

Article

User Analytics in Online Social Networks: Evolving from Social Instances to Social Individuals

Gerasimos Razis ^{1,*} , Stylianos Georgilas ¹, Giannis Haralabopoulos ² and Ioannis Anagnostopoulos ¹

¹ Computer Science and Biomedical Informatics Department, University of Thessaly, 35131 Lamia, Greece

² Henley Business School, University of Reading, Reading RG6 6UD, UK

* Correspondence: razis@uth.gr

Abstract: In our era of big data and information overload, content consumers utilise a variety of sources to meet their data and informational needs for the purpose of acquiring an in-depth perspective on a subject, as each source is focused on specific aspects. The same principle applies to the online social networks (OSNs), as usually, the end-users maintain accounts in multiple OSNs so as to acquire a complete social networking experience, since each OSN has a different philosophy in terms of its services, content, and interaction. Contrary to the current literature, we examine the users' behavioural and disseminated content patterns under the assumption that accounts maintained by users in multiple OSNs are not regarded as distinct accounts, but rather as the same individual with multiple social instances. Our social analysis, enriched with information about the users' social influences, revealed behavioural patterns depending on the examined OSN, its social entities, and the users' exerted influence. Finally, we ranked the examined OSNs based on three types of social characteristics, revealing correlations between the users' behavioural and content patterns, social influences, social entities, and the OSNs themselves.

Keywords: online social networks; Twitter; Facebook; Instagram; social influence; social analytics; behavioural patterns; social individuals



Citation: Razis, G.; Georgilas, S.; Haralabopoulos, G.; Anagnostopoulos, I. User Analytics in Online Social Networks: Evolving from Social Instances to Social Individuals. *Computers* **2022**, *11*, 149. <https://doi.org/10.3390/computers11100149>

Academic Editor: Fernando Bobillo

Received: 20 September 2022

Accepted: 5 October 2022

Published: 7 October 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Our big data era is characterised by an overflowing information overload, mainly due to the abundance of content and resources, which are often of questionable completeness or quality. Different types of stakeholders, from citizens to decision-makers and from start-ups to governments, need to utilise a variety of resources [1] to fully satisfy their data and informational needs, as content can be diffused across different OSNs and websites [2]. Usually, each resource is focused on specific aspects and, consequently, information stakeholders must rely on multiple heterogeneous sources [3] to acquire an in-depth perspective on a subject.

The same principle applies to the online social networks (OSNs), an integral part of our everyday lives in terms of communication, news discovery and data consumption. Three of the most popular services are Twitter, Facebook, and Instagram. Nowadays, every minor or major event is published and is instantly accessible and visible to the world. The innate human desire for belonging and socialising is reflected by the fact that, in early 2022, approximately 4.5 billion people worldwide were OSN users (<https://www.statista.com/statistics/454772/number-social-media-user-worldwide-region/>, accessed on 29 August 2022). Moreover, according to the Global Web Index [4], in 2020, the average number of OSN accounts maintained by Millennials or Generation Z (born between 1981 and 2012) was 8.9, an increase of 44% from the 6.2 accounts in 2015. This growth of multi-networking is attributed to the specialisation of the individual OSNs. Instagram specialises in the sharing of photos, YouTube in videos, Twitter in short textual messages, and LinkedIn in work and business-related content. However, OSNs constantly upgrade, improve, and

expand their services to increase the loyalty of their users and prevent their transition to other platforms.

Consequently, end-users need to maintain accounts in multiple OSNs to gain access to a wide range of services for a complete social networking experience, since each OSN has a different philosophy in terms of the services offered, dissemination of content, user interactions, and user experience. Thus, the study of the consistency of, or variation in, individuals' behaviour in each OSN via multiple social instances and potential differentiation between the types of shared content would constitute a research topic of interest. To the best of our knowledge, there is no existing comprehensive study of this nature.

There are three aims of this study. Firstly, contrary to the literature, we examine users' behavioural and disseminated content patterns under the assumption that accounts maintained by users in several OSNs should not be regarded as distinct accounts but rather as the same social individual with multiple social instances. Secondly, we propose a framework for collecting and storing OSN users' disseminated content from multiple social platforms. Our social analysis, enriched with information about the examined users' social influences, reveals the existence of behavioural patterns that depend on the OSN, its social entities (e.g., links, hashtags, mentions), and the users' exerted influences. Such a study on user behaviour in multiple OSNs that co-considers their social influences has not been reported in the literature. Thirdly, we rank the examined OSNs and the social influence groups based on three types of social characteristics, namely social entities, social acceptance, and social conversation. These characteristics expose correlations between the behavioural patterns of the same social individuals (users), the degree of social influences, the appearance of social entities, and the OSNs themselves.

In this study, a predefined list of social individuals who maintained accounts (social instances) in all three examined OSNs, including Twitter, Facebook, and Instagram, were examined. We identified 57 social individuals who satisfied the multi-OSN criterion. Their accounts in all three OSNs were identified under semi-supervision. Our study differs from the majority of the literature, where random accounts are selected from multiple OSNs to perform the social analysis. Contrary to this process, we examined the behavioural patterns of the same individuals (OSN users) across all three OSNs where accounts (their social instances) are maintained. Thus, any observed consistencies or differences in the user behavioural patterns in the examined OSNs are significant, since the conclusions are drawn from the exact same group of individuals rather than from unrelated and random accounts.

The remainder of this work is organised as follows. In the next section, an overview of the related studies is presented, describing the benefits of the homogenised social content along with the research directions deriving from the investigation of social profiles appearing in multiple OSNs. Then, in Section 3, an overview of the architecture of our service is analytically described, with the purpose of the identification of the social individuals' behavioural and disseminated content patterns in OSNs. In Section 4, we define a methodology for the quantification of social influence, which is used in our social analysis. In Section 5, we analytically present and discuss the results of the proposed framework, co-considering the influential OSN dynamics of the social individuals. In Section 6, we rely on specific OSN-related metrics to identify the behavioural patterns of the same users in multiple OSNs by examining the users' exerted influences. Finally, in Section 7, we present the conclusions of our study by summarising the outcomes and proposing future directions.

2. Related Work

It is known that users maintain accounts in multiple OSNs and that homogenised content enhances analytics' capabilities, as this topic has been investigated since the rise of OSNs. Specifically, in 2008, the authors of [5] relied on semantic web technologies and the FOAF (Friend of a Friend) ontology to merge multiple accounts into distinct social profiles. To achieve this, reasoning techniques were employed in combination with the available FOAF semantified social data provided by 11 blogging websites. The results revealed that a social profile can consist of up to six accounts. The FOAF ontology was also utilised

in [6] to generate user graphs describing different OSNs. A graph-based similarity metric was employed to identify social accounts belonging to the same user. Recently, a plethora of profile matching algorithms [7–10] for identifying user profiles in multiple OSNs were also developed.

The same area is also investigated in [11], where public information on four OSN accounts were collected from aggregator services, which enriched the original social data, to enable the disambiguation and merging of distinct social profiles belonging to the same user from different OSNs. It was reported that the “Name” field was the most discriminative feature in the disambiguation process.

The need to unify the information found in multiple OSNs, along with the outdated view of focusing on a single OSN for efficient user and event analysis, are discussed in [3]. The authors suggested the inclusion of multiple OSNs and heterogeneous data sources when aiming to develop a comprehensive understanding of events, thus proposing a unified semantic model for event analysis. Similar to other works in the field, an ontological model for transforming and homogenising the retrieved social information was applied.

Along the same lines, the study reported in [12] presents an information system for collecting social content from several heterogeneous OSNs, also relying on an ontological model, with the aim of merging social profiles from different OSNs. The social information discussed and analysed in this work is similar to that which we collected in ours.

The use of ontologies, along with the derived benefits of integrating raw isolated social data, are discussed in [13]. Specifically, a semantic data integration methodology for providing unified access to, and enabling the analysis of, data on three OSNs was presented. Moreover, it was reported that, due to heterogeneity, the data preparation and pre-processing steps tend to be the most time-consuming processes.

As we can see, research on, and corporal interest in, unifying OSN profiles is significant due to the subsequent benefits, which can be applied in a plethora of applications. However, contrary to most works, we do not investigate the approaches of such a unification but rather the benefits deriving from the unified profiles and, in particular, from our proposed novel social analytics. Our social individual profiles, presented in Section 4.1, were semi-automatically identified on Twitter, Facebook, and Instagram, relying on the “Name” field of their social instances, namely the accounts, as suggested by [11].

Recent studies have focused on assessing and comparing the behavioural patterns of OSN users across multiple platforms with the aim of drawing multi-discipline practical and theoretical conclusions. Specifically, in [14], the image-sharing practices of five OSNs (Facebook, Twitter, Instagram, Snapchat, and WhatsApp) were investigated to identify users’ gender-based patterns. It was revealed that gender differences exist in terms of image sharing, as users adjust their communication approach variably depending on the platform. The examined users did not necessarily maintain an account in each OSN platform.

The authors of [15] examined the online bridging and bonding social capital of four OSNs (Twitter, Facebook, Instagram, and Snapchat) by analysing the responses that their users provided in a survey. Some of these users maintained multiple accounts in the OSNs, but the analysis was performed only on the platform that they most frequently used. The study concluded that characteristics such as trust in the OSN, social relationship strength, homophily, social comparison, and privacy concerns affect the frequency of use of the OSNs and the social capital. The examined users did not necessarily maintain an account in all the OSNs. Similarly, in [16], a survey was utilised to identify the usage patterns and preferences of Twitter, Facebook, and Instagram users. The results showed a correlation between the personal characteristics and demographics of the users and their OSN usage and preferences.

The study reported in [17] is closely related to our work, as it not only discussed social influence, but also proposed a mechanism for quantifying this influence across multiple OSNs, including Twitter, Facebook, and Instagram. To this end, a set of manually selected OSN accounts were utilised, whereas the social data were collected from the OSN APIs. The

same rationale is also followed in our work, as presented in Section 3. The study concludes with a regression model relying on a set of features affecting social influence.

The three OSNs of Twitter, Facebook, and Instagram were also investigated in [18] with the aim of identifying correlations between OSN usage patterns and user loneliness and well-being. Based on the responses of a survey, it was reported that, depending on the OSNs used (one or more), different psychological patterns appeared, affecting the social interactions. In the same line of thought, the authors of [19] assessed the correlation of OSN usage with romantic relationship happiness. A prerequisite for the participants was their possession of active accounts in all three examined platforms. It was revealed that active Twitter and Instagram use was negatively associated with romantic relationship happiness.

The effects of OSN usage on college social adjustment were investigated in [20]. To this end, the interactions of Twitter, Facebook, and Instagram college users were categorised into two groups: friends and family or OSN users who are complete strangers. It was revealed that users communicating primarily with familiar persons adjusted better to college society compared to those interacting with strangers.

Finally, the Twitter, Facebook, and Instagram posts of the leaders of three political Canadian parties were examined in [21] during the election period. This study discussed how the OSN platform characteristics, disseminated content, and available actions in regard to the messages affect user engagement. The analysis revealed that engagement is not produced uniformly across the OSNs, even when stimulated by the same person discussing the same topic.

3. User Analytics in OSNs: Our Approach

In this section, we present an architectural overview of our service with the aim of identifying OSN users' behavioural and disseminated content patterns under the assumption that accounts maintained by users in each of these OSNs are not regarded as distinct accounts, but rather as the same individual with multiple social instances/personas. Moreover, we present our database schema for storing multi-instanced social information and the employed APIs used for acquiring the social content from the three investigated OSNs in detail.

3.1. Service Architecture

The three-layered architecture of our service, along with the relevant data flows, are presented in Figure 1. Our proposed service consists of multiple interconnected components. Its design facilitates its maintenance and expandability due to the decoupled layers, while every layer locally comprises all technical and implementation dependencies. Specifically, the gathered social information is stored and accessed by and from the dedicated "Persistence Storage" component, which consists of a relational database in the "Data" layer, represented by the green highlighted rectangle.

All business logic, data gathering, conversion, querying, and processing tasks are handled by the "Processing" layer, represented by the blue highlighted rectangle. Therefore, this layer consists of three components: (a) the "OSN APIs" layer (Section 3.2), responsible for the collection of social information from the three investigated OSNs (Twitter, Facebook, Instagram); (b) the "Data Parsing and Analysis" layer, where the raw social information provided by the "OSN APIs" component is transformed (e.g., social accounts, disseminated content); and (c) the "Data Querying" layer, handling the service's and/or users' data requests and facilitating the database search operations. The collected social content is then converted into the standardised format required by the "Persistence Storage" and "Social Analytics" components, in order to measure the social influences of the Twitter accounts (Section 4). Finally, the "Visualization" layer, represented by the grey highlighted rectangle in Figure 1, consists of the "Social Analytics" component, providing valuable insights regarding the OSN users' behavioural and disseminated content patterns, accessible via the web-based "Interactive User Interface (UI)" component. Moreover, this component allows several administrative actions to be performed and provides access to reports regarding

the progress and status of the various processes (e.g., gathering of social data, OSN API requests, and so on) and configuration parameters, as well as access to the database content.

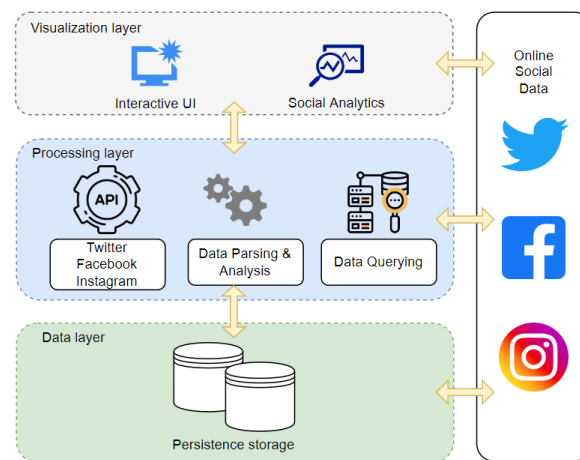


Figure 1. The three-layered architecture of our service.

As can be seen in the right-hand side of Figure 1, the processed or raw collected OSN content can be accessed by all layers of the service in order to enable the enriched or original social data to be analysed, when needed, by the appropriate components.

3.2. Database Model

Figure 2 presents the relational database model of the “Persistence Storage” component of our system, defining the logical structure of our database, in the format of an Entity–Relationship (ER) diagram. An ER diagram is a flowchart illustrating how the conceptual entities within a system relate to each other. Such a diagram relies on a defined set of symbols and lines to depict the interconnections between the conceptual entities, their relationships, and attributes. Our database model consists of five types of tables, as analysed below:

1. OSN accounts: the tables “Twitter”, “Facebook”, and “Instagram” store the basic metadata regarding the social accounts.
2. OSN posts: the tables “Tweets”, “FacebookPosts”, and “InstagramPosts” store the basic metadata regarding the social messages.
3. OSN entities: the tables “Hashtags”, “Links”, and “Media” store these social entities.
4. User alias: the table “User” allows us to analyse the seemingly unconnected social instances, namely the OSN accounts, as social individuals. Essentially, this table assigns an alias to an entity (social individual) that maintains accounts in all three examined OSNs (social instances).
5. Joining tables: these supportive structures are created to represent the many relationships in our schema, since relational database management systems (RDBMS) do not support the direct implementation of said relationships between tables. Indicative examples are the “Links2Tweets”, “Media2Fb”, and “Hash2Insta” tables.

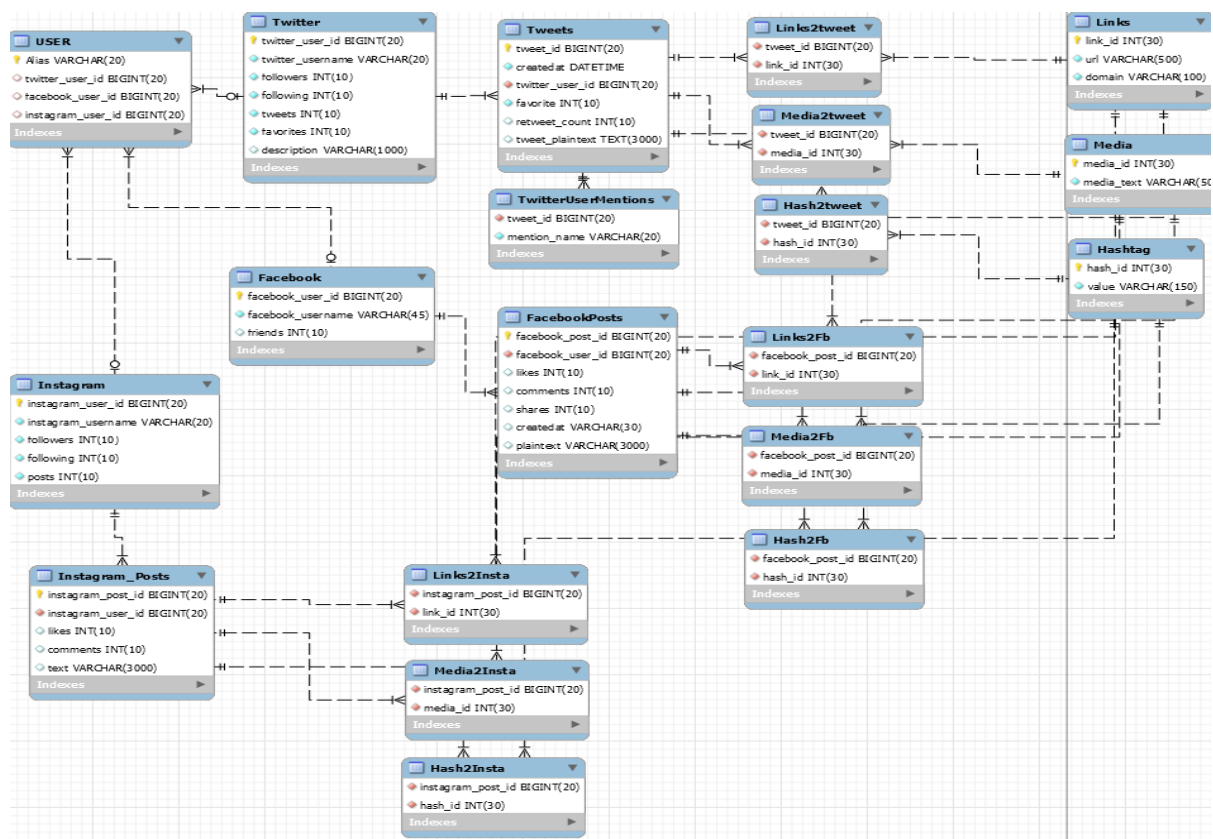


Figure 2. The Entity–Relationship Diagram of our database schema.

Quite possibly, a different database model which stores all social profiles from the three OSNs in a single table and all social posts in a separate table could have been adopted. However, having considered such a design, we identified data, auditing, and architectural issues in our specific case. As one can see, the information and available metadata of the examined OSNs are not uniform. This data heterogeneity is categorised into three types [1]: syntactic, schematic, and semantic. A different model would lead to either inconsistencies or a series of blank records/cells in the database. Moreover, the existence of distinct table sets for each OSN leads to easier data auditing and error management tasks. Finally, having distinct tables for each OSN not only facilitates the unification of the “OSN accounts” type tables to create a generic entity (“User Alias” type table) but also demonstrates the scalability of the model, allowing us to incorporate additional OSNs.

3.3. OSN Data Acquisition

To collect the required information from the three OSNs described in the ER diagram of Section 3.2, the Python programming language was employed. Specifically, as presented in Figure 3, three separate services were implemented for communicating with the APIs of the examined OSNs (Twitter, Facebook, Instagram). These services were orchestrated by a primary harvesting service and were running in parallel for the purpose of the collection of the social content. These services are an integral part of the “Processing” layer of our architecture, presented in Figure 1, as they represent the “OSN APIs” component.

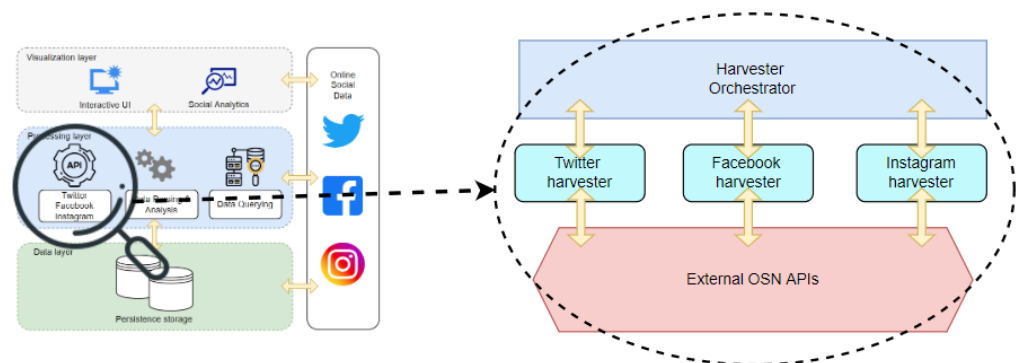


Figure 3. The orchestration of the OSN harvesters.

4. Defining Influential Metrics in OSNs

One of the research directions investigated in our study is the identification of behavioural patterns and correlations of the examined OSNs and social entities (e.g., links, hashtags, mentions) with the users' exerted social influences. However, the identification of influential entities, along with the measurement of an individual's influences in all types of networks, ranging from biological to computer influences and from corporate to OSN influences, are challenges that are applicable to a wide range of scientific areas. Despite the fact that having many direct peers, or, simply put, "friends" and "followers", is a good indication of popularity, our previous studies ([22,23]) revealed that such entities are not necessarily the most influential ones, and additional factors have to be considered. Our influence measurement relies on three social pillars: social degree, social activity, and social acceptance.

If depicted in a directed graph, the accounts of Twitter, and practically those of any OSN, would be represented by vertices, whereas their friendship or follow-up relations would be represented by directed edges. However, this structural information alone is not sufficient for defining a robust and accurate influence measurement. As already mentioned, the degree of popularity, or more formally, the number of "Followers" that a Twitter account has, does not indicate a high social impact, even though a small number of these accounts can be more influential than others. The social degree (active or passive account) must also be considered, as measured by the number of "Following" edges, indicating the accounts that a user is following. In a case where the number of "Following" is lower than the number of "Followers", then the account is considered as an active one and is more interested in authoring and disseminating new social content rather than consuming it. Thus, the "Followers to Following" (FtF) ratio is introduced. However, since this ratio may result in equal rates independent of the actual values of its components, the factor of the "Order of Magnitude" (OOM) of the "Followers" is incorporated into the described social degree measurement. The rationale is that higher values of the "Followers" social characteristic should indicate an account's greater influence. Finally, the "Tweet Creation Rate" (TCR) is introduced for calculating the social activity of the accounts, relying on the timeframe in which their most recent 100 tweets were authored. In the event of accounts with the same social degree value, then the one with the highest social activity value is the most active, thus tending to exert a stronger social influence in the OSN.

The aforementioned information is retrieved via the Twitter API, whereas each social message is accompanied by additional types of metadata, two of which are the "Retweet" and "Like" (which was "Favourite" in the past). These metrics are crucial for the well-defined measurement of social influence and social acceptance and are regarded as a means of expressing support, as well as reflecting the assessment of certain content, thus enhancing the social dynamics of real influencers. We rely on these metadata to provide a qualitative overview of an OSN account's content in terms of its impact and likeability from the perspective of the rest of the OSN by generating h-index values based on the established

scholarly domain metric [24]. Specifically, the “Retweet100 h-index” and “Like100 h-index” factors are calculated based on an account’s 100 most recent tweets.

An account with a “Retweet100 h-index” value equal to h means that “at least h tweets have been shared at least h times”. Consequently, we know that these reposting actions generated at least $h \cdot h$ new social posts, which must be attributed to the original author’s account. The normalised versions of these new messages are factored into our influence measurement, represented by the “Adjusted Tweets” factor, and this process is extensively analysed in our previous work [22].

Finally, the described Influence Metric relies on the aforementioned factors and characteristics of the OSN accounts, as defined in Equation (1). As mentioned above, any social messages derived from other accounts’ actions must be attributed to the original author’s account. Hence, the value of the “Adjusted Tweets” is added to the number of tweets. To avoid outlier values, the “FtF” ratio is placed inside a base-10 log, while this ratio is increased by one to prevent it from being zeroed in cases of equal “Followers” and “Following” values. Our Influence Metric is modular, as new factors can be easily incorporated; dynamic, as it depends on the most recent social activity by incorporating the factor of the social acceptance of the content; and configurable. According to analyses reported in [22,23], values of k equal to 100 provide the most accurate measurement of social influence.

$$\text{Influence Metric} = \frac{\text{tweets}_k + \text{AdjustedTweets}_k}{\text{Hours}_{\text{since } k_{th} \text{ tweet}}} * \text{OOM}(\text{Followers}) * \log_{10} \left(\frac{\text{Followers}}{\text{Followees}} + 1 \right) \quad (1)$$

4.1. Rating the Dataset’s Twitter Accounts

As already mentioned, our analysis focuses on entities (social individuals) maintaining accounts (social instances) in all three examined OSNs, namely Twitter, Facebook, and Instagram. In our dataset, 57 entities satisfy this criterion, and their accounts, existing in all OSNs, were discovered under semi-supervision (as in [17,21]) based on the “Name” field of their social instances, as suggested by previous literature [11]. By relying on our Influence Metric methodology ([22,23]), we quantified the degree of influence of each of these entities for the OSN of Twitter. Specifically, to perform the analysis and evaluation of Section 5.2.2, we organised the 57 examined social individuals into three groups according to their Influence Metric scores, as follows:

- “Medium” group (Table 1): Influence Metric score range [30, 47);
- “High” group (Table 2): Influence Metric score range [47, 65);
- “Very High” group (Table 3): Influence Metric score range [65, 82).

Table 1. Group of “Medium” Influence Metric score.

OSN Alias	Influence Metric Score	OSN Alias	Influence Metric Score	OSN Alias	Influence Metric Score
TurtleRock	31.26	Liverpool	41.34	Airbnb	42.01
Sport24	42.62	Dell	43.06	Oreo	43.3
Coursera	43.56	FIBA	44.19	Mega	44.34
Orange	44.39	LEVIS	44.48	Oracle	44.64
Ellinofreneia	45.07	Arduino	46.42	Wikipedia	46.52
TurtleRock	31.26	Liverpool	41.34	Airbnb	42.01

Table 2. Group of “High” Influence Metric score.

OSN Alias	Influence Metric Score	OSN Alias	Influence Metric Score	OSN Alias	Influence Metric Score
PCMag	51.61	CocaCola	52.33	Ford	52.45
JamieOliver	53.09	Nike	53.25	Kathimerini	53.75
RedBull	53.62	Skai	54.34	TimOreilly	54.36
LKing	55.07	News247	55.45	Yahoo	55.93
Dropbox	57.18	PIglesias	57.45	MLS	57.72
Bulls	58.09	Naftemporiki	59.04	BMW	59.07
Microsoft	59.38	MatteoRenzi	59.45	EUComm	61.1
Marvel	61.29	OnePlus	61.6	AmericanAir	62.14
NRJ	62.52	FTimes	63.27	HuffPost	63.3
McDonalds	63.74	BBCSport	64.16	-	-

Table 3. Group of “Very High” Influence Metric score.

OSN Alias	Influence Metric Score	OSN Alias	Influence Metric Score	OSN Alias	Influence Metric Score
Starbucks	66.3	VSecret	66.98	Samsung	71.03
CR7	71.95	Chelsea	72.06	Google	72.2
RedDevils	72.31	Barca	72.34	9GAG	72.56
WhiteHouse	73.01	Time	75.59	CNN	78.78
KatyPerry	81.78	-	-	-	-

5. Experimental Results

In this section, we present and discuss the details of our experimental methodology, along with the analyses and assessment of the experimental results of our proposed framework, presented in Section 3.

5.1. Experimental Methodology

The objective of our framework is to identify the behavioural and disseminated content patterns under the assumption that accounts maintained by users in multiple OSNs should not be regarded as distinct accounts, but rather as the same individual with multiple social instances. To this end, a predefined list of 57 social individuals were examined, who maintain accounts (social instances) in all three examined OSNs: Twitter, Facebook, and Instagram. Since we are examining the behavioural patterns of the same individuals in these OSNs, any observed consistencies or differences in their patterns are significant.

In terms of the social data acquisition process, a data-gathering service is responsible for collecting, transforming, and storing the content from the examined OSNs, including the disseminated social entities (i.e., hashtags, links, domains, multimedia content, likes, reposts, and comments) and account details, along with their social influence degrees. By combining this information, we can investigate the important research questions of (a) whether users receive social acceptance uniformly in all OSNs, and (b) the correlations of the examined OSNs and social entities with the users’ exerted social influence. To this end, we provide multiple analyses of the collected social data of the examined individuals, co-considering their influential dynamics in the OSNs. The indicative analyses include the following:

- Distribution of social entities (per OSN, per social instance, or per influence group);
- Ranking differences between social entities (per social instance);
- Overlap of social entities (per OSNs);
- Number of social entities (per post);
- Correlation between OSN activity and social influence.

5.2. Assessment of the Experimental Results

In this section, we present and discuss the experimental results of our social analysis performed on the examined social individuals, who were required to maintain active accounts in the three OSNs of Twitter, Facebook, and Instagram. Moreover, the results of our specific analyses are further evaluated by investigating the effects of the examined individuals' social influences.

5.2.1. Investigating the Behavioural Patterns of the Social Individuals

Our social analysis begins with the investigation of hashtag use. For each OSN, we selected the top 100 hashtags with the most occurrences using the collected social messages, and in Figure 4 we present their trendlines: the trendline of Twitter is black, Facebook is blue, and Instagram is purple. As we can observe, in the case of all three OSNs, a typical power law distribution appears, meaning that there are very few hashtags with a substantial number of occurrences, and the majority of them appear sparsely.

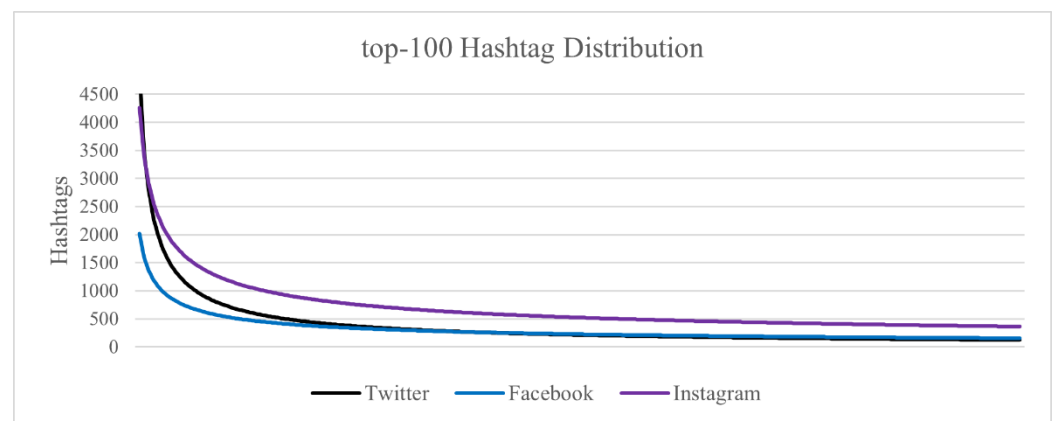


Figure 4. The distribution of the top 100 hashtags with the most occurrences in the three OSNs.

A similar power law distribution is presented in Figure 5, analysing the number of hashtags disseminated by each social instance, i.e., the OSN accounts held by the social individuals. However, these distributions do not provide any insights into the social individuals' degree of hashtag usage on their accounts in each OSN. A further analysis, presented in Figure 6, revealed that the usage of hashtags in the OSNs by the same set of individuals is not consistent. Specifically, by comparing the rankings of the instances according to the usage of hashtags in each OSN, it can be observed that, on average, there is a deviation of approximately 14 to 16 positions (in absolute numbers) in the rankings.

This finding validates our hypothesis, asserting that social individuals do not behave uniformly in all OSNs but, via their social instances, display multiple behavioural patterns. Specifically, the difference in the rankings is 15.8 positions between Twitter and Facebook, 15.2 between Twitter and Instagram, and 13.8 between Facebook and Instagram. The reduced difference in the ranking positions between Facebook and Instagram can be attributed to the fact that these OSNs belong to the same organisation (Meta) and, from the end-users' perspective, share a different philosophy compared to Twitter.

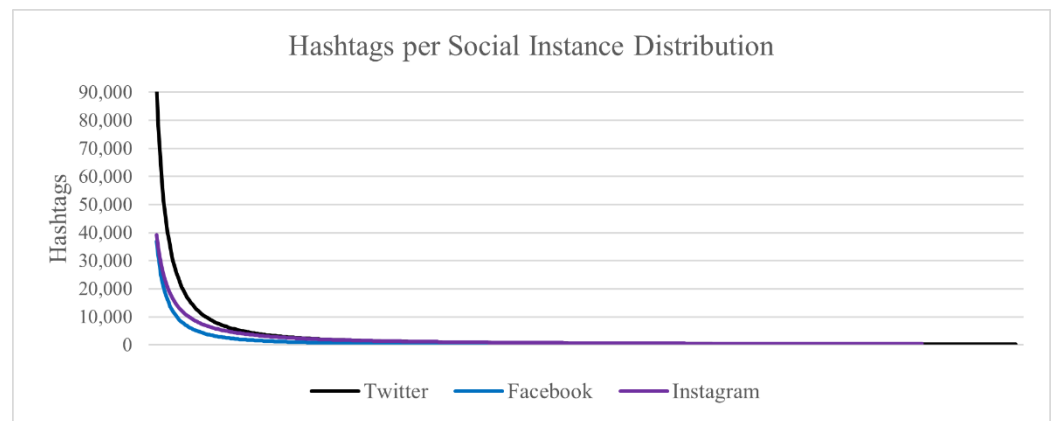


Figure 5. The distribution of hashtags disseminated in each OSN.

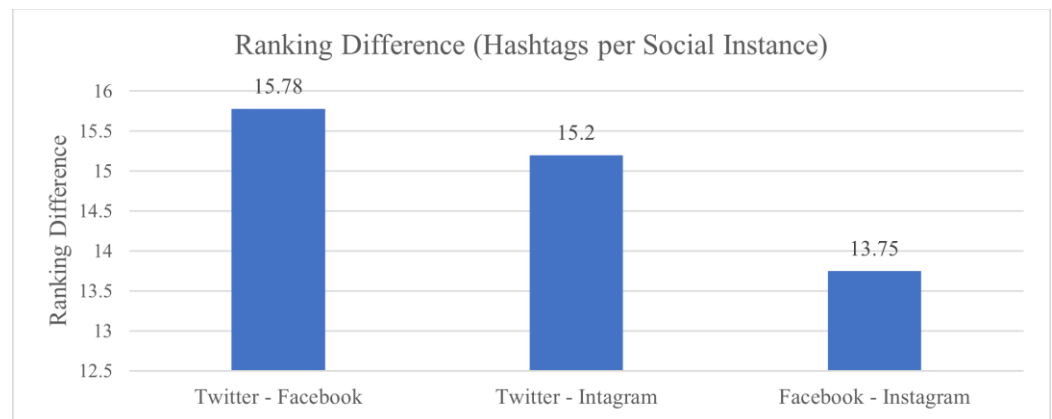


Figure 6. The difference between rankings in each OSN according to the usage of hashtags.

Considering that all hashtags were collected from the posts of the same users in different OSNs, we measured the overlap between hashtags in these OSNs. As presented in Figure 7, 61.8% of the Facebook hashtags appear on Twitter, but only 13.5% of the latter appear on Facebook. This is a direct result of the fact that a wider variety of hashtags are created and shared on Twitter, whereas the reuse of the same hashtags is more extensive on Facebook. However, it can be observed that the overlap between Facebook and Instagram hashtags is much higher (30.1% and 13.5%, respectively) compared to Twitter. As in the previous case, these two OSNs are correlated, since they belong to the same organisation (Meta) and, from the end-users' perspective, share a different philosophy compared to Twitter.

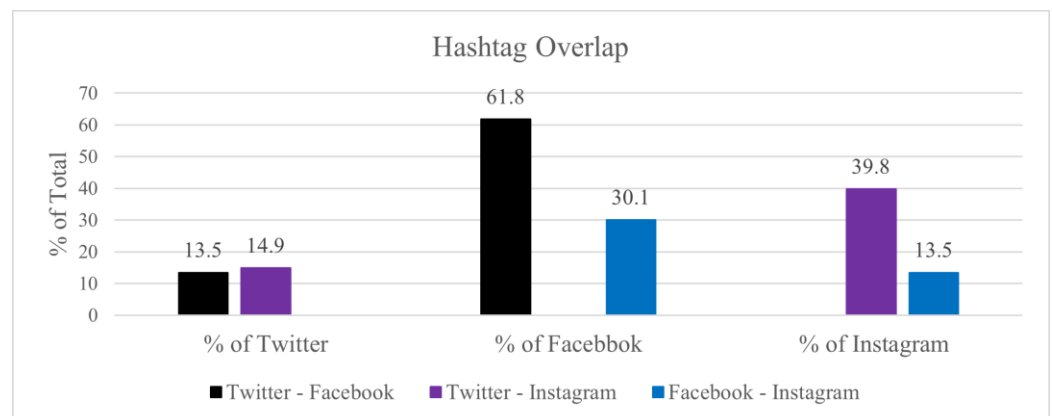


Figure 7. The hashtag overlap of the three OSNs.

Continuing our hashtag-related analysis, we identified the users associated with the top 10% of posts containing the most hashtags per OSN. Figures 8–10 present the aliases of these accounts representing at least 1%, while the rest are placed in the generic “Others” group. On Twitter, 10 distinct individuals were found, while on Facebook we found 9 and on Instagram 14. The social individual “Naftemporiki” appears in all three charts, and five individuals appear in two. Specifically, three individuals (“BMW”, “Orange”, and “EUComm”) appear on Facebook and Instagram, whereas two individuals (“Oracle” and “RedDevils”) appear on Twitter and Instagram.

No individuals common to Twitter and Facebook were discovered for this subset, thus indicating, again, the lack of correlation between these OSNs. On the contrary, Instagram seems to be more correlated with the other OSNs. Out of the 33 individuals with a share that is greater than 1%, 26 are unique. Thus, approximately 21% of the individuals contribute to more than one OSN. Moreover, in the case of Twitter and Facebook, more than 50% of the examined posts with the most hashtags were posted by individuals related to (Greek) news agencies, namely “Naftemporiki” and “Skai”, whereas on Instagram, the domain of the individuals is more diverse.

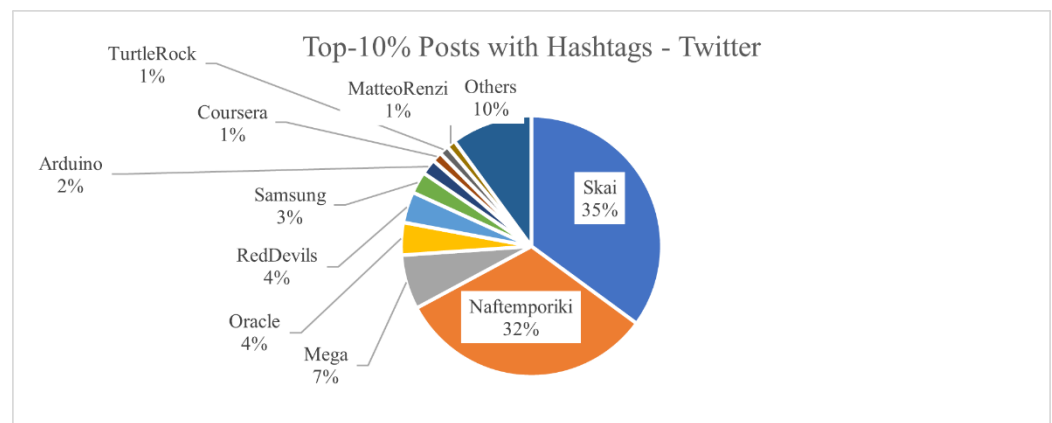


Figure 8. The distribution of the individuals who posted the top 10% of posts with the most hashtags on Twitter.

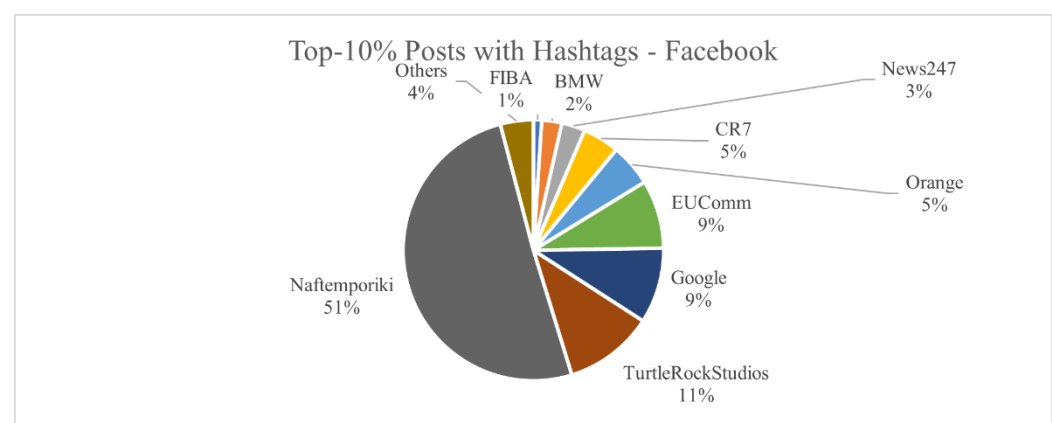


Figure 9. The distribution of the individuals who posted the top 10% of posts with the most hashtags on Facebook.

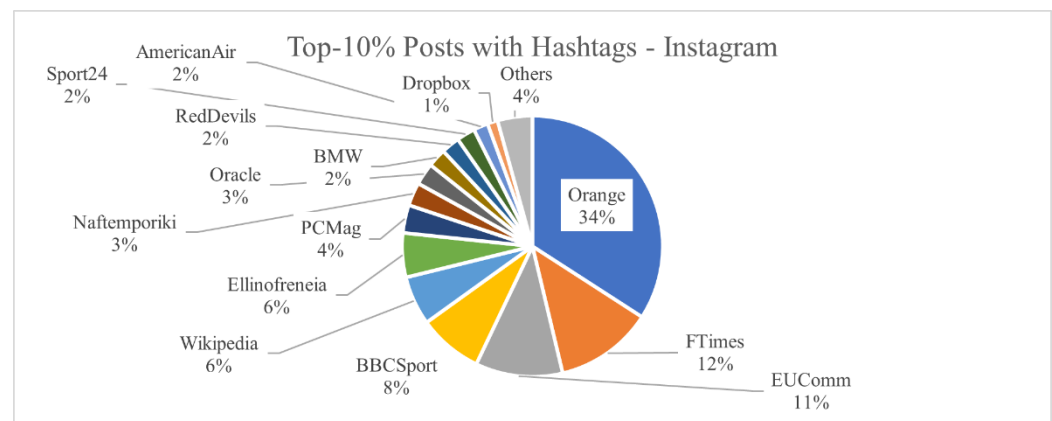


Figure 10. The distribution of the individuals who posted the top 10% of posts with the most hashtags on Instagram.

Figure 11 presents the average number of hashtags found in the posts of the three examined OSNs. We can observe that the use of hashtags on Facebook is very sparse, with only 0.2 hashtags per post, whereas the same individuals contribute almost three times more hashtags than this on Twitter. The latter finding is in line with our analysis of Figure 7. Finally, on Instagram, an emphatic increase in hashtags per post of 2.5 is observed for the same individuals, almost 4 times as much as Twitter and 12.5 times compared to Facebook.

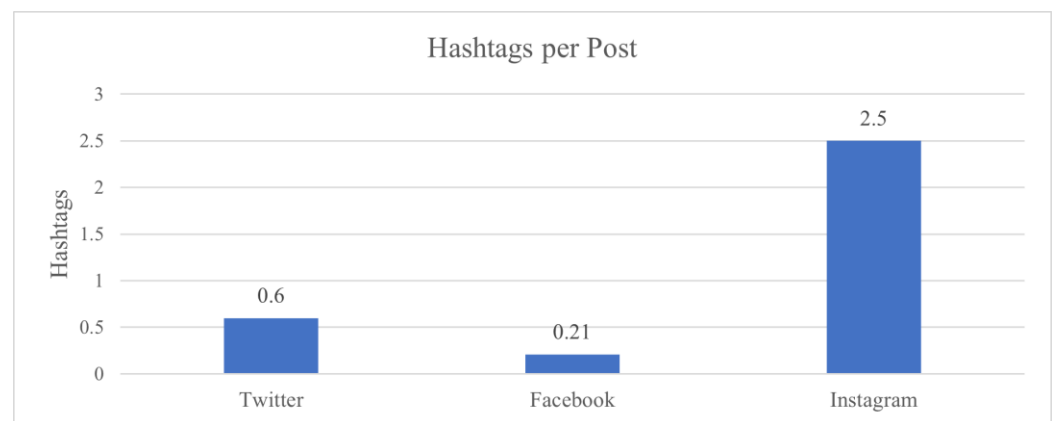


Figure 11. The distribution of hashtags per post on the three OSNs.

Our social analysis of the examined social individuals and the OSNs continues with the investigation of the use of hyperlinks. Currently, Instagram does not allow for the inclusion of hyperlinks in its posts. For the remaining OSNs, we selected the top 100 links with the most occurrences using the collected social messages, and in Figure 12 we present their trendlines: the trendline of Twitter is black and that of Facebook is blue. The vertical axis represents the occurrences of the links. As we can observe, for both OSNs, a typical power law distribution appears, meaning that there are very few links with a substantial number of occurrences, and the majority of them appear sparsely.

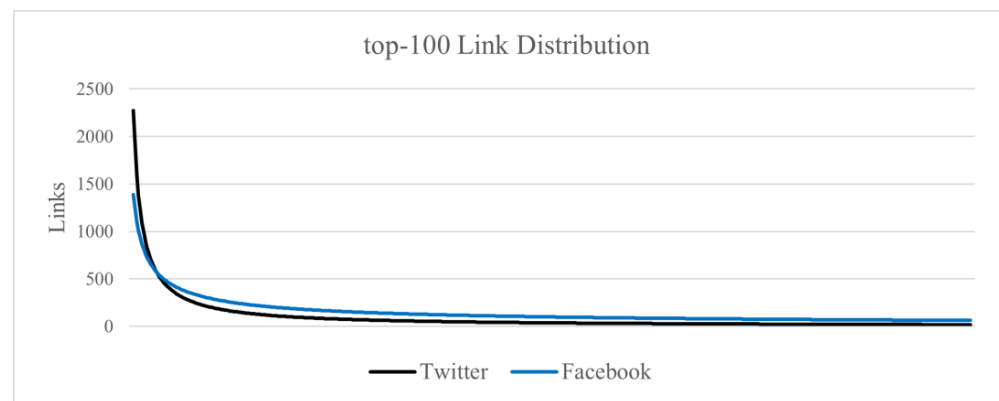


Figure 12. The distribution of the top 100 links with the most occurrences on Twitter and Facebook.

Figure 13 presents the average number of links found in the posts on Twitter and Facebook. We can observe that our examined set of social individuals use, on average, 0.73 links on Facebook posts, almost 52% higher compared to their Twitter messages. It is evident that Facebook is the preferred OSN for the dissemination of link content.

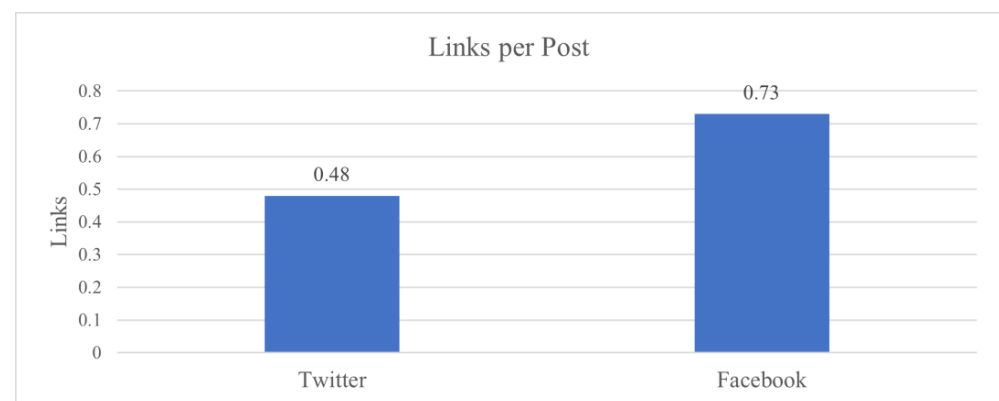


Figure 13. The distribution of links per post on Twitter and Facebook.

By further investigating the disseminated links, we extracted their domains and analysed the top 250 with the most occurrences. In Figure 14, the distribution (on the logarithmic scale) of the domains is presented, along with their trendlines: Twitter is in black and Facebook is in blue. Similar to the previous distributions, in the case of both OSNs, a typical power law distribution appears. There are very few domains with a substantial number of occurrences, and the majority of them appear sparsely.

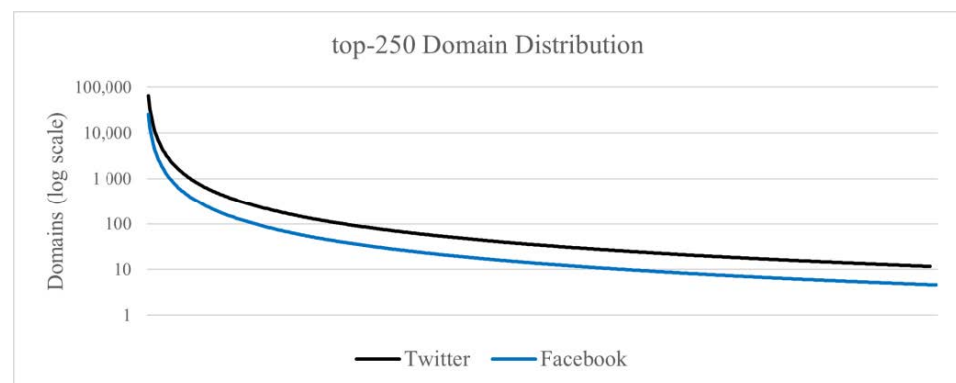


Figure 14. The distribution (on the logarithmic scale) of domains on Twitter and Facebook.

As in the case of hashtags, considering that all the links were collected from the posts of the same users on different OSNs, we measured the overlap between their domains. As presented in Figure 15, 63.4% of the Facebook domains appear on Twitter, but only 21.7% of the latter appear on Facebook. This is despite the fact that, on Twitter, less links and domains appear per post, and they are more diverse, whereas the reuse of the same domains is more extensive on Facebook.

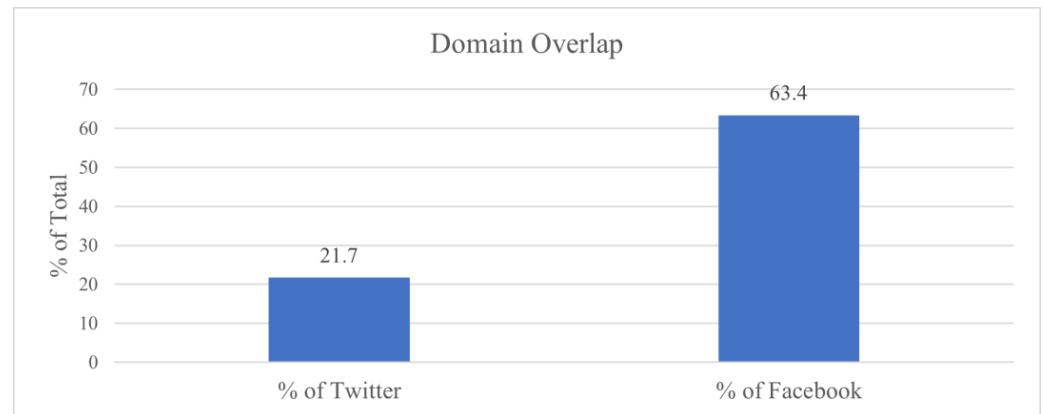


Figure 15. The domain overlap between Twitter and Facebook.

The rise in multimedia content, such as images and videos, has rapidly increased in the exchanged social messages [14,25] on all OSNs. Thus, we also analysed this type of content. Figure 16 presents the average number of multimedia content types found in the posts of the three examined OSNs. Twitter API characterises this type of content as “media”, a term we also used for reference purposes. As we can observe, the use of media on Twitter is very sparse, with only 0.31 media per post, whereas the same individuals share almost 60% more media on Facebook. Finally, on Instagram, an emphatic increase in the number of media per post by the same individuals of 1.64 can be noticed, which is more than five times that of Twitter and more than three times that of Facebook. This metric validates the common assumption that Instagram is an OSN predominantly used for disseminating multimedia content.

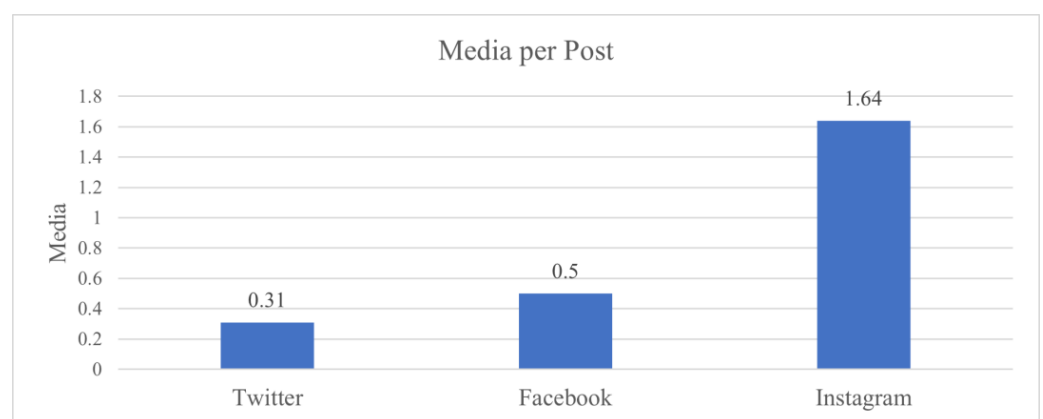


Figure 16. The distribution of media per post on the three OSNs.

As in the previous cases, for each OSN, we selected the top 100 media (represented by their URLs) with the most occurrences using the collected social messages, and in Figure 17 we present their trendlines: black for the trendline of Twitter and blue for Facebook. Instagram does not appear in this graph, as there is no single instance of multimedia content with more than one occurrence. Instagram generates new URLs for each reposted image or video. As we can observe, in the case of both Twitter and Facebook, a typical

power law distribution appears, meaning that there are very few media files being shared a substantial number of times, and the majority of them appear sparsely.

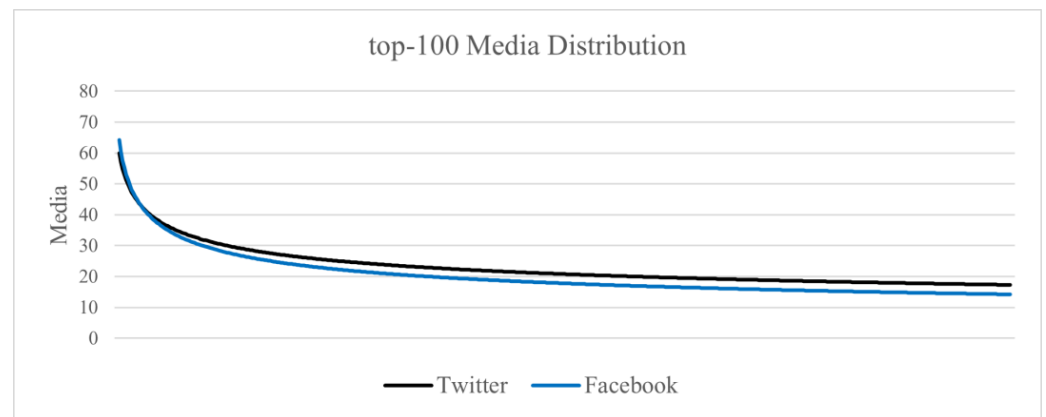


Figure 17. The distribution of multimedia content on Twitter and Facebook.

As already mentioned, social acceptance is an important characteristic of the disseminated messages, measured by the number of received Likes and Reposts. However, we must consider an important research question, namely: are users receiving this acceptance uniformly in all OSNs? Figure 18 presents the received average number of likes of the posts created by the same social individuals in the three examined OSNs. On Instagram, this number is overwhelming compared to those of Facebook and Twitter, being greater by one and two orders of magnitude, respectively. This analysis portrays Instagram as enabling the reception of positive feedback considerably more effectively compared to the other platforms.

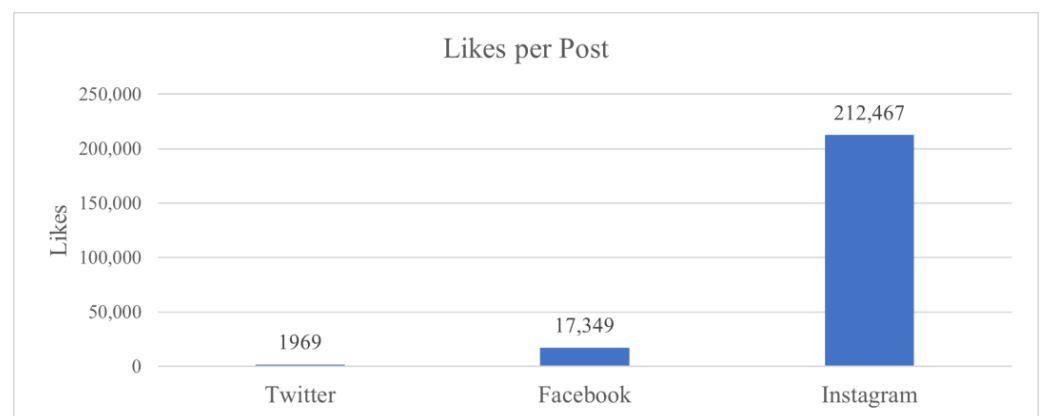


Figure 18. The distribution of likes per post on the three OSNs.

Reposting actions are the second means of identifying social acceptance. Similar to the previous figures, Figure 19 presents the average number of reposts of the messages created by the same social individuals in the three examined OSNs. Contrary to the case of the “Likes”, we can observe that this action on Facebook is sparse, with 58 reposts per message, whereas the messages of the same individuals are reposted at a rate of more than six times this number on Twitter, specifically 385 times. Currently, Instagram does not offer the ability to reshare posts.

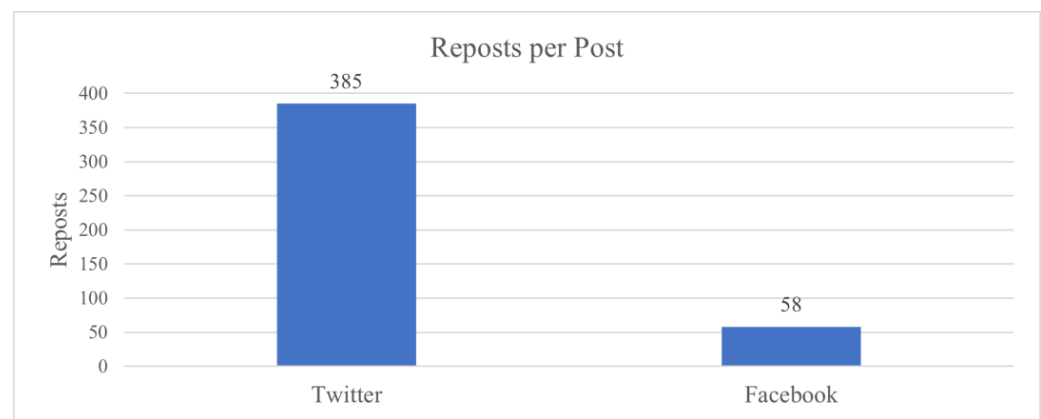


Figure 19. The distribution of reposts per post on Twitter and Facebook.

Our final analysis involves the ability of the examined social individuals to engage other users in public conversations through a social message. Figure 20 presents the average number of comments found in the posts of the three examined OSNs. We can observe that the Twitter engagement of users via the posting of comments (or “Replies”, as named by this platform) is sparse, with 369 comments per post, whereas the same individuals produce two times more comments on Facebook, with 856 per original post. Finally, on Instagram, an emphatic increase in the number of comments per post of 1,848 can be observed for the same individuals, being five times as much as that of Twitter and more than two times that of Facebook. This comparison clearly indicates (a) the tendency of users to react to Instagram posts rather than Twitter or Facebook posts and (b) that Instagram seems to be the platform most often used for engaging other users in public discussions.

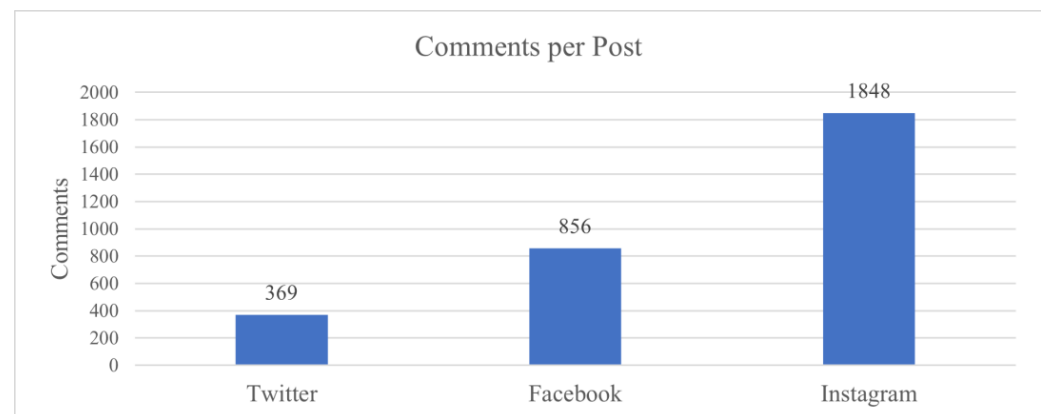


Figure 20. The distribution of comments per post on the three OSNs.

5.2.2. Investigating the Behavioural Patterns of the OSN Influence Groups

Here, we further evaluate the results of specific analyses described in the previous section by investigating the effects of the examined individuals’ social influences. As presented in Section 4.1, three social influence groups were created, including those of “Medium”, “High”, and “Very High” influence.

Our first analysis considers the influence factor according to the average number of hashtags per post for each OSN. As presented in Figure 21, we use one group of bars for all users (represented by the “All” label) and three for each influence group (represented by the “Medium”, “High”, and “Very High” labels). Twitter is represented by the black bar, Facebook by the blue bar, and Instagram by the purple bar. We can observe that the use of hashtags by each social group is similar to their combined distribution (“All”). Moreover, it is revealed that the social instances of the “Medium” influence group share on average 3.27 hashtags per post on Instagram, 32% more than the corresponding value of the “High”

group and 80% more than the “Very High” group. The “High” group uses 0.67 hashtags per Twitter post, which is 24% more than the “Medium” group and 36% more than the “Very High” group. The “Very High” group tends to hardly ever use hashtags, especially on Instagram.

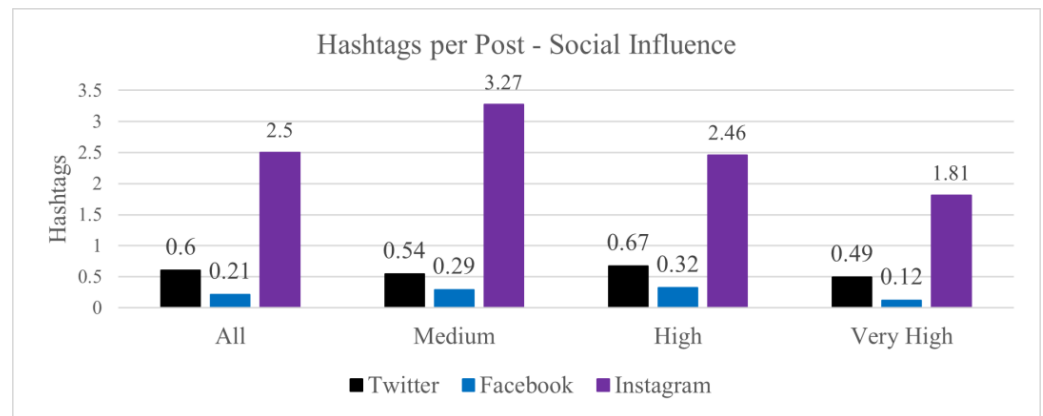


Figure 21. The distribution of hashtags per post on the three OSNs per social influence group.

For the hashtag analysis, we relied on the top 500 posts with the most hashtags in each OSN. Specifically, on the one hand, we calculated the average number of hashtags per post, while on the other hand, we calculated the average number of unique hashtags in these posts. By observing one occurrence of each hashtag per post, we aimed to investigate the tendency of the social groups to abuse hashtag use. Figures 22–24, on the vertical axis, present the average number of hashtags per post, while on the horizontal axis, the four categories of all users, along with the three influence groups of Twitter, Facebook, and Instagram, respectively, are depicted. The black, blue, and purple bars of Figures 22–24 represent the average number of possibly repeated hashtags, whereas the orange bar represents the unique number of hashtags per post. We can observe that, on Twitter (Figure 22) and Facebook (Figure 23), no major differences can be observed between groups. However, on Instagram (Figure 24), significant differences can be observed. Specifically, 2.5% of the hashtags disseminated by the “High” group appeared more than once in a single post, as do 3.2% of the “Medium” group’s hashtags. This pattern is extensively observed in the “Very High” group, where almost 7% of the hashtags appeared multiple times in a single post. This demonstrates that Instagram is not only the OSN with the most extensive use of hashtags, but also the group in which hashtags are more likely to appear more than once in the posts.

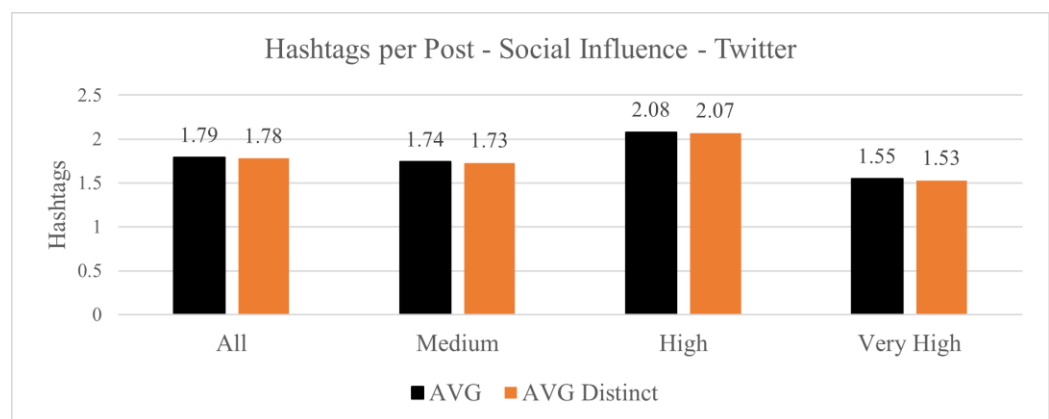


Figure 22. The distribution of the total and distinct hashtags per post on Twitter per social influence group.

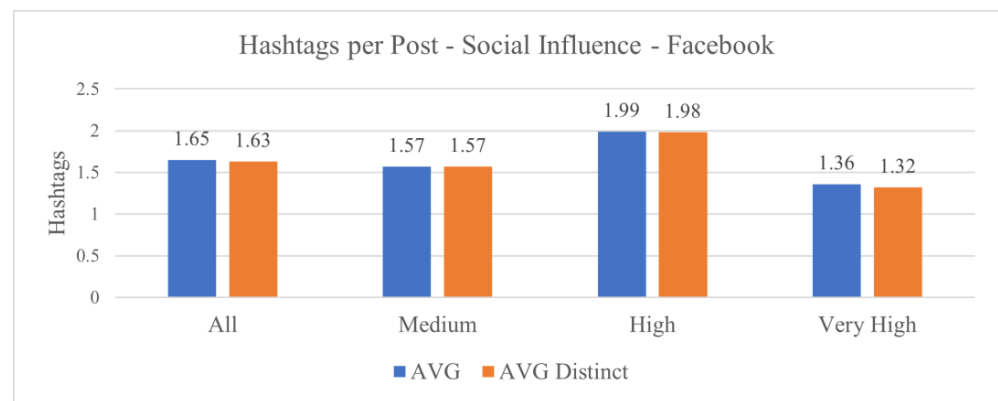


Figure 23. The distribution of the total and distinct hashtags per post on Facebook per social influence group.

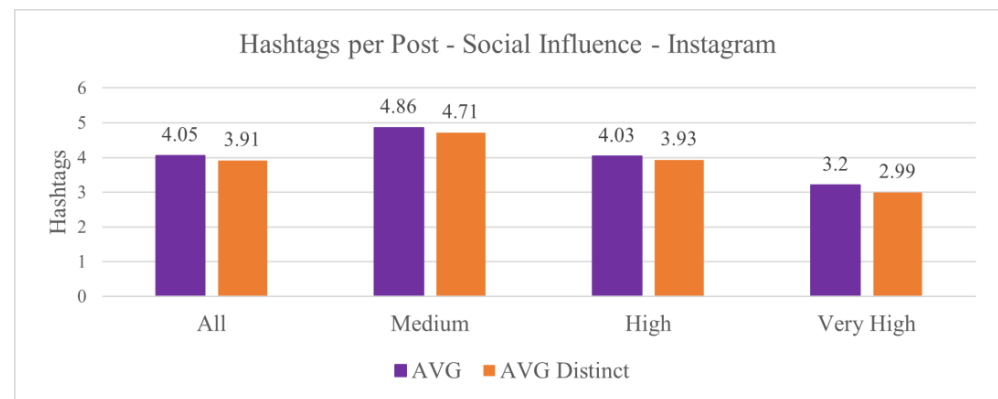


Figure 24. The distribution of the total and distinct hashtags per post on Instagram per social influence group.

The next examined social entity are hyperlinks. Figure 25 presents the average number of links found in posts on both Twitter and Facebook for all social individuals combined, as well as those for each influence group. Instagram is not presented, as hyperlinks are not currently allowed in the posts. We can observe that the use of links by each social group is similar to their combined distribution (represented by the “All” label). Moreover, it is revealed that the “Medium” and “High” groups include, on average, 36% more links per post on Facebook compared to Twitter, whereas the “Very High” group includes 88%. It is evident that the “influencers”, namely the “Very High” group, prefer Facebook for the dissemination of such content.

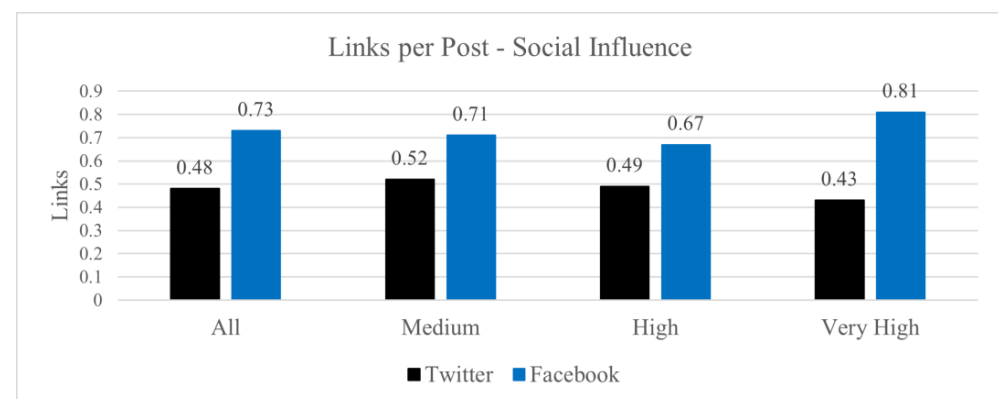


Figure 25. The distribution of links per post on Twitter and Facebook per social influence group.

The next examined entity is multimedia (media) content. In Figure 26, we present the average number of multimedia content posts on the three examined OSNs for all social individuals combined, as well as for each influence group. As we can observe, all the values show exceedingly small deviations between the influence groups.

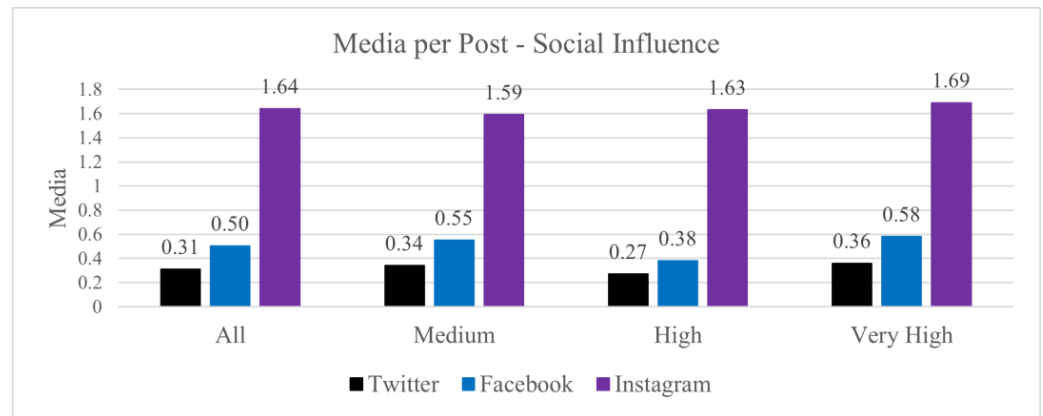


Figure 26. The distribution of media per post on the three OSNs per social influence group.

As already mentioned, social acceptance is an important characteristic of OSNs, measured by the number of received Likes and Reposts. In order to extend the analysis of the previous section, we measured the average number of likes that the posts received in the three examined OSNs for all the social individuals combined, as well as for each individual influence group. This allowed us to identify whether the influence groups are accepted uniformly in each OSN.

As we can see in Figure 27, a higher degree of influence leads to a higher number of received likes. Moreover, it is evident that, on Instagram, the number of received likes is substantially increased compared to the other OSNs, as the degree of influence is also increased. On top of that, Twitter and Facebook show no major differences between the “Medium” and “High” groups, as opposed to the “Very High” group, whose values are increased by one order of magnitude.

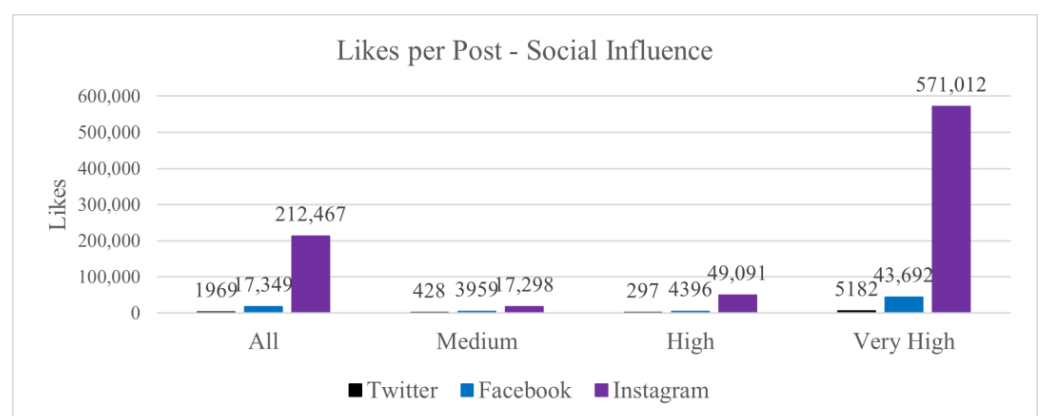


Figure 27. The distribution of likes per post on the three OSNs per social influence group.

The second aspect of social acceptance is the number of times a message has been reposted in the OSNs. We measure the average number of reposts that the posts received on Twitter and Facebook for all the social individuals combined, as well as for each individual influence group. Currently, Instagram does not offer the ability to reshare any posts; thus, it is absent from this comparison. As in the case of likes, a higher degree of influence leads to a higher number of reshared social messages, as presented in Figure 28. Moreover, on Twitter, the number of reposts is substantially increased compared to Facebook, as the

degree of influence is also increased. This indicates that the more social influence a social individual exerts on its direct and indirect peers, the more reposts they will receive.

Considering the conclusions drawn from Figures 27 and 28, we can observe that the influence groups do not receive a uniform social acceptance in all OSNs. Furthermore, this acceptance is affected by, and closely related to, the social individuals' exerted influences.

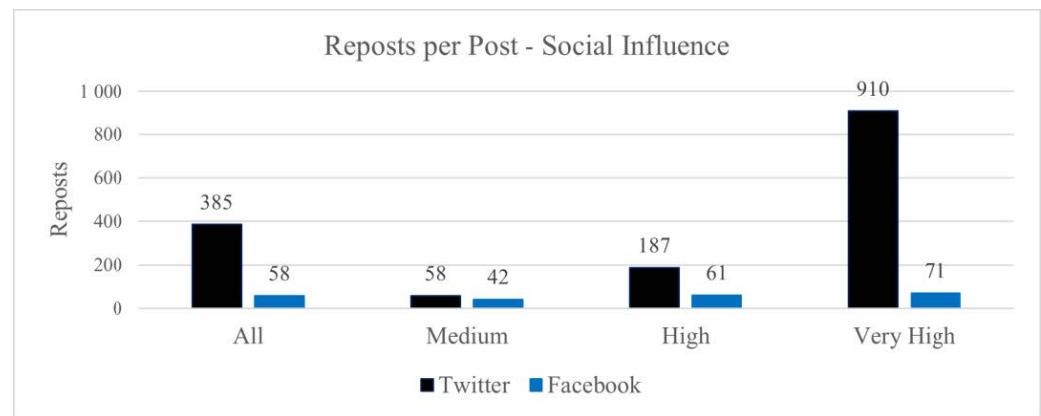


Figure 28. The distribution of reposts per post on the three OSNs per social influence group.

The next examined social aspect is the ability of asocial individuals to engage other users in public conversations, stimulated by a specific social message. In Figure 29, we present the average number of comments per post on the three examined OSNs for all social individuals combined, as well as for each influence group. As we can observe, the social instances of the “Medium” group receive more comments on Twitter compared to Facebook (+21%) and Instagram (+34%). However, the groups of higher social influence receive a substantial number of increased comments on Facebook and Instagram by up to six times. This comparison clearly indicates that (a) the non-influencers participate in public discussions more often on Twitter rather than the other OSNs, (b) higher social influence leads to a greater number of public discussions, and (c) Instagram seems to be the best platform for engaging other users in public discussions.

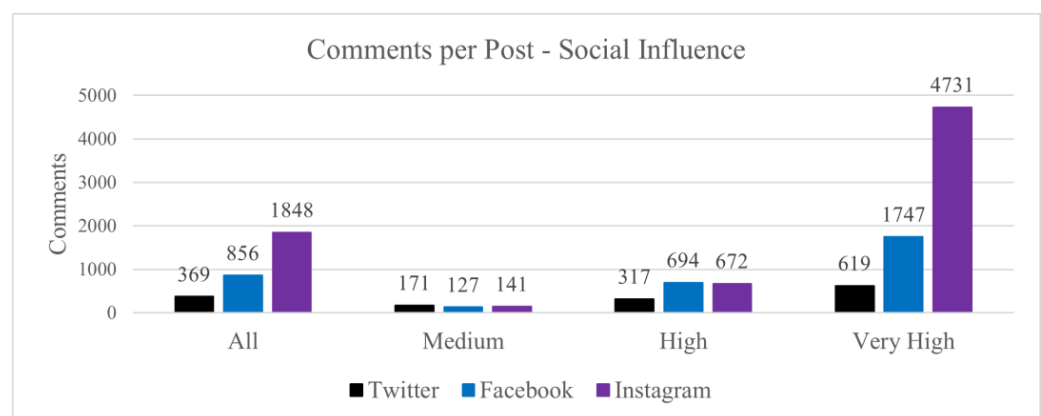


Figure 29. The distribution of comments per post on the three OSNs per social influence group.

In Section 4, we introduced social activity as a factor contributing to a social account's influence. To analyse the examined social individuals' activity uniformly in the three OSNs, we calculated their “OSN Activity” values. Specifically, we relied on the dates of the most recent and older disseminated social messages, as well as the number of social messages posted between these dates. Figure 30 presents the distribution of the Twitter (in black) and Facebook (in blue) “OSN Activity” values. Unfortunately, due to an issue with the collected data, we were not able to calculate this metric for Instagram. It can be observed that both

OSNs follow a typical power law distribution, meaning that there are very few social instances that are considerably active, and the majority of them show a minimal activity.

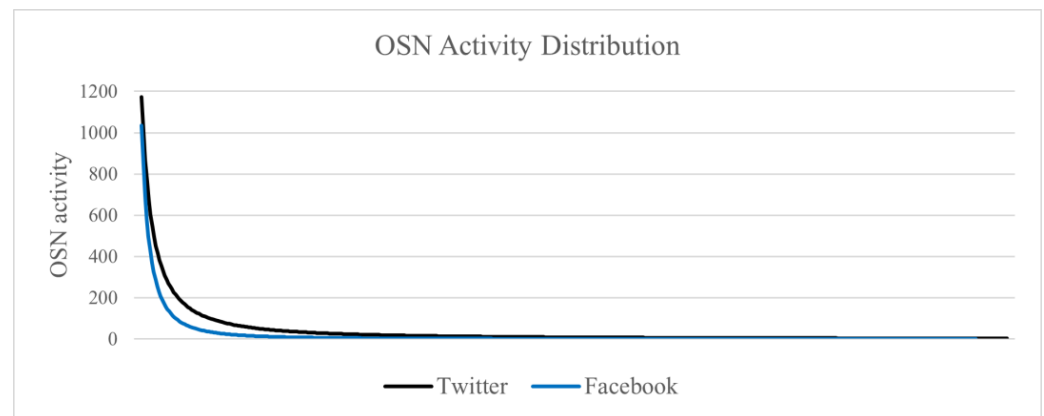


Figure 30. The distribution of social activity values on Twitter and Facebook.

However, these distributions do not provide any insights into the social individuals' activity in terms of the exerted influence. A further analysis, presented in Figure 31, reveals that the social activity of the influence groups is not consistent across Twitter and Facebook. Specifically, by comparing the rankings of the instances in each group, it can be observed that, on average, there is a deviation of approximately three positions (in absolute numbers) in the rankings. This finding means that social individuals do not behave uniformly in all OSNs but, via their social instances, display multiple distinct behavioural patterns. Specifically, the ranking difference of the “Medium” group is 11.7 positions, and it is 7.4 for the “High” group and 13.1 for the “Ver High” group. These rankings indicate that the “High” influence group displays greater consistency in its activity across these two OSNs compared to the other groups. On the contrary, the “Very High” group displays the highest deviation of the rankings, leaning towards one of these OSNs.

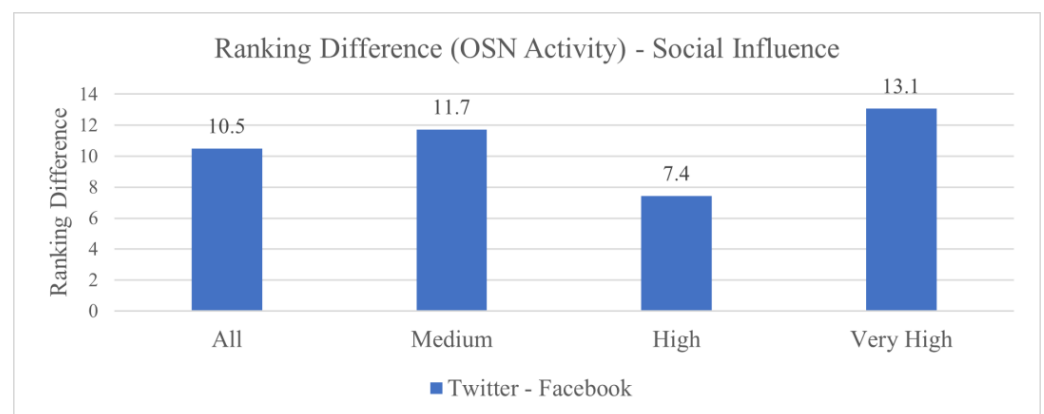


Figure 31. The difference between rankings on Twitter and Facebook according to the social activity per social influence group.

6. Identifying Social-Influence-Stimulated Behavioural Patterns in OSNs

In this section, we further investigate the outcomes of our analyses in Section 5 and present the rankings of the examined OSNs according to the extent of usage of three types of social characteristics: (a) social entities, including the number of hashtags, links, and multimedia content (photos and videos); (b) social acceptance, including the number of likes and reposts; and (c) social conversation, including the number of comments and replies to original posts. These metrics reflect the behavioural patterns of the same social individuals (users) in the different OSNs. Furthermore, we categorise the examined individuals into

three groups based on their social influence, as presented in Section 4.1, to further analyse the usage of these social characteristic types and to rank the influence groups in each OSN.

The three types of social characteristics consist of six metrics, analysed in Section 5.2.1, and are presented in the first two columns of Table 4. These metrics are associated with their respective figures, where further details can be found. The remaining three columns represent the examined OSNs, namely Twitter, Facebook, and Instagram. For each of these networks, we ranked the usage of the six metrics from higher to lower, as “1st”, “2nd”, and “3rd”. In the case where an OSN metric could not be ranked, due to its not being offered by the platform, the value “N/A” was assigned. To calculate the rankings in Table 4, all the examined social individuals, along with their content in the three OSNs, were analysed without any further categorization or filtering.

Table 4. Ranking of the OSNs according to the usage of social characteristics.

Social Characteristic Type	Metric (Average Number of)	OSN		
		Twitter	Facebook	Instagram
Entities	Hashtags per Post (Figure 11)	2nd	3rd	1st
	Links per Post (Figure 13)	2nd	1st	N/A
	Media per Post (Figure 16)	3rd	2nd	1st
Acceptance	Likes per Post (Figure 18)	3rd	2nd	1st
	Reposts per Post (Figure 19)	1st	2nd	N/A
Conversation	Comments per Post (Figure 20)	3rd	2nd	1st
Ranking Points		4	6	6

As we can observe, Instagram prevailed in all four metrics available on the platform. Currently, Instagram does not allow for the inclusion of hyperlinks in posts or offer the ability to reshare any posts; thus, the metrics related to these aspects cannot be calculated (denoted by the “N/A” value). In our study, this was interpreted as a missed opportunity for content dissemination on the platform’s part. In order to uniformly quantify the rankings of these metrics and facilitate the ranking process, we assigned two points for every first ranking, one point for every second ranking, and zero points for every third ranking, while one point was subtracted for each “N/A” value. According to the six rankings in Table 4, Instagram and Facebook are ranked first, with six points, despite Instagram not participating in the two metrics, whereas Twitter is ranked in the last position, with four points. Therefore, our analysis suggests that Instagram is the OSN with the most extensive use of social entities and conversations, as it prevailed in all four metrics available on the platform despite not participating in two metrics, which decrease its score, while Twitter does not seem to be preferred over the other two platforms by the same users.

Continuing our analysis, in Tables 5–7, we present the same six social metrics for each of the examined OSNs, considering the influence groups of the examined social individuals, as analysed in Section 5.2.2. These metrics are also associated with their respective figures, where further details can be found. Specifically, in Table 5, the metrics calculated for the Twitter content are presented along with the rankings of the three influence groups. By following the same scoring rationale used previously, we can see that, on Twitter, the group of “Very High” influence is ranked first, with eight points, while the groups of “Medium” and “High” influence follow with five points.

We can also observe that the “influencers” (“Very High” group) tend to comment on or respond to their Twitter posts and generally stimulate public conversations, in contrast

to the least influential group. Moreover, as discussed in Section 4, social acceptance is a factor affecting a user's exerted influence and is measured by the number of received likes and reposts. In Table 5, the “influencers” are ranked in the first position for this social characteristic type. On the contrary, the least influential users tend to disseminate more hashtags and hyperlinks in their social messages compared to the “influencers”, who prefer the inclusion of multimedia content.

Table 5. Ranking of the Twitter metrics per social influence group.

Social Characteristic Type	Metric (Average Number of)	Twitter Social Influence Group		
		Medium	High	Very High
Entities	Hashtags per Post (Figure 21)	2nd	1st	3rd
	Links per Post (Figure 25)	1st	2nd	3rd
	Media per Post (Figure 26)	2nd	3rd	1st
Acceptance	Likes per Post (Figure 27)	2nd	3rd	1st
	Reposts per Post (Figure 28)	3rd	2nd	1st
Conversation	Comments per Post (Figure 29)	3rd	2nd	1st
Ranking Points		5	5	8

Similarly, Table 6 presents the metrics of the Facebook content, along with the rankings of the three influence groups. Once again, the group of “Very High” influence is ranked first, with 10 points, whereas the groups of “High” and “Medium” influence are placed in the second and third positions, with five and three points, respectively. As in the case of Twitter, on Facebook the “influencers” can also stimulate public conversations initiated by their posts, contrary to the least influential group. This can be attributed to the fact that the latter group does not prefer Facebook for publicly participating in conversations, since a separate private chatting service is offered by the platform. On the other hand, the “influencers” seem to be more engaged with their audiences on Facebook. The social acceptance factor seems to be in line with the influence degree, as the least influential group is ranked in the last position for this social characteristic type. Moreover, we can observe a common pattern in the “Hashtags per Post” and “Media per Post” metrics on Twitter and Facebook.

Table 6. Ranking of the Facebook metrics per social influence group.

Social Characteristic Type	Metric (Average Number of)	Facebook Social Influence Group		
		Medium	High	Very High
Entities	Hashtags per Post (Figure 21)	2nd	1st	3rd
	Links per Post (Figure 25)	2nd	3rd	1st
	Media per Post (Figure 26)	2nd	3rd	1st
Acceptance	Likes per Post (Figure 27)	3rd	2nd	1st
	Reposts per Post (Figure 28)	3rd	2nd	1st
Conversation	Comments per Post (Figure 29)	3rd	2nd	1st
Ranking Points		3	5	10

Finally, Table 7 presents the metrics of the Instagram content along with the rankings of the three influence groups. Once again, the group of “Very High” influence is ranked first, with six points, the group of “High” influence follows with four points, and the group of “Medium” influence comes last, with two points. Instagram does not allow for the inclusion of hyperlinks in posts or offer the ability to reshare any posts; thus, the related metrics cannot be calculated (“N/A” value). In this case, no points are subtracted, since the comparison is performed using the same platform.

On Instagram, the social acceptance factor seems to be in line with the influence degree, as the “influencers” are ranked in the first position for this social characteristic type, whereas the least influential users appear in the last position. As in the case of Facebook, the “influencers” can stimulate public conversations, contrary to the least influential group, since the former seem to be more engaged with their audiences in this OSN as well. Moreover, the least influential group uses more “Hashtags per Post” compared to not only the other social groups but also the other two examined OSNs. Finally, on all three OSNs, the group of “Very High” influence disseminates more multimedia content compared to the other groups. Among the others, the use of this content type can be attributed to promotional purposes and marketing campaigns.

Table 7. Ranking of the Instagram metrics per social influence group.

Social Characteristic Type	Metric (Average Number of)	Instagram Social Influence Group		
		Medium	High	Very High
Entities	Hashtags per Post (Figure 21)	1st	2nd	3rd
	Links per Post	N/A	N/A	N/A
	Media per Post (Figure 26)	3rd	2nd	1st
Acceptance	Likes per Post (Figure 27)	3rd	2nd	1st
	Reposts per Post	N/A	N/A	N/A
Conversation	Comments per Post (Figure 29)	3rd	2nd	1st
Ranking Points		2	4	6

In summary, there is a pattern in the rankings of these influence groups. The group of “Very High” influence dominates in all the OSNs, the group of “High” influence is placed in the second position (being ranked in the second position on Facebook and Instagram), and the group of “Medium” influence is ranked third (collecting less points in total across all OSNs). As discussed in Section 4, social acceptance is a factor affecting a user’s exertion of influence and, as we can see in Tables 5–7, the “influencers” are ranked in the first position for this social characteristic type. This group, unsurprisingly, seem to attract the social attention and acceptance of others and stimulate social conversations via their posts, while they also tend to avoid the use of hashtags compared to the other two influence groups.

Our conclusions are summarised in Table 8, presenting an overview of the values of the six OSN metrics, as derived from the majority vote of their values in Tables 5–7. In a “majority vote” system, a value is assigned in a case where more than half of the votes are cast. In our case, the number of the votes is three, including one for each examined OSN, whereas the actual values of the votes are the ranking positions. The group of “Very High” influence is ranked first, with eight points, the group of “High” influence follows with five points, and the group of “Medium” influence comes last, with two points. The majority vote value of the “Links per Post” metric for the three groups cannot be calculated, since a different ranking was assigned to each group for the examined OSNs.

Table 8. Majority voting ranking of the OSN metrics per social influence group.

Social Characteristic Type	Metric (Average Number of)	Social Influence Group		
		Medium	High	Very High
Entities	Hashtags per Post	2nd	1st	3rd
	Links per Post	N/A	N/A	N/A
	Media per Post	2nd	3rd	1st
Acceptance	Likes per Post	3rd	2nd	1st
	Reposts per Post	3rd	2nd	1st
Conversation	Comments per Post	3rd	2nd	1st
Ranking Points		2	5	8

7. Conclusions and Future Work

In this study, we proposed a framework relying on a layered and expandable architecture (Section 3.1) for the identification of the OSN users' behavioural and disseminated content patterns under the assumption that accounts maintained by users in multiple OSNs should not be regarded as distinct accounts, but rather as the same individuals with multiple social instances. A predefined list including social individuals were examined, who maintained accounts (social instances) in the three examined OSNs, namely Twitter, Facebook, and Instagram. Our dataset consists of 57 social individuals who satisfied this criterion, and their accounts in the three OSNs were identified under semi-supervision. Thus, any observed consistencies or differences in the users' behavioural and content patterns in the examined OSNs are significant, since the conclusions are drawn from an analysis of the same group of individual rather than unrelated or random accounts. Indicative conclusions drawn from our analyses of Section 5.2.1 are presented below:

- Social individuals do not behave uniformly in all OSNs but, via their social instances, display multiple behavioural patterns.
- In the case of all three OSNs, a typical power law distributions occurs in regard to the dissemination of the social entities of hashtags, links, domains, and multimedia content. This means that there are very few such entities with a substantial number of occurrences, and most of them appear very sparsely.
- On Twitter, a wider variety of hashtags are created and shared, whereas on Facebook, the reuse of the same hashtags is more frequent.
- The use of hashtags on Facebook posts is very low per user, while the same user includes on average almost three times more hashtags on Twitter and 12 times more on Instagram.
- The users react more to Instagram posts via engaging in public conversations or marking them with a "Like" compared to Twitter or Facebook posts.
- Facebook posts include more hyperlinks compared to Twitter posts authored by the same individuals.
- The use of multimedia content in Twitter posts is very low, whereas the same individuals include on average almost two times more such content on Facebook and five times more on Instagram.
- Regarding social acceptance, on Instagram it is considerably easier to receive positive feedback through "Likes" compared to Facebook (placed second) and Twitter (placed third). However, the same social individuals' posts are reposted more than six times more frequently on Twitter compared to Facebook. Instagram does not support reposting.

The conclusions drawn from our influence-related analyses of the "Medium", "High", and "Very High" groups in Section 5.2.2 are presented below:

- The influence groups do not receive uniform social acceptance in all three OSNs, and this acceptance is affected by, and closely related to, the social individuals' exerted influence.

- The non-influencers participate in public discussions more often on Twitter rather than the other two OSNs.
- Higher social influence leads to the stimulation of a greater number of public conversations.
- The “influencers” display a more consistent usage of a single OSN compared to the least influential groups.
- The “influencers” receive more social acceptance and tend to avoid the use of hashtags compared to the other influence groups.

Our three contributions are presented as follows. Firstly, we proposed a framework for collecting and storing OSN users’ disseminated content from multiple social platforms. Our social analysis, enriched with information about the examined users’ social influences, revealed the existence of behavioural and content dissemination patterns depending on the examined OSN, its social entities (e.g., hashtags, multimedia content, hyperlinks), and the users’ exerted influences. Secondly, the results of our analyses confirmed our research hypothesis, asserting that an abundance of behavioural and disseminated content patterns are displayed in each OSN, even by the same actors. Therefore, accounts maintained by users in multiple OSNs should not be regarded as distinct accounts but rather as the same social individuals with multiple social instances, who frequently behave differently on each platform. Thirdly, we ranked the examined OSNs and the derived social influence groups based on three types of social characteristics, namely social entities, social acceptance, and social conversation. This ranking revealed correlations between the behavioural and content patterns of the same social individuals (users), the social influence degree, the appearance of social entities, and the OSNs themselves.

Compared to the related literature, our study differs in two important aspects. Firstly, we did not perform the social analysis using accounts that were randomly selected from multiple OSNs but examined the behavioural patterns of the same individuals across three OSNs, where their social instances are relayed from one OSN to another (Section 5.2.1). These expanded user profiles allowed us to acquire an in-depth perspective across multiple platforms and more effectively capture interests, characteristics, behavioural and content dissemination patterns. This information is crucial for social metrics, profiling, and recommendation systems. Secondly, our proposed social analysis framework utilises the users’ exerted social influences (Section 5.2.2) to reveal correlations between, and patterns among, the users, the interactions, the disseminated entities, the social acceptance, and the OSNs themselves in a way that has not been reported before.

Going forward, our plan is to extend this study and expand its scope by incorporating additional OSNs, such as YouTube and LinkedIn. Due to their diverse nature, by focusing on videos and business profiles, new types of information can be acquired, further enriching the results of this study. Moreover, inspired by [17], we will aim to further improve our Influence Metric ([22,23]) and apply it across all examined OSNs, considering our main finding that social individuals display different behavioural patterns via their social platform identities and do not receive uniform social acceptance in each OSN.

Author Contributions: Conceptualization, G.R.; methodology, G.R., S.G. and G.H.; software, S.G.; validation, G.R., G.H. and I.A.; formal analysis, G.R.; investigation, G.R.; resources, S.G.; data curation, G.R. and S.G.; writing—original draft preparation, G.R.; writing—review and editing, G.H. and I.A.; visualization, G.R. and S.G.; supervision, I.A.; project administration, G.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is contained within the article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Breslin, J.G.; Decker, S.; Harth, A.; Bojars, U. SIOC: An approach to connect web-based communities. *Int. J. Web Based Communities IJWBC* **2006**, *2*, 133–142. [\[CrossRef\]](#)
2. Zhang, Y.; Li, G.; Chu, L.; Wang, S.; Zhang, W.; Huang, Q. Cross-media topic detection: A multi-modality fusion framework. In Proceedings of the 2013 IEEE International Conference on Multimedia and Expo (ICME), San Jose, CA, USA, 15–19 July 2013; pp. 1–6. [\[CrossRef\]](#)
3. Fang, M.; Li, Y.; Hu, Y.; Mao, S.; Shi, P. A Unified Semantic Model for Cross-Media Events Analysis in Online Social Networks. *IEEE Access* **2019**, *7*, 32166–32182. [\[CrossRef\]](#)
4. Mander, J.; Kavanagh, D.; Buckle, C. GlobalWebIndex's Flagship Report on the Latest Trends in Social Media. GlobalWebIndex. 2020. Available online: <https://www.gwi.com/hubfs/Downloads/2019%20Q2-Q3%20Social%20Report.pdf> (accessed on 20 August 2022).
5. Golbeck, J.; Rothstein, M. Linking Social Networks on the Web with FOAF: A Semantic Web Case. In Proceedings of the 23rd AAAI Conference on Artificial Intelligence, AAAI Press, Chicago, IL, USA, 13–17 July 2008; pp. 1138–1143.
6. Rowe, M. Interlinking distributed social graphs. In Proceedings of the WWW2009 Workshop on Linked Data on the Web, LDOW 2009, Madrid, Spain, 20 April 2009.
7. Bennacer, N.; Nana Jipmo, C.; Penta, A.; Quercini, G. Matching User Profiles Across Social Networks. In *Advanced Information Systems Engineering, Proceedings of the CAiSE 2014, Thessaloniki, Greece, 16–20 June 2014*; Lecture Notes in Computer Science; Jarke, M., Mylopoulos, J., Quix, C., Rolland, C., Manolopoulos, Y., Mouratidis, H., Horkoff, J., Eds.; Springer: Cham, Switzerland, 2014; Volume 8484, pp. 424–438. [\[CrossRef\]](#)
8. Panchenko, A.; Babaev, D.; Obiedkov, S. Large-Scale Parallel Matching of Social Network Profiles. In *Analysis of Images, Social Networks and Texts, Proceedings of the 4th International Conference, AIST 2015, Yekaterinburg, Russia, 9–11 April 2015*; Communications in Computer and Information Science; Khachay, M., Konstantinova, N., Panchenko, A., Ignatov, D., Labunets, V., Eds.; Springer: Cham, Switzerland, 2015; Volume 542, pp. 275–285. [\[CrossRef\]](#)
9. Halimi, A.; Ayday, E. Profile Matching Across Online Social Networks. In *Information and Communications Security, Proceedings of the 22nd International Conference, ICICS 2020, Copenhagen, Denmark, 24–26 August 2020*; Lecture Notes in Computer Science; Meng, W., Gollmann, D., Jensen, C.D., Zhou, J., Eds.; Springer: Cham, Switzerland, 2020; Volume 12282, pp. 54–70. [\[CrossRef\]](#)
10. Diao, M.; Zhang, Z.; Su, S.; Gao, S.; Cao, H. UPON: User Profile Transferring across Networks. In Proceedings of the 29th ACM International Conference on Information & Knowledge Management (CIKM '20), Virtual, 19–23 October 2020; Association for Computing Machinery: New York, NY, USA, 2020; pp. 265–274. [\[CrossRef\]](#)
11. Malhotra, A.; Totti, L.; Meira, W., Jr.; Kumaraguru, P.; Almeida, V. Studying User Footprints in Different Online Social Networks. In Proceedings of the 2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, Istanbul, Turkey, 26–29 August 2012; pp. 1065–1070. [\[CrossRef\]](#)
12. Moshkin, V. The approach to building a graph knowledge base using social media data. In Proceedings of the 2020 IEEE 14th International Conference on Application of Information and Communication Technologies (AICT), Tashkent, Uzbekistan, 7–9 October 2020; pp. 1–6. [\[CrossRef\]](#)
13. Hiba, S.; Taieb, H.; Ali, M.; Mohamed, B.A. SNOWL model: Social networks unification-based semantic data integration. *Knowl. Inf. Syst.* **2020**, *62*, 4297–4336. [\[CrossRef\]](#)
14. Thelwall, M.; Vis, F. Gender and image sharing on Facebook, Twitter, Instagram, Snapchat and WhatsApp in the UK: Hobbying alone or filtering for friends? *Aslib J. Inf. Manag.* **2017**, *69*, 702–720. [\[CrossRef\]](#)
15. Phua, J.; Jin, S.V.; Kim, J. Uses and gratifications of social networking sites for bridging and bonding social capital: A comparison of Facebook, Twitter, Instagram, and Snapchat. *Comput. Hum. Behav.* **2017**, *72*, 115–122. [\[CrossRef\]](#)
16. Shane-Simpson, C.; Manago, A.; Gaggi, N.; Gillespie-Lynch, K. Why do college students prefer Facebook, Twitter, or Instagram? Site affordances, tensions between privacy and self-expression, and implications for social capital. *Comput. Hum. Behav.* **2018**, *86*, 276–288. [\[CrossRef\]](#)
17. Arora, A.; Bansal, S.; Kandpal, C.; Aswani, R.; Dwivedi, Y. Measuring social media influencer index- insights from facebook, Twitter and Instagram. *J. Retail. Consum. Serv.* **2019**, *49*, 86–101. [\[CrossRef\]](#)
18. Ye, S.; Ho, K.K.W.; Zerbe, A. The effects of social media usage on loneliness and well-being: Analysing friendship connections of Facebook, Twitter and Instagram. *Inf. Discov. Deliv.* **2021**, *49*, 136–150. [\[CrossRef\]](#)
19. Delle, F.A.; Clayton, R.B.; Jordan Jackson, F.F.; Lee, J. Facebook, Twitter, and Instagram: Simultaneously examining the association between three social networking sites and relationship stress and satisfaction. *Psychol. Pop. Media*, 2022, online ahead of print. [\[CrossRef\]](#)
20. Chia-chen, Y.; Lee, Y. Interactants and activities on Facebook, Instagram, and Twitter: Associations between social media use and social adjustment to college. *Appl. Dev. Sci.* **2020**, *24*, 62–78. [\[CrossRef\]](#)
21. Boulianne, S.; Larsson, A.O. Engagement with candidate posts on Twitter, Instagram, and Facebook during the 2019 election. *New Media Soc.* 2021, online ahead of print. [\[CrossRef\]](#)
22. Razis, G.; Anagnostopoulos, I. InfluenceTracker: Rating the impact of a Twitter account. In *Artificial Intelligence Applications and Innovations, Proceedings of the Artificial Intelligence Applications and Innovations Workshops (AIAI 2014), Rhodes, Greece, 19–21 September 2014*; IFIP Advances in Information and Communication Technology; Iliadis, L., Maglogiannis, I., Papadopoulos, H., Sioutas, S., Makris, C., Eds.; Springer: Berlin/Heidelberg, Germany, 2014; Volume 437, pp. 184–195. [\[CrossRef\]](#)

23. Razis, G.; Anagnostopoulos, I. Semantifying Twitter: The Influence Tracker Ontology. In Proceedings of the 9th International Workshop on Semantic and Social Media Adaptation and Personalization (SMAP), Corfu, Greece, 6–7 November 2014; pp. 98–103. [\[CrossRef\]](#)
24. Hirsch, J.E. An index to quantify an individual's scientific research output. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 16569–16572. [\[CrossRef\]](#) [\[PubMed\]](#)
25. Razis, G.; Theofilou, G.; Anagnostopoulos, I. Latent Twitter Image Information for Social Analytics. *Information* **2021**, *12*, 49. [\[CrossRef\]](#)