social networks. A strong usage of interactive content will reshape social media as 3D content has the power to take the users to an entirely new and unimaginable world. Several social media players are quickly joining AR and VR such as Snapchat, Instagram, and Facebook [42].

Besides AR possibilities, other trends involve audio features. For example, HearMeOut is a voice-based social network that allows users to record and share audio posts. Moreover, the use of hands- and eyes-free social media allows users to do both multitasking and experiment greater 'reality' [43,44]. The social media landscape tends to witness significant trends and changes. With the continued technological innovation and changing user demands, platforms will focus more on personalized experiences, integration of AR and VR, and the enrichment of video content [45]. Entrepreneurs and marketers should stay up to date with these trends to be ahead of the curve and effectively engage their target audiences [46].

According to recent articles and reports [47,48] other trends to be considered are related with the increasing role of influencers; the need for more dynamic and interactive content; conditions for developing social media commerce/shopping as these platforms have wider target audiences; and usage of social media for mobile purposes and hybrid events among others. Companies that offer richer experiences to their customers will be much better equipped to face major challenges and stand up alongside competition.

5.2. Implications and Future Work

WDS can be used to gain insights toward key performance indicators. For instance, post interaction can be a good variable for analyzing which content, times, and places are more attractive and then rethink destination marketing strategies. Moreover, with these data, other more advanced analyses can be developed especially through programming connections either between WDS tools or between them and other platforms. Text networks are an example to build sentimental analyses and predict customer behavior more assertively.

Today, tourism organizations can choose from a rich set of marketing instruments. Marketers can now deal with paid (content they pay such as an ad or sponsorship), owned (content they create and control, like own Facebook page or website), and earned (content that others create like reviews or mentions) media [49]. However, these instruments can be better gathered and managed through the dashboard of the right WDS tool.

Currently any digital business model should approach SMA needs [50]. The focus of this study is relevant for advanced consultancy in tourism as it highlights web data extraction optimization to enhance the knowledge base for tourism innovation and revival. Given the importance of social media throughout the entire customer journey, tourism man-

agers need to know and understand the repercussions of their efforts. In fact, measuring social media performance has been considered as a major challenge of modern management [18,50]. In contrast to traditional media, social media resemble 'living' organisms [51] as these platforms involve content, motives, structure, roles and interactions. Therefore, more metrics should be developed and tested to understand and monitor their dynamics. Tourism associations should invest in hiring qualified staff, both on tourism dynamics and social media behavior. Consequently, invest in tools designed to provide a workspace and think-tank where unstructured data can be quick and easily analyzed. Moreover, consider and explore the developments toward advanced cognitive tools. The way people acquire and broaden their knowledge is changing toward cognitive acceleration. Cognitive computing models and systems can decipher unstructured data and draw valuable conclusions. These systems are capable of reasoning, decision-making, and experience-based learning [41]. They can communicate with people in natural language, comprehend real situations, and give personalized answers.

References

1. González-Padilla, D.; Tortolero-Blanco, L. Social media influence in the COVID-19 Pandemic. *Int. Braz. J. Urol.* **2020**, *46*, 120–124. [CrossRef] [PubMed]
2. Duro, J.; Perez-Laborda, A.; Turrion-Prats, J.; Fernández-Fernández, M. COVID-19 and tourism vulnerability. *Tour. Manag. Perspect.* **2021**, *38*, 100819. [CrossRef]
3. Ashley, C.; Tuten, T. Creative Strategies in Social Media Marketing: An Exploratory Study of Branded Social Content and Consumer Engagement. *Psychol. Mark.* **2015**, *32*, 15–27. [CrossRef]
4. Zhao, B. *Web Scraping. Encyclopedia of Big Data*; Springer: Berlin/Heidelberg, Germany, 2017; pp. 1–3. [CrossRef]
5. Kaisler, S.; Armour, F.; Espinosa, J.; Money, W. Big Data: Issues and Challenges Moving Forward. In Proceedings of the 46th Hawaii International Conference on System Sciences 2013, Wailea, HI, USA, 7–10 January 2013; pp. 995–1004. [CrossRef]
6. Bar-Ilan, J. Data collection methods on the Web for infometric purposes-A review and analysis. *Scientometrics* **2001**, *50*, 7–32. [CrossRef]
7. Mitchell, R. *Web Scraping with Python: Collecting More Data from the Modern Web*, 2nd ed.; O'Reilly: Sebastopol, CA, USA, 2018.
8. Kusumasari, B.; Prabowo, N. Scraping social media data for disaster communication: How the pattern of Twitter users affects disasters in Asia and the Pacific. *Nat. Hazards* **2020**, *103*, 3415–3435. [CrossRef]
9. Kaburuan, E.; Lindawati, A.; Surjandy; Sinswantini; Putra, M.; Utama, D. A Model Configuration of Social Media Text Mining for Projecting the Online-Commerce Transaction (Case: Twitter Tweets Scraping). In Proceedings of the 7th International Conference on Cyber and IT Service Management (CITSM) 2019, Jakarta, Indonesia, 6–8 November 2019; pp. 1–4. [CrossRef]
10. Kaur, C.; Sharma, A. Social Issues Sentiment Analysis using Python. In Proceedings of the 5th International Conference on Computing, Communication and Security (ICCCS) 2020, Patna, India, 14–16 October 2020; pp. 1–6. [CrossRef]
11. Raman, D.; Jayalakshmi, S.; Arumugam, K.; Raj, A.; Balaji, D.; Brightsingh, R. Implementation of Data Analysis and Document Summarization in Social Media Data Using R and Python. In Proceedings of the 4th International Conference on Inventive Research in Computing Applications (ICIRCA) 2022, Coimbatore, India, 21–23 September 2022; pp. 1457–1464. [CrossRef]
12. Bhardwaj, B.; Ahmed, S.; Jaiharie, J.; Dadhich, R.; Ganesan, M. Web Scraping Using Summarization and Named Entity Recognition (NER). In Proceedings of the 7th International Conference on Advanced Computing and Communication Systems (ICACCS) 2021, Coimbatore, India, 19–20 March 2021; pp. 261–265. [CrossRef]
13. Dansana, D.; Adhikari, J.; Mohapatra, M.; Sahoo, S. An Approach to Analyse and Forecast Social media Data using Machine Learning and Data Analysis. In Proceedings of the International Conference on Computer Science, Engineering and Applications (ICCSEA) 2020, Gunupur, India, 13–14 March 2020; pp. 1–5. [CrossRef]

14. Camargo-Henríquez, I.; Núñez-Bernal, Y. A Web Scraping based approach for data research through social media: An Instagram case. In Proceedings of the V Congreso Internacional en Inteligencia Ambiental, Ingeniería de Software y Salud Electrónica y Móvil (AmITIC) 2022, San Jose, Costa Rica, 14–16 September 2022; pp. 1–4. [CrossRef]

15. Zou, J.; Le, D.; Thoma, G. Locating and parsing bibliographic references in HTML medical articles. *Int. J. Doc. Anal. Recognit.* **2010**, *13*, 107–119. [CrossRef]

16. Korab, P. Text Network Analysis: Generate Beautiful Network Visualisations. 2022. Available online: https://towardsdatascience.com/text-network-analysis-generate-beautiful-network-visualisations-a373dbe183ca (accessed on 21 May 2023).

17. Alaei, A.R.; Becken, S.; Stantic, B. Sentiment Analysis in Tourism: Capitalizing on Big Data. *J. Travel Res.* **2019**, *58*, 175–191. [CrossRef]

18. Boegershausen, J.; Datta, H.; Borah, A.; Stephen, A.T. Fields of Gold: Scraping Web Data for Marketing Insights. *J. Mark.* **2022**, *86*, 1–20. [CrossRef]

19. Màrquez-Domínguez, C.; López López, P.; Arias, T. Social networking and political agenda: Donald Trump's Twitter accounts. In Proceedings of the 12th Iberian Conference on Information Systems and Technologies (CISTI) 2017, Lisbon, Portugal, 21–24 June 2017; pp. 1–6. [CrossRef]

20. Tarai, J.; Kant, R.; Finau, G.; Titifanue, J. Political Social Media Campaigning in Fiji's 2014 Elections. *J. Pac. Stud.* **2015**, *35*, 89–114.

21. Rullo, L.; Nunziata, F. "Sometimes the Crisis Makes the Leader?" A Comparison of Giuseppe Conte Digital Communication before and during the COVID-19 Pandemic. *Comun. Politica* **2021**, *3*, 309–332. [CrossRef]

22. Mabillard, V.; Zumofen, R.; Pasquier, M. Local governments' communication on social media platforms: Refining and assessing patterns of adoption in Belgium. *Int. Rev. Adm. Sci.* **2022**, 1–17. [CrossRef]

23. Martínez, T. Comunicación y diabetes, un camino para la reflexión. *RedMarka-Rev. De Mark. Apl.* **2022**, *26*, 96–113. [CrossRef]

24. Jayasingh, S.; Venkatesh, R. Customer Engagement Factors in Facebook Brand Pages. *Asian Soc. Sci.* **2015**, *11*, 19. [CrossRef]

25. Huertas, A.; Marine-Roig, E. User reactions to destination brand contents in social media. *Inf. Technol. Tour.* **2016**, *15*, 291–315. [CrossRef]

26. Caldevilla-Domínguez, D.; Barrientos-Báez, A.; Padilla-Castillo, G. Dilemmas Between Freedom of Speech and Hate Speech: Russophobia on Facebook and Instagram in the Spanish Media. *Politics Gov.* **2022**, *11*, 1–13. [CrossRef]

27. Martínez-Fernández, V.; Amboage, E.; Burneo, M.; Benitez, V. La gestión de los medios sociales en la dinamización de destinos turísticos termales: Análisis crosscultural de modelos aplicados en España, Portugal y Ecuador. *Hologramática* **2015**, *2*, 47–60.

28. Sánchez-Jiménez, M.; Matos, N.; Correia, M. Evolution of the presence and engagement of official social networks in promoting tourism in Spain. *J. Spat. Organ. Dyn.* **2019**, *7*, 210–225.

29. Sánchez-Jiménez, M. Análisis de la comunicación digital oficial en la promoción turística de Brasil. *3c TIC-Cuad. De Desarro. Apl. A Las TIC* **2020**, *9*, 17–39. [CrossRef]

30. Lee, M. Evolution of hospitality and tourism technology research from Journal of Hospitality and Tourism Technology: A computer-assisted qualitative data analysis. *J. Hosp. Tour. Technol.* **2021**, *13*, 62–84. [CrossRef]

31. Pereira, P. Social Media Influencers in Travel and Tourism. Master's Thesis, Master Course in Information Management. Nova Information Management School, Lisbon, Portugal, 2023.

32. Phaujdar, A. 9 Best Web Scraping Tools. 2021. Available online: https://hevodata.com/learn/web-scraping-tools/ (accessed on 22 May 2023).

33. Rizkallah, J. The Big (Unstructured) Data Problem. 2017. Available online: https://www.forbes.com/sites/forbestechcouncil/2017/06/05/the-big-unstructured-data-problem/?sh=cd00fa3493a3 (accessed on 23 March 2023).

34. Selz, D. Unstructured Data Is Key to True Customer Insight. 2017. Available online: https://www.linkedin.com/pulse/unstructured-data-key-true-customer-insight-dorian-selz (accessed on 23 March 2023).

35. Chen, S.; Kang, J.; Liu, S.; Sun, Y. Cognitive computing on unstructured data for customer co-innovation. *Eur. J. Mark.* **2020**, *54*, 570–593. [CrossRef]

36. Marr, B. How Much Data Do We Create Every Day? *The Mind-Blowing Stats Everyone Should Read.* 2018. Available online: https://www.forbes.com/sites/bernardmarr/2018/05/21/how-much-data-do-we-create-every-day-the-mind-blowing-stats-everyone-should-read/?sh=4de1a9aa60ba (accessed on 12 March 2023).

37. Ruan, Z.; Siau, K. Digital Marketing in the Artificial Intelligence and Machine Learning Age. Americas Conference on Information Systems. 2019. Available online: https://www.semanticscholar.org/paper/Digital-Marketing-in-the-Artificial-Intelligence-Ruan-Siau/5d0764dbe4cb3beb6c194b49a4eae1a991a72cd8 (accessed on 13 March 2023).

38. Kim, H.; Chan, H.; Gupta, S. Examining information systems infusion from a user commitment perspective. *Inf. Technol. People* **2016**, *29*, 173–199. [CrossRef]

39. Changchit, C.; Chuchuen, C. Cloud computing: An examination of factors impacting users' adoption. *J. Comput. Inf. Syst.* **2018**, *58*, 1–9. [CrossRef]

40. Biedrzycki, N. Cognitive Computing. What Can It Be Used for? 2020. Available online: https://towardsdatascience.com/cognitive-computing-what-can-it-be-used-for-8af4721928f5 (accessed on 26 May 2023).

41. Frackiewicz, M. The Role of NLP in Cognitive Computing. 2023. Available online: https://ts2.space/en/the-role-of-nlp-in-cognitive-computing/ (accessed on 26 May 2023).

42. Rao, L. Instagram Copies Snapchat Once again with Face Filters. 2017. Available online: https://tinyurl.com/ybcuxxdv (accessed on 5 June 2023).

43. Perry, E. Meet HearMeOut: The Social Media Platform Looking to Bring Audio Back into the Mainstream. 2018. Available online: https://tinyurl.com/y8yxbzah (accessed on 5 June 2023).
44. Katai, L. 3 Reasons Why Audio Will Conquer All Social Media. 2018. Available online: https://www.adweek.com/performance-marketing/3-reasons-why-audio-will-conquer-social-media/ (accessed on 5 June 2023).
45. Shahid, M.Z.; Li, G. Impact of Artificial Intelligence in Marketing: A Perspective of Marketing Professionals of Pakistan. *Glob. J. Manag. Bus. Res.* **2019**, *19*, 27–33.
46. Dwivedi, Y.; Ismagilova, E.; Hughes, D.; Carlson, J.; Filieri, R.; Jacobson, J.; Jain, V.; Karjaluoto, H.; Kefi, H.; Krishen, A.; et al. Setting the future of digital and social media marketing research: Perspectives and research propositions. *Int. J. Inf. Manag.* **2021**, *59*, 102168. [CrossRef]
47. Zoho Social. Social Media Marketing Trends for 2022. 2021. Available online: https://www.zoho.com/social/journal/social-media-marketing-trends-2022.html (accessed on 5 June 2023).
48. NBBJ. Social Media Is Evolving Quickly, and Your Business Needs to Also. 2022. Available online: https://www.northbaybusinessjournal.com/article/industrynews/social-media-is-evolving-quickly-and-your-business-needs-to-also/ (accessed on 5 June 2023).
49. Corcoran, S. Defining Earned, Owned and Paid Media. 2009. Available online: https://www.forrester.com/blogs/09-12-16-defining_earned_owned_and_paid_media/ (accessed on 12 March 2023).
50. Wozniak, T.; Stangl, B.; Schegg, R.; Liebrich, A. Do Social Media Investments Pay Off? Preliminary Evidence from Swiss Destination Marketing Organizations. In Proceedings of the ENTER eTourism Conference 2016, Bilbao, Spain, 2–5 February 2016.
51. Peters, K.; Chen, Y.; Kaplan, A.; Ognibeni, B.; Pauwels, K. Social media metrics-A framework and guidelines for managing social media. *J. Interact. Mark.* **2013**, *27*, 281–298. [CrossRef]