Understanding Privacy Concerns and Behavioral Intentions on Social Media: Exploring the Role of Marketing Comfort and the Potential of Blockchain Solutions

Auteur : Najib, Elias
Promoteur(s) : Standaert, Willem
Faculté : HEC-Ecole de gestion de l'Université de Liège
Diplôme : Master en sciences de gestion, à finalité spécialisée en international strategic marketing
Année académique : 2022-2023
URI/URL : http://hdl.handle.net/2268.2/18678

Understanding Privacy Concerns and Behavioral Intentions on Social Media: Exploring the Role of Marketing Comfort and the Potential of Blockchain Solutions

Jury:

Supervisor:

Willem STANDAERT

Reader:

Nicolas NEYSEN

Master thesis by:

Elias NAJIB

For a master's degree in management specializing in International Strategic Marketing

Academic Year: 2022/2023

Acknowledgements:

I would like to express my heartfelt gratitude to my supervisor, Pr. Willem STANDAERT, for his unwavering guidance, invaluable insights, and continuous support throughout the journey of conducting this research. Your dedication and mentorship have been instrumental in shaping the direction of this thesis.

I am also deeply appreciative of my friends, whose encouragement and camaraderie have provided me with the motivation to overcome challenges and persevere through the various stages of this work. Your discussions and perspectives have enriched my understanding and added depth to my work.

To my family, I extend my deepest appreciation for your constant encouragement, understanding, and belief in my abilities. Your unwavering support has been my pillar of strength, enabling me to navigate the hardships of my academic pursuits.

I am profoundly grateful to all the respondents who generously contributed their time and insights to participate in the survey, making this study possible. Your valuable input has been the cornerstone of this research.

Lastly, I want to acknowledge the valuable contributions of all the individuals, resources, and references that have played a role in shaping the foundation of this thesis. Each piece has contributed to the overall knowledge that this work embodies.

Thank you all for being an integral part of this journey.

Table of Contents:

I. Introduction: 5

II. Literature review: 7

1. Defining Social Media: 7

2. Defining Social media marketing: 10

3. Data collection / Privacy concerns: 11

4. Defining Blockchain: 13

5. Defining Blockchain-based social media: 14

III. Theoretical framework and hypotheses development: 15

1. Theoretical framework: 15

1.1 Concerns for social media information privacy (CFSMIP): 15

1.2 Marketing comfort: 16

1.3 Behavioral intentions: 16

1.4 Privacy calculus theory: 16

2. Hypothesis development: 17

3. Methodology: 18

4. Context of research: 18

IV. Research design: 19

1. Data collection: 19

2. Survey Structure: 20

3. Scales and measurements: 20

4. Respondents: 22

V. Results: 24

1. Descriptive statistics: 24

2. Preliminary tests: 25

3. Hypothesis testing: 28

VI. Discussion: 30

VII. Conclusion: 33

VIII. Limitations and Future Research: 33

IX. Executive summary: 34

X. Appendix: 36

XI. References: 52

I. Introduction:

In a world rapidly transformed by the proliferation of social media, the pillars of human connectivity, communication, and self-expression have undergone a profound change (Rautela, 2021). These digital spaces have created lots of chances for people to connect and exchange, disregarding limitations such as background, culture, or distance. Thanks to Internet-based technological advancements that align with the principles of Web 2.0, these platforms enable users to generate and exchange their own content while also fostering interaction and collaboration among individuals (Kaplan & Haenlein, 2010). However, the rapid evolution of the Internet in the context of social media has led to the creation of the attention economy, simply defined as a way to monetize internet surfers' attention. Traditional businesses as we know generate revenue by selling a product or a service, while the attention economy has now allowed social media platforms, search engines, and any web-based application with a significant volume of visitors, to generate profit by selling advertising spaces (Carpentier, 2023), as well as the data of individuals that visit said platforms, which in turn will be used to customize advertisements in a way that is often intrusive and unethical (Jungwoo, Cheong, & Kun Chang, 2022).

This thesis ventures into the complex dynamics of users' concerns for social media information privacy, their subsequent behavioral intentions, and the mediating role of marketing comfort. By exploring these dynamics, the study unravels how users perceive the privacy landscape and how this perception shapes their behavior within the digital spaces of social media.

Through this paper, a study of the variables responsible for social media users' decision to disclose their personal information will be conducted, aimed at investigating how the concerns for information privacy leads to specific behavioral reactions from the users, whereas how the concerns for social media information privacy systematically leads to negative feelings about disclosing private information on such platforms is tested. The relationship between the two variables is also studied from a different angle, in such a way that we consider the users comfort with marketing practices to mediate the relationship.

The integration of marketing comfort as a mediating variable introduces a novel dimension to the privacy-behavior relationship. While the relationship between concerns for social media information privacy and behavioral intentions is well-recognized, the potential role of users' comfort with marketing practices in mediating this relationship remains relatively uncharted. This study contributes to filling this gap by examining the mediating effect of marketing comfort, thereby enriching the theoretical understanding of the privacy-behavior relationship.

In sum, this study aims to contribute to the scholarly understanding of the intricate relationship between concerns for social media information privacy, behavioral intentions, and marketing comfort. Moreover, by considering the viability of blockchain as a potential solution, this research offers insights that hold relevance for both academia and the ever-evolving landscape of digital technology in the realm of social media. This research aims to fulfill the following objectives:

1. Investigate the relationship between concerns for social media information privacy and users' behavioral intentions regarding disclosing personal information on social media platforms.

2. Explore the potential of marketing comfort as a mediating variable in the relationship between concerns for privacy and behavioral intentions.

3. Assess the feasibility and desirability of blockchain-based social media platforms as a potential solution to privacy issues.

By addressing these objectives, the study aspires to enhance the theoretical understanding of the relationship between users' concerns, intentions, and comfort levels within the context of social media information privacy. Furthermore, the research contributes to the discourse on the viability of blockchain technology as a potential remedy for privacy-related challenges.

To achieve these objectives, a quantitative causal research approach was adopted, employing an online survey administered to a diverse sample of 103 social media users. The survey collected data on participants' concerns for social media information privacy, their behavioral intentions, marketing comfort levels, and their familiarity with the privacy benefits of blockchain technology.

The paper follows a mainstream approach, starting with a review of the literature around the subjects of social media and advertising practices in these spaces, followed by an overview of empirical studies in the context of data privacy in social media, more specifically the users concern for the often-unauthorized usage of their personally identifiable information for marketing purposes. The literature review is then concluded by a brief overview of a potential solution to the previously mentioned privacy issues by relying on the principle of decentralization, using blockchain as a foundational technology to control the mechanisms of collection and usage of the users' data. Next the theoretical foundations of the study are defined, which proves crucial in formulating meaningful hypotheses to be tested. This leads to an overview of the research design, and the data collection method employed. Finally, an overview of the results of the study is presented, followed by a discussion to offer managerial and scholarly implications, then finishing with a conclusion englobing our main findings and limitations to be taken into account.

II. Literature review:

1. Defining Social Media:

Since the birth of the Internet in 1983, it has been constantly developed to facilitate social interactions, especially with the emergence of the Web 2.0 functionalities, which allowed for an evolutionary leap forward in the social aspect of Internets use (Obar & Wildman, 2015). When we mention Web 2.0, it refers to a revolution to how the World Wide Web was used; it depicts how content and applications were no longer created and published by individuals, they were instead continuously modified by all users in a collaborative way (Kaplan & Haenlein, 2010). Indeed, the Web 2.0 did not come from any specific updates to its predecessor, the Web 1.0, it only refers to a new way the internet has been used, which in turn was the milestone that allowed for the creation of the modern social media platforms as we know them today, such as Facebook, Instagram and TikTok (Kenton, 2022). This change is mainly highlighted by the shift of the status of the user from a consumer to a participant (Obar & Wildman, 2015).

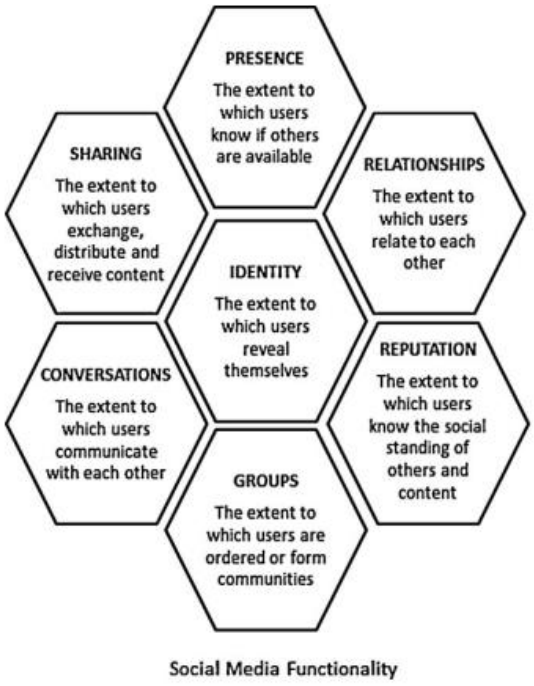
Social media can be seen as an extension of traditional word of mouth networks (Sajithra & Rajindra, 2013). As explained by Sajithra & Rajindra (2013), the introduction of techonolgy into the equation has allowed anyone with an access to internet to share his opinion and be part of social media. But what is social media? Today social media platforms are an integral part of our daily life, as of April 2023 we could count 4.8 billion social media users throughout the world, representing 59.9% of the world's population (Petrosyan, 2023). We may be able to identify whether a platform is a social media, as there are already some notorious names that come into our mind whenever we think of the concept, but it is important to have a formal definition of what social media is.

Multiple research fields have offered different definitions of social media, mainly the communication discipline, as well as other related ones such as public relations, information science and mass media (Carr & Hayes, 2015). What they all have in common is the fact that social media implies the use of online or internet technology (Wolf, Sims, & Yang, 2018) to emphasize user generated content or interaction (Kaplan & Haenlein, 2010). According to Wolf, Sims & Yang (2018), researchers should account for three mains components to properly define social media : the technological aspect, the ideological one (transparancy, sharing and integration) as well as the functional component. The technological aspect refers to the necessity to have access to internet and a device capable of connecting you to it, the ideological component, according to the aforementioned paper, is mainly enabled by the emergence of Web 2.0, as it allows for the programmatic integration of multiple applications thanks to the constant growth of standard integration protocols. The functional component refers to the possibilities opened to us by social media, including and not limited to creating a personal profile, viewing your personal network as well as your networks relationships (Boyd & Ellison, 2007), sharing and viewing content, networking and interacting with others (Kapoor, et al., 2017).

Taking into account the aforementioned components, the following definition includes all the important elements that should be highlighted to identify a platform that can be considered as a social media from one that is not :

Social media are internet based applications built on Web 2.0 (Huang & Benyoucef, 2013), that, according to Carr & Hayes (2015) : "Allow users to opportunistically interact and selectively self-present, either in real-time or asynchronously, with both broad and narrow audiences who derive value from user-generated content and the perception of interaction with others".

Figure 1 - The honeycomb of social media



As inclusive as this definition can be to social media in all shapes and forms, the ability to differentiate between social media platforms based on the functionalities they offer remains important. A simple example is the fact that Twitter and YouTube both fit correctly into this definition, but any user of both platforms knows that each one of them serves a different purpose. To fill this gap, we refer to Figure 1, which describes the seven functional building blocks of social media, with each block describing a specific aspect of the social media user experience (Kietzmann, Hermkens, McCarthy, & Silvestre, 2011).

The paper describes each building block as the following:

1. **Identity:** Identity represents the extent to which a social media user reveals personally identifiable information to his peers, such as the name, age, gender, occupations, locations and many more (Kietzmann, Hermkens, McCarthy, & Silvestre, 2011). We also include in the identity building block the concept of "self-disclosure", which was introduced to describe the conscious or unconscious revealing of personal information such as thoughts, feelings, likes and dislikes (Kaplan & Haenlein, 2010). It is an important concept that will help us build the model for our research later.
2. **Conversations:** This block represents the extent to which social media users communicate with one another. According to Kietzmann, Hermkens, McCarthy, & Silvestre (2011), these conversations have varying purposes depending on the social media platform, it could be to meet people with the same interests and opinions, to date, to build self-esteem, to always be on top of the latest news, or to share important messages to support humanitarian, environmental, economic and political causes.
3. **Sharing:** The sharing block of the framework represents the extent to which users can share, send and receive content with each other (Kietzmann, Hermkens, McCarthy, & Silvestre, 2011).

4. **Presence:** The presence building block represents the ability of a user to see whether another user is available (Kietzmann, Hermkens, McCarthy, & Silvestre, 2011). Like all other building blocks, the presence status varies from a social media platform to another. An example is how platforms like Facebook allow users to know whether their peers are offline or online, while other platforms such a Discord allows users to see the live activity of their peers. Aside from the traditional, online, offline, away statuses, it is possible to see which music a user is listening to on Spotify, or even which video game they are playing. This feature takes us back to the integration concept we discussed earlier, and how the Web 2.0 enabled social media platforms to integrate functionalities from multiple other applications (Wolf, Sims, & Yang, 2018).
5. **Relationships:** This building block depicts the ability of users to create relationships with one another on the platform, Kietzmann, Hermkens, McCarthy, & Silvestre (2011) describe these relationships as any form of association that allows users to communicate, share content, meet up, or only add each other in their friends list, for the case of Instagram and Twitter the followers list.
6. **Reputation:** The reputation building block represents the ability to identify a users standing in the platform. This can be done by evaluating the number of connections or followers they have, or also by the quality of the content they share, which can be assessed by the number of views, likes, reshares and many mechanics that vary from a platform to another (Kietzmann, Hermkens, McCarthy, & Silvestre, 2011).
7. **Groups:** The groups functional building block depicts the ability of users to create and be part of communities. Two different types of groups exist, the first one is the ability of each user to categorize his network into different sub-groups, such as close friends, friends, fans, followers and so on. The second one represents the grouping system that is used widely across most social media platforms, which is the ability to create communities or clubs for different subjects, interests and purposes (Kietzmann, Hermkens, McCarthy, & Silvestre, 2011).

The importance of these functional building blocks lies in the fact that they enable the differentiation between different types of social media, enabling each type to specialize in some functionalities over the others. No social media platform today focuses only on one of the blocks, but each one chooses a mix of three to four primary blocks (Kietzmann, Hermkens, McCarthy, & Silvestre, 2011) they will focus on to deliver services suitable to the population they aspire to serve. If we take as an example the three platforms LinkedIn, YouTube, and Facebook, they may all fall under the definition of social media we cited before, but LinkedIn puts at the core of its site the identity, relationships and reputation building blocks, while YouTube will focus more on the sharing, conversations, groups, and reputation building blocks. Facebook represents an example of a platform that englobes all seven building blocks, the user will be the one to define the degree to which he will benefit from each functional building block.

However, other researchers would argue that the backbone of all social media platforms lies in the identity building block, more specifically the user profiles (Boyd & Ellison, 2007). We indeed find proof of that in the fact that all social media platforms have some level of built-in functionalities for identifying the user's identity, differing by how personal they are but also by whether the information is disclosed willingly or unwillingly. An example of that is the platform Reddit, which allows users to join communities and exchange information and opinions on multiple matters. While the platform asks users for their personal information such as name, age, and gender, all of that information is not disclosed to other users, as each one is only identifiable by his username, which most of the time is not linked by any means to the users' real name. But we find that users can be identified by information they do not disclose directly, such as the communities they join (referred to as Sub-Reddits), the type of content they like or dislike (upvote or downvote for the case of Reddit), which allows the platform

in turn to suggest new types of content they could consume or new communities they may want to join. The reason the identity functional building block serves as a backbone is the fact that it enables social connections between users (Obar & Wildman, 2015), which is crucial as the social aspect is a must to build a successful social media platform.

With the increasing usage of social media across the world, notorious platforms not only served their social purpose, but they also grew to being a huge business phenomenon (Obar & Wildman, 2015). Facebook has been at the top of the list for many years now, it counted 2.9 billion monthly active users as of January 2023, next was YouTube with 2.5 billion, followed by WhatsApp and Instagram with each 2 billion monthly active users (Dixon, 2023). A common rule for businesses is to always be where customers are, and the previous numbers speak for themselves. With such a volume of users, no business aspiring to be successful can ignore how important it is to build a presence on those platforms. The next chapter will explain how businesses expanded their presence into this social phenomenon and used it not only to communicate with prospects and customers, but also to understand their needs and behaviors and adapt their communication strategies accordingly, in what we can call social media marketing.

2. Defining Social media marketing:

The introduction of social media into our daily life not only changed the way we live (Quan-Haase & Young, 2010), connect, share, and communicate with one another (Van Dijck, 2012), but also with businesses. Today social media can be used to share ratings of restaurants and hotels, recommendations for which product would best suit your needs, and advice for which businesses should be avoided due to mediocre service or products of inferior quality. It is therefore crucial for businesses to maintain a good perception of the quality of the products and services they offer, wherever large masses of audiences are present, and social media remains the biggest example of platforms where large gatherings of people and encounters of opinions occur in this digital era.

However, it is not only a necessity for businesses to improve their presence in social media, but it is now, thanks to the biggest players in the field, a massive opportunity to grow their customer base and target a worldwide audience. Indeed, Meta products represent the best examples for this empowerment, as Facebook and Instagram today allow businesses to target audiences in a precise manner using user personal information collected either willingly, or unwillingly disclosed by users, in what we call today social media marketing.

Before getting into social media marketing, it is necessary to define the bigger concept that englobes it, which is digital marketing. Digital marketing, simply put, regroups all marketing efforts that use an electronic device or internet (Desai & Vidyapeeth, 2019). According to Desai & Vidyapeeth (2019), digital marketing is an umbrella term that regroups many approaches to marketing, including and not limited to search engine optimization, content marketing, email marketing, affiliate marketing, marketing automation, and social media marketing. They define social media marketing as a business' efforts to promote their brand and content on social media platforms to improve brand awerness and capture more leads online. However, the applications of social media marketing may vary from a business to another, one of the main differences we can register is between the private and public sectors, whereas in the former companies often use social media marketing to communicate and sell products and services, while in the latter it is mainly used to share content of all types and encourage users interaction with the posts (Royle & Laing, 2014).

Aside from the mainstream use of social media as a communication medium between businesses and customers, such platforms are a rich source of data that companies can use for their benefit. Marketers today use data available on social media for opinion mining, targeted advertising, and for customer

relationships (Jacobson, Gruzd, & Hernandez-Garcia, 2020). Opinion mining refers to using data collected from social media platforms to extract insights and cutomers opinions on new or existing products or subjects in general, which can then be used to reorient the business towards providing solutions that would appeal more to their customer base. Targeted advertising refers to communicating different advertisements to each segment of the customer base, or only focusing on the segments that would be responsive to the type of content offered. Customer relationships refers to using social media to build and foster relationships with customers (Jacobson, Gruzd, & Hernandez-Garcia, 2020).

The possibilities social media has opened to marketeers are endless, and businesses today can only benefit from the growing number of users on such platforms. However, for businesses to use social media marketing effectively and ethically, social media users users need to be comfortable with those practises (Jacobson, Gruzd, & Hernandez-Garcia, 2020), and that includes the collection and usage of data for targeted advertising, as well as the exposure to advertisement, which is most of the time unsolicited.

The collection and usage of social media users data for marketing ends has been a trending subject for many years now. But the event that brought the most attention to this practice was the Facebook-Cambridge Analytica scandal in 2014, in which personally identifiable information of more than 87 million unaware Facebook users has been used by the firm Cambridge Analytica, using a personality test that was meant to evaluate their "OCEAN" psychological profile, meant to measure their Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism (Isaak & Hanna, 2018). The data was leaked and used to influence the presidential campaigns of Ted Cruz and Donald Trump. It was only four years after the act that Christopher Wylie, a former Cambridge Analytica employee, decided to blow the whistle on his former employer and denounce the misuse of the data, then legal measures were taken against Facebook, which was held accountable for the breach of information.

What's important to highlight in the previous story is the fact that, the"leaked" data was used to alter the presidential campaigns of the United States of America, and it was only four years after the act was done that information about the scandal came to light, and only because a former employee decided to reveal it. This fact alone gives us a good idea on how it is nearly impossible for social media users to know in real time where their data is stored, why it is collected and who is using it. The next chapter will help us uncover the journey user data goes through from the moment they sign up to a new social media platform.

3. Data collection / Privacy concerns:

Instead of analyzing each social media platform data collection and usage schemes, we will focus in this paper on uncovering how Facebook does it, or rather Meta, the father company which englobes Facebook, Instagram, WhatsApp, and many more social networking sites, with a special focus on Facebook and Instagram given the similarities between them and the massive volume of monthly users.

The Facebook-Cambridge Analytica scandal did not end after the fine was paid, as the event only marked the beginning of mass investigations into Meta's data collection practices. It was concluded that, for Facebook's case, the behavioral advertising business model is both the company's source of profit, but also the root of many problems (Rohit, 2019). Rohit (2019), a Federal Trade Commission former commissioner, believes that behavioral targeting works by using people's past behavior to manipulate their future actions using ads they would respond positively to, which makes Facebook's practices not only a breach of the users right for privacy, but also a huge threat to societies, by the fact that a company such as Facebook is now capable of manipulating the masses actions and opinions without them being conscious of it.

There are different types of data Facebook harvests from its users, and each type is collected in a different manner. In this paper, we will be focusing on Personally Identifiable Information. Personally Identifiable Information (PII) is information which can be used to trace an individual's identity (Krishnamurthy & Wills, 2009), it regroups mostly information we give voluntarily, such as the name, age, gender, and location, but also information that is generated throughout each visit of the social media platform, and this kind is the most valuable to businesses nowadays, as it is the fuel behind behavioral targeting. Behavioral targeting, defined in a general internet context, refers to the ability to target users based on their behavior while browsing online websites, it uses information such as the pages a user visited, or the searches they have made (Srimani & Srinivas, 2011). In the more specific context of social media, we can also include behaviors such as the type of content a user views, likes, and dislikes, the time of days when the platform is mostly used, and anything that can help advertisers segment users in a way that benefits their business.

While the kind of information collected can differ from one site to another, the same mechanism is applied for the collection of such data, which falls below the term we often hear about: Cookies. Whenever we visit any website, it is common to encounter a pop-up window that will let us choose whether we want to accept cookies, and which types of cookies we would like to accept, but what are these cookies? Also called HTTP cookies, they refer to small text files that websites place on the users device when surfing the internet (Koch, s.d.). According to the GDPR's website, cookies are harmless by nature and are there to facilitate a users browsing experience. However, it is when third party websites get access to data generated by such cookies that users privacy is in danger, as they enable websites to obtain information to identify users and send highly specific ads depending on their browsing activity. In the social media context, the same rules apply, but an additional layer of danger is applied, as the behavioral data collected by cookies Is linked to a users profile, and therefore instead of only knowing a users browsing history by an unknown ID, Facebook and other social media platforms can put a name, age, gender and location on such behavioral information, and can therefore target a single user with ads specifically tailored for him, instead of only putting him in a bucket and sending different ads to each segment (Srimani & Srinivas, 2011).

Many efforts have been made to give back the right of privacy to internet users in general. In Europe, the General Data Protection Regulation's (GDPR) 30th recital clearly talks about the uses of cookies, it states the following:

"Natural persons may be associated with online identifiers provided by their devices, applications, tools and protocols, such as internet protocol addresses, cookie identifiers or other identifiers such as radio frequency identification tags. This may leave traces which, in particular when combined with unique identifiers and other information received by the servers, may be used to create profiles of the natural persons and identify them."

Moreover, the ePrivacy Directive (EDP) was amended in 2009, and came to add a new layer of protection to the users data, as it was responsible for solving the transparency aspect users lacked before, the directive was what allows users today to receive cookie consent pop-ups on every website they visit (Koch, s.d.), in which they can see which types of cookies they will be receiving, and the ability to choose whether or not they accept them. However, as its name suggests, it is only a directive, and a proper regulation is yet to be released, as well as the fact that no specific regulations exist for the social media context, it is therefore a lawless land, and users continue to have their privacy breached, as their data is not only used to personalize the user experience, but also to allow advertisers to send tailored messages of what attracts them the most (Srimani & Srinivas, 2011) across multiple platforms. A concrete example of that is something everyone has gone through at least once while browsing the internet, a user could look up something on the internet, Nike shoes for example, only for informative

purposes. When done with his search, he could go into Instagram or Facebook, and find a sponsored advertisement for Nike shoes sold by some online vendor.

Despite some improvements after the data breach incident in 2014, Facebook's privacy remains partial. It tracks users' online activities, including browsing history, without consent, even when logged off. Unique ID numbers enable tracking even with privacy settings or deregistration (Lulandala, 2020).

Centralized organizations such as Facebook amass large quantities of personal and sensitive information from users who have little or no control over the data that is stored about them and how it is used (Zyskind, Oz, & Pentland, 2015). Many solutions to this issue of privacy breaches on social media platforms are being explored today, some may not be full proof, or require more development into the technologies behind them, others just did not receive a high level of acceptance from users yet. What most of these solutions have in common is one specific characteristic, referred to as decentralization, which means that the internet is controlled by many, no one actor can own it, control it or switch it off for everyone (Who Controls the Internet?, s.d.). In the next chapter we will be exploring a technology that is known for the highest degree of decentralization ever achieved in the context of the Internet, which is the Blockchain.

4. Defining Blockchain:

Blockchain is well known for being the technology behind Bitcoin and all cryptocurrencies in a general sense, the most successful application of it to this day, but it’s important to highlight it is not the only one, and unlike the internet, it is not unique, as there are many blockchains created. But what is a blockchain? By definition, it is a distributed database, or often referred to as ledger, that is shared among multiple computers in a network (Hayes, 2023). As its name indicates, it can be referred to as blocks in a chain where each new block is added to the previous one before it by consensus from all the nodes of the network. Once the details of the transactions or events are fed into a blockchain, it is impossible to tamper the details, as they are shared with all the members of the network (Chatterjee & Chatterjee, 2017). As previously mentioned, Bitcoin is the reason for the blockchain technology’s current fame, as it was its first major successful application, born thanks to Satoshi Nakamoto, the anonymous person that released the paper on how the Bitcoin cryptocurrency works (Nakamoto, 2008). In his paper, Nakamoto introduces the blockchain as a solution to have completely safe monetary transactions without the intervention of a third-party, he called it a system for electronic transactions without relying on trust.

According to Chatterjee & Chatterjee (2017), there are five main advantages to using blockchain: Immutability, irreversibility, ditributed system, resilience, and absence of a centralized authority. Immutability refers to the difficutly of altering a block after it has been added to the blockchain. Irreversibility is a feature that makes double spending impossible, by double spending we mean using a flaw to duplicate money. The distributed nature of the system is what allows each member of the network to possess a copy of the ledger, which is what makes it nearly impossible to change the information entered into a block after it has been added to the chain. Resilience refers to its high defense against any potential cyberattacks. And finally the absence of a centralized authority, the most important feature for our paper, refers to the absence of an organization that will own or control the data as it is exchanged between users. In the case of social media, Facebook, Instagram or Twitter would no longer be in control, they would just be a name that regroups members of a social media platform.

Blockchain technology was acknowledged to be revolutionary for many fields, including but not limited to supply chain management, smart contracts, real estate, healthcare, banking, and e-voting (Hayes, 2023), as it does not only offer decentralization, but also security, transparency, and user privacy. But

according to many researchers, the technology is still at its early stages, and still has much hidden potential to be exploited, as true blockchain-led transformation is still many years away (Iansiti & Lakhani, 2017).

The next big step for blockchain was in the year 2015 with the launch of the Ethereum blockchain, which marked the beginning of the use of smart contracts and dApps (History of Blockchain, 2023). dApps, an acronym for decentralized applications, refers to software programs that run on a network of computers that could be empowered by blockchain or a simple peer-to-peer connection of nodes (Frankenfield, 2023). The benefit of decentralized apps, when run on a blockchain network, is that they allow the platform to be free from control from one single authority. The launch of such technology has marked a revolution in applications such as self-executing contracts, multi-user games, and social media platforms. In the context of social media, the decentralized nature of the platform allows for four main benefits, transparency, censorship resistance, user privacy (Frankenfield, 2023), as well as content curation, as it is easy to implement a rewarding system using cryptocurrencies into such platforms (Guidi, 2020).

5. Defining Blockchain-based social media:

As we discussed in the previous chapter , current social media platforms suffer from many privacy-related issues, as all data are stored in the server(s) of the social media platform, Facebook for example, and users must trust that it protects all their personal information (Guidi, Conti, Passarella, & Ricci, 2018). However, the Facebook-Cambridge Analytica scandal was enough proof that social media users cannot rely only on trust, as the giant players in social media can easily manipulate users' data as they please and hide any breaches of security if they ever happen. Since trust in a central organism is not a reliable option anymore, the most logical solution would be to decentralize the social media platforms. Researchers explored this solution and are coming up with an array of terms to define the concept, including and not limited to decentralized online social media, decentralized online social networks, and even more specifically blockchain-based social media, which is the term we will be using going further into this paper. Before defining blockchain-based social media, it is logical to start by defining a decentralized online social media, as a blockchain-based social media is in essence one of the many forms a decentralized online social media can take. A decentralized online social media is defined as an online social media that exploits the decentralized nature of its social services using a distributed platform, such as a blockchain or a peer-to-peer network (Datta, Buchegger, Vu, Strufe, & Rzadca, 2010). Overall, The rupture of trust between social media users and the platforms has already caused the collapse of some early social media services, alongside other reasons such as technical issues and management problems due to the exponentially increasing number of users (Datta, Buchegger, Vu, Strufe, & Rzadca, 2010), all issues that can be solved by decentralizing such platforms.

Blockchain's integration into social media is still a nascent idea, many new players have adhered to it, but most of them come up with solutions that are still under development (Guidi, 2020). The common traits between them all is that they offer three main advantages: Improved transparency, data ownership is given back to the users, and incentivizing actions taken in the platform (Ivey, 2023). Many social media platforms of such nature have been developed to this day, the most prominent ones being SteemIt, Peepeth, Lit, Verasity, Sapien, SocialX and many more. Just like regular social media platforms we see today, each one of these is special compared to the others from different facets, such as the functionalities they offer, the type of content exchanged, and the degree of decentralization. For the sake of our research, we will be taking into account the platform SteemIt, which can be considered as inspired by Reddit for the similarity of functionalities they offer (Guidi, 2020), as it is mainly a platform for discussions. The choice of SteemIt was not arbitrary, it is justified by its high number of users

compared to other blockchain-based social media platforms, as it currently counts over 1 million registered users.

SteemIt is a website based on the Steem blockchain, which is a blockchain that was designed specifically to create decentralized social media platforms (Cryptopedia, 2022). The Steem blockchain relies on three different kinds of currency units: Steem, Steem Power (SP), and Steem Dollars (SBD), and each serves a different purpose in the platform. Unlike other blockchains, which use the Proof-of-Work (PoW) as a consensus protocol, Steem uses a Delegated Proof of Stake (DPoS) for its robustness in the context of social networks. Another main difference is that instead of block miners, nodes responsible for adding a block are referred to as witnesses, whereas twenty-one random witnesses are selected per round to create and sign blocks of transactions (Guidi, 2020).

Aside from its decentralized nature that allows for higher user privacy, SteemIt focuses on curating its content by rewarding users for the quality of what they post, which is decided by none other than the other members of the community. Just like Reddit, SteemIt relies on the upvote/downvote system to decide whether a post contains relevant and reliable information or not. A slight difference between the two platforms lies in the fact that the upvote or downvote of a user has a weight that is decided by his Voting Power. This mechanism was implemented to give more value to a user's vote, as it limits the number of contents voted by a user in a period of time (Guidi, 2020).

If we go back to the seven fundamental building blocks of social media designed by Kietzmann, Hermkens, McCarthy, & Silvestre (2011), we mention how each social media platform choses a mix of the bulding blocks it would like to focus on, which gives birth to a unique experience for the users. Reddit and Steemit being fundamentally the same, we can deduct that they both focus on the same building blocks, which are conversation, reputation and groups. Consequently, the fact that we chose Facebook as our model for a centralized social media platform and SteemIt for the decentralized ones may appear to not make sense, as each platform focuses on different building blocks and therefore serves a different purporse for its users, but it is important to highlight that we only consider those platforms for the mechanism they use for collecting and using user data, therefore we are not testing our researches theories on the SteemIt platform for its rewarding system or its absence of censorship, but only for its use of the blockchain technology to store user data. This gap is due to the fact that there are no platforms that could be compared to Facebook which are blockchain-based, our research should however be able to be generalized to a context where such platform is created in the future.

III. Theoretical framework and hypotheses development:

1. Theoretical framework:

Based on the literature review we conducted above, we can extract many relevant constructs to our research, whereas they will help us develop our model. The following part aims at defining the most important ones from the literature they were taken from, then use them to formulate our hypotheses and finally the conceptual model we will follow in our research.

1.1 Concerns for social media information privacy (CFSMIP):

Information privacy has been subject to concern for many years now, even before the emergence of social media platforms as we know them today. One of the major studies on this matter was done by Smith, Milberg, & Burke (1996), whereas the paper enabled future studies in this stream by developing a valid instrument to measure the different dimensions around an individuals concerns about an organizations information privacy practices. However, it had yet to be adapted to the current digital era, charaterized by massive amounts of data collected everyday from different types of organizations.

At that time, interesting dimensions were developped, which we can use for our study, Smith, Milberg, & Burke (1996) have explained them as follows:

- **Collection:** The collection dimension measures the concern that large amounts of personally identifiable information are being collected and stored by organizations.
- **Secondary use (internal & external):** These two dimensions refer to using data collected from individuals for a different purpose than the original one behind the collection, refered to as secondary. The difference between the internal and external dimensions is that the internal one refers to the secondary use of the data by the same organization that collected it, while the external one refers to the secondary use of the data by an external party to whom the data has been disclosed.
- **Unauthorized Access:** This dimension measures an individuals degree of concern that their data can be accessed by people not authorized to do so.
- **Errors:** The errors dimension measures an individuals concern that their data is not protected against deliberate and accidental errors.

The aforementioned scales were used by many researchers to measure the concern for information privacy (CFIP), which is general to any organization that meddles with a large number of individuals data. This model was later adapted to the social media context, to measure the mediating role of the concern for information privacy between social media anxiety and behavioral intentions (Osatuyi, 2015).

1.2 Marketing comfort:

The idea of "marketing comfort" was developed as a way to figure out how at ease social media users are with the thought of their information being used for specific purposes. In our previous review of existing research, we highlighted three main purposes: Targeted advertising, opinion mining, and customer relationships. This concept was introduced by researchers Jacobson, Gruzd, and Hernandez-Garcia in 2020. They aimed to understand how people feel about their data being utilized for these purposes on social media platforms. This concept provides valuable insights into users' sentiments and perceptions regarding these different marketing activities. By examining the level of marketing comfort, we can use our research to gain a clearer picture of users' attitudes towards data usage in the realm of social media marketing.

1.3 Behavioral intentions:

In the past, many researchers studied something called "behavioral intentions." They used this idea to figure out how people react when they must share personal information. This could be when they're making a bank account or joining a social media site. Some researchers, like Smith, Milberg, and Burke (1996), Stewart and Segars (2002), and Osatuyi (2015), looked at this and developed the construct as well as tested it as a valid instrument. They helped us understand how people make choices about giving away their personal info or keeping it private, related to other constructs.

1.4 Privacy calculus theory:

The Privacy Calculus theory, originally formulated by Laufer and Wolfe in 1977, is a concept that considers information privacy as an economic commodity that can be exchanged for other goods or benefits. This theory suggests that individuals engage in a cost-benefit analysis when deciding whether to disclose their personal information. When faced with the choice of sharing their data, individuals weigh the potential benefits they might gain against the perceived risks and costs associated with potential negative outcomes. The costs in this context are often linked to the potential risks that might arise from the exposure of personal information, such as unauthorized access, misuse, or loss of control

over the data. The theory is based on the idea that individuals consider not only immediate gains but also future consequences when making decisions about privacy (Laufer & Wolfe, 1977). Even though it was developed decades before the current data-driven digital era, this concept still is applicable and describes well the behavior of social media users, we will therefore have to take it into account when conducting our research, as it may explain any contradictory behavior regarding the concern for social media information privacy and the users' behavioral intentions.

2. Hypothesis development:

The aim of our research being to examine the typical social media users' reaction towards privacy risks when using such platforms. We formulate the following hypotheses based on the concepts we explored in the literature review and the above theoretical framework:

H1: *A higher concern for social media information privacy will lead to negative behavioral intentions from social media users.*

H1a: A higher concern for information access in social media will lead to negative behavioral intentions from social media users.

H1b: A higher concern for errors of data collection and storage in social media will lead to negative behavioral intentions from social media users.

H1c: A higher concern for data collection in social media will lead to negative behavioral intentions from social media users.

H2: *A higher concern for social media information privacy will lead to less marketing comfort from social media users.*

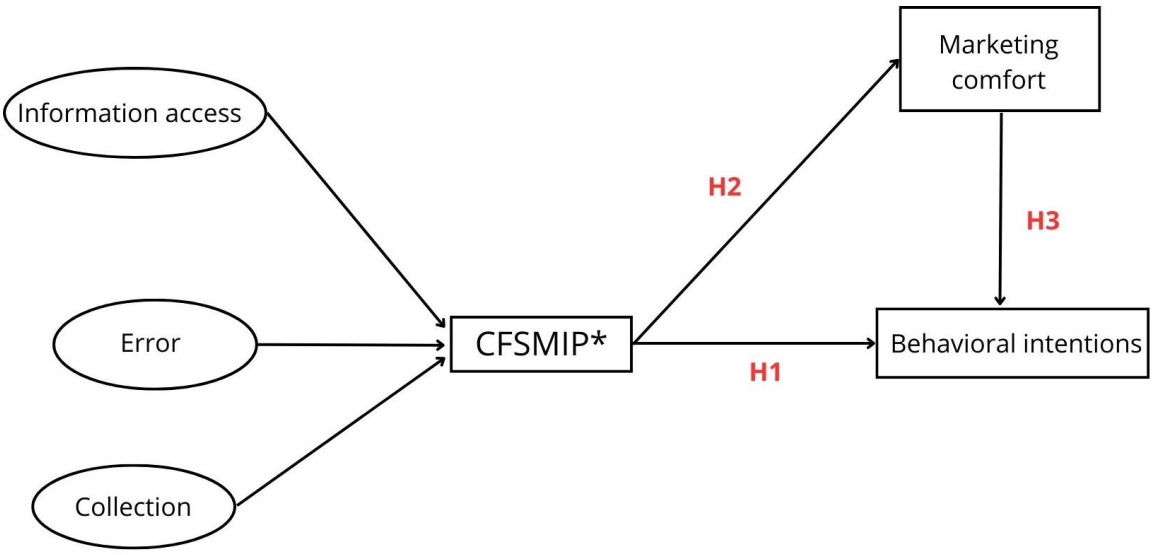
H2a: A higher concern for information access in social media will lead to less marketing comfort from social media users.

H2b: A higher concern for errors of data collection and storage in social media will lead to less marketing comfort from social media users.

H2c: A higher concern for data collection in social media will lead to less marketing comfort from social media users.

H3: *Marketing comfort mediates the relationship between concern for social media information privacy and social media users' behavioral intentions.*

Based on the concepts we explored as well as the hypotheses formulated, the following model describes the relationships we would like to test. From this model we can see that we used the same instruments as the papers from Stewart and Segars (2002), and Osatuyi (2015), as they were tested and validated instruments to measure the concern for information privacy, and for our case the concern for social media information privacy (CFSMIP). Also based on our hypothesis, we suspect that the concerns for social media information privacy will impact the behavioral intentions of social media users indirectly, with the marketing comfort construct acting as a mediating variable between the two. The information access instrument was developed to include both the "Unauthorized access" and the "Secondary use (internal & external)" constructs, it includes the same items as the ones used by previous research, the only difference is that Osatuyi (2015) grouped them into one for the sake of clarity.



*Concerns for social media information privacy

Figure 2 Conceptual model

3. Methodology:

In this section, we will outline the methodology employed to come to relevant conclusions regarding our research question, in regards of the effects of the concerns for data privacy in a social media context on the behavior of social media users towards such platforms. The next chapters will encompass the research design, data collection methods, sampling of the participants and finally the data analysis techniques we will utilize in our study.

4. Context of research:

The goal of our study is to understand the mechanics responsible for social media users allowing or disallowing their data to be collected and used for different marketing purposes. More specifically, we aim through this paper to understand the effects the concerns for social media information privacy has on the users comfort towards three different kinds of marketing practices we extracted from the literature, which are targeted advertising, opinion mining about products and services, and building customer relationships, as well as the mediating effect this marketing comfort has on the relationship between concerns for social media information privacy and a typical social media users behavioral attitudes towards these platforms from a privacy point of view. We also include in the insights we would like to gather the level of understanding social media users have on blockchain technology and its benefits from a privacy perspective, as well as whether they would join a new type of social media platforms if they offered them increased privacy. The aim of this last series of questions about blockchain may not contribute to analyzing the relationships between the constructs we defined in our model, but they will prove crucial to knowing the potential such an improvement would have in the future of social media.

To test our hypotheses and reach relevant conclusions, quantitative data will be gathered through the means of an online survey. From there on, the data gathered will be analyzed to test the relationships

between our constructs, as well as some important information about the respondents' social media usage frequency, the platforms mostly used, and their demographics. The next parts will dive deeper into the details of how the survey was conducted to collect quantitative data.

IV. Research design:

1. Data collection:

The research design section aims to describe the methods chosen to conduct our study. For our case, the data used is of the quantitative type, gathered using an online survey published using the Qualtrics platform, the details around the survey will be discussed in the next chapter. Our study can be divided into two objectives, a primary one and a secondary one. The primary one, as explained earlier, is to describe the causal relationship between the variables in our model, while the secondary one is to understand social media users view on the blockchain technology and the level of acceptance it would receive in the future if platforms that offer the same functionalities as the current main ones (Facebook, Instagram, Twitter…) are created, but as blockchain-based platforms which offer much more privacy.

The reason for not merging these two objectives into one by simply analyzing the causal relationship between the same variables we have, only this time for blockchain-based platforms, is that the technology currently has a low level of adoption in the context of social media, and therefore it is not possible to gather opinions about concerns for social media information privacy and marketing comfort when we suspect that most social media users do not use or even know about such alternatives yet. For this reason, we include questions about the knowledge social media users have on blockchain technology and its applications to increase data privacy, which will help future research in defining the directions that should be taken in regards of blockchain-based social media platforms.

Coming back to the primary objective of this study, the causal relationships we research are between three variables, the concerns for social media information privacy, the marketing comfort, and the behavioral attitudes. In this setup, we consider our dependent variable to be the behavioral attitudes one, as it is the construct we would like to measure using the other independent variables. The independent variable that we hypothesize has a direct and indirect effect on the dependent variable is the concerns for social media information privacy, the direct effect will be analyzed through the validation or rejection of the first hypothesis H1, while the indirect effect will be tested using the second hypothesis H2, which states that the concerns for social media information privacy indirectly affects behavioral attitudes through the mediating effect of the marketing comfort variable, consequently the latter will be considered a mediating variable in our model, its effect is translated into our second hypothesis H2. On another hand, literature has provided us with the CFSMIP variable as a construct that should be measured using three other constructs, which are error, collection, and information access, with information access including both the unauthorized access and secondary use variables (Osatuyi, 2015). The effects of these three constructs on the CFSMIP construct are not included in our hypotheses, since they do not contribute to the results we would like to come to in our research, it is however necessary for us to analyze the effects of these variables on the construct and possibly remove some items to improve the instrument.

After defining our objectives, we can conclude that our research design is of the conclusive type, which as explained by Malhotra, Nunan, & Birks, (2017), is a research design that is used to test hypotheses and examine specifc relationships between variables. More specifically, this study will be conducted using a causal research design, as it is used to analyze cause-and-effect relationships between variables. Conclusive research can be conducted using multiple methods, the one chosen for this study is an

online survey, for its high potential of reach across multiple platforms. The next chapter will dive deeper into the surveys structure.

2. Survey Structure:

To collect our quantitative data, we rely on an online survey conducted using the platform Qualtrics. The reason for choosing this platform over the other options available to us is the limitless functionalities it offers, and the high compatibility with smartphone users respondents, as it is suspected that most respondents will answer from their phone, given that the link to the survey was mainly shared throughout the main social media platforms used by the population this research aspires to study, which are Facebook, Instagram and LinkedIn.

Prior to officially sharing the survey on the platforms cited above, we carried out a trial run with 3 participants, who were master students in the same field as myself. This trial run was important because it helped us make the survey better. We looked at the way the survey was set up, found any possible problems, and made sure the questions were easy to understand, related to the topic, and didn't show any favoritism. We listened carefully to what they said and made a lot of changes based on their suggestions. Thanks to their help, the final survey turned out much better than before.

The survey was introduced by a concise paragraph where the objective of the study is explained, and the anonymity of the respondents is highlighted. The first question right after the introduction is a simple one, but crucial to our filtering of the respondents. The nature of this study dictates that the respondents should be social media users to be able to give meaningful insights into their perceptions of privacy and their comfort towards marketing practices on social media platforms, therefore the questionnaire starts with the question “Do you use social media”, with the options Yes or No as answers. The goal being to filter out the respondents using this question, those who answer yes will be directed to the rest of the questionnaire, while does who answer no have their survey immediately finished.

After this first step of filtering is passed, the respondents were confronted with a set of structured questions, aiming at assessing multiple dimensions of their social media usage, such as the time spent every day, the most prominent platforms used, the reason why they use social media and their general privacy preferences across their social media accounts. Next the respondents were faced with the core questions of the study, whereas they needed to share their level of agreement with multiple statements, aiming at measuring the variables CFSMIP, marketing comfort, and behavioral attitudes, for a total of 19 statements spread out into multiple pages for the sake of clarity.

To achieve the studies secondary objective as explained in an earlier chapter, three questions were asked separately to assess the level of knowledge respondents have around the blockchain technology and its use for increased privacy, as well as whether they would switch to a blockchain-based platform if it offered increased data privacy. The aim being to understand if a nascent technology such as this one could be potentially adopted in the future, in correlation to the level of concern for social media user information privacy as well as the familiarity with blockchain. Finally, the questionnaire ends with questions aimed at collecting basic socio-demographic information about the respondents, including gender, country of residency, age, and employment status. The full survey is provided in appendix 1.

3. Scales and measurements:

As explained in the theoretical framework chapter, the literature has provided us with validated instruments to be used for our research, Table 1 below provides an overview of the six constructs that are measured using the questionnaire, and the respective items for each one of them, which are enlisted as statements for the respondents to express their level of agreement with:

Construct	Scale	Item	Statement
CFSMIP Information access: Unauthorized Access (Smith, Milberg, & Burke, 1996) ; (Osatuyi, 2015) ; (Jacobson, Gruzd, & Hernandez-Garcia, 2020)	7-point Likert scale ranging from 1 (strongly disagree) to 7(strongly agree)	UAC1	Social media databases that contain personal information should be protected from unauthorized access – no matter how much it costs.
		UAC2	Social media sites should take more steps to make sure that unauthorized people cannot access personal information on their site.
		UAC3	Databases that contain personal information should be highly secured.
CFSMIP Information access: Secondary Use (Smith, Milberg, & Burke, 1996) ; (Osatuyi, 2015) ; (Jacobson, Gruzd, & Hernandez-Garcia, 2020)	7-point Likert scale ranging from 1 (strongly disagree) to 7(strongly agree)	SUS1	Social media sites should not use personal information for any purpose unless it has been authorized by the individuals who provided the information.
		SUS2	When people give personal information to social media sites for some reason, these sites should never use the information for any other purpose.
		SUS3	Social media sites should never share personal information with third-party entities unless authorized by the individual who provided the information.
CFSMIP Error (Smith, Milberg, & Burke, 1996) ; (Osatuyi, 2015) ; (Jacobson, Gruzd, & Hernandez-Garcia, 2020)	7-point Likert scale ranging from 1 (strongly disagree) to 7(strongly agree)	ERR1	Social media sites should take more steps to make sure that personal information in their database is accurate.
		ERR2	Social media sites should have better procedures to correct errors in personal information.
		ERR3	Social media sites should devote more time and effort to verifying the accuracy of the personal information in their databases before using it for recommendations.
CFSMIP Collection (Smith, Milberg, & Burke, 1996) ; (Osatuyi, 2015) ; (Jacobson, Gruzd, & Hernandez-Garcia, 2020)	7-point Likert scale ranging from 1 (strongly disagree) to 7(strongly agree)	COL1	It usually bothers me when social media sites ask me for personal information.
		COL2	It usually bothers me when social media sites ask me for my current location information.
		COL3	It bothers me to give personal information to so many people on social media.
		COL4	I am concerned that social media sites are collecting too much personal information about me.
Marketing Comfort (Jacobson, Gruzd, & Hernandez-Garcia, 2020)	7-point Likert scale ranging from 1 (Extremely uncomfortable) to 7(Extremely comfortable)		*How comfortable would you be if information about you or posed by you publicly on social media is used for:*
		MCAD	Targeted advertising.
		MCOM	Gathering user opinion about products or services.
		MCCR	Building customer relationships.
Behavioral intentions (Stewart & Segars, 2002) ; (Osatuyi, 2015)	5-point Likert scale ranging from 1 (Very unlikely) to 5 (Very Likely)		*How likely are you, within the next three years to:*
		BIN1	Provide personal information on social media sites.
		BIN2	Make your social media account public for others to find easily.
		BIN3	Share information on social media sites with your friends.

Table 1 Research construct and items

As described in the table above, all questions dealing with the study variables will be in the form of a Likert scale, with the objective of measuring the degree of agreement among questionnaire respondents. The Likert scale is educational and social science research's most basic and widely used

psychometric tool (Joshi, 2015). In 1932, the scale was developed to measure 'attitude' in a scientifically accepted and validated manner (Edmondson, 2005).

According to Heale & Twycross (2015), validity is the degree to which an idea is accurately measured in a quantitative study. Three types of evidence can be employed to demonstrate the construct validity of a research instrument. These are homogeneity (the instrument measures only one construct), convergence (occurs when the instrument measures concepts that are comparable to those measured by other instruments), and theory evidence (observed when behavior resembles theoretical assertions of the constructs under consideration). The second quality parameter in a quantitative inquiry is instrument accuracy. In other words, the degree to which a research instrument consistently provides consistent results when used in the same setting several times. The three reliability characteristics are homogeneity, stability, and equivalence (Heale & Twycross, 2015).

While many papers previously validated and used the instruments listed above, they were adapted many times repeatedly to different contexts, from research subjects that weren't related to any digital component. The CFSMIP construct for instance started out as CFIP, which refers to concern for information privacy (Osatuyi, 2015). Alongside the behavioral attitudes construct, they were used by past researchers to study the concerns for information privacy in a general context for individuals who share their general information to create a bank account, subscribe to a newsletter, and other applications which do not require extensive amounts of information. The instruments were later adapted to reflect the concerns for privacy in a social media context, where data collected is much more personal and intrusive, and the risks of privacy breaches are much higher.

4. Respondents:

The survey conducted has gathered 103 answers over the course of a week, thanks to the use of social media as well as the website Survey Circle, the target number of answers was reached in a record time, after which the survey was closed to avoid any additions to the database. The first question of the survey, which aimed at filtering out respondents who do not use social media, did not exclude many answers, as only 2 people claimed they do not use social media, which made sense as the main channels through which the survey was distributed were the social media platforms Facebook, Instagram and LinkedIn, a respondents would therefore have access to said spaces to be able to answer.

The data cleaning and validation process did not have any effect on the remaining 101 recorded responses, since no incomplete answers were given, and no respondent seemed to finish the questionnaire in an unusually short duration. The remaining 101 answers were therefore considered as valid for use in our research.

The socio-demographic questions revealed an overview of the respondent's gender, age, location, and employment status distributions. The sample had an even gender distribution, composed of 51 female respondents and 50 male ones, while no answers were recorded from people who identified as non-binary / third gender. The highest percentage of respondents were in Belgium, covering 40.6% of the total sample, while 29.7% of the respondents answered "Other", which included mainly answers from Morocco. Then came France with a percentage of 16.9% of the total sample. Overall, the data states that 70.3% of the respondents were in the Western Europe region. The questionnaire was mainly answered by a young population with ages ranging between 18 and 34, whereas 69.3% of the sample fell into the age range between 18 and 24 years old, while 27.7% was from the range between 25 and 34 years old. The remaining 3% belonged to respondents between the ages of 45 and 55 years old, which highlights the level of interest emanating from a younger population for a subject such as this one. As for the employment status of the sample, the data reveals that 48.5% of the respondents were

students, while 37.6% were employed. Further details about the socio-demographic description of the sample are presented In Table 2 below:

Gender	Female	50,5%
	Male	49,5%
	Non-binary / third gender	0,0%
	Prefer not to say	0,0%
Age	Under 18	0,0%
	Between 18 and 24	69,3%
	Between 25 and 34	27,7%
	Between 35 and 44	0,0%
	Between 45 and 55	3,0%
	Above 55	0,0%
Location	Belgium	40,6%
	Luxembourg	11,9%
	Netherlands	1,0%
	France	16,8%
	Other	29,7%
Employment status	Student	48,5%
	Employed	37,6%
	Unemployed	13,9%
	Retired	0,0%

Table 2 Socio-demographic statistics

V. Results:

In this section, we will go through the results collected thanks to the online survey, starting with descriptive statistics gathered using the social media usage and blockchain familiarity questions, then conducting the necessary preliminary tests to test the reliability and validity of the data collected using the scales taken from the literature, and finally using the results extracted from the software SPSS, validating, or rejecting our hypotheses.

1. Descriptive statistics:

The first and second to last sections of the questionnaire aimed at assessing the social media usage dynamics exhibited from the respondents, and their level of familiarity with the blockchain technologies benefits in terms of privacy, as well as the likelihood of respondents to switch from the social media platforms they currently use for one that offers additional privacy using blockchain.

The social media usage dynamics questions were five in total, with the objective of collecting data on daily spent time using social media, social media platforms mostly used, reasons for using social media, posting frequency, and overall privacy settings maintained across used platforms. The sample showcased a tendency towards using social media for more than 3 hours every day, whereas 44.6% of the respondents claim they use social media for 3 to 5 hours a day, and 30.7% for more than 5 hours a day. The "platforms mostly used" question revealed interesting insights about the respondents' habits, as all 101 of them use Instagram, while 51.49% of them use Facebook, followed by LinkedIn with 43.56% of the respondents using it. The reasons the respondents used social media were mostly for entertainment, staying connected with friends and family, developing their personal and professional network, and sharing photos and videos, with respective percentages of 90.1%, 71.29%, 55.45%, and 54.46%. The final questions regarding the respondents posting and privacy tendencies revealed that 55.4% of the respondents rarely posted on social media, and 52.5% of them only kept their basic information public. As for the questions concerning the familiarity with blockchain technology, they revealed that 69.3% of the respondents were familiar with it, while the question about the familiarity with the benefits of the blockchain in terms of privacy revealed that 26.7% of respondents were moderately familiar with it, while 25.7% and 27.7% respectively were slightly familiar and not at all familiar with it. Finally, 34.7% of the respondents answered that they were somewhat likely to switch to a blockchain based platform if it offered additional privacy, while 30.7% of respondents were neither likely nor unlikely to switch to a blockchain-based platform. Further details about the descriptive statistics are provided on Table 3 below:

Daily usage	Less than 1 hour	2%
	1-2 hours	22,8%
	3-5 hours	44,6%
	More than 5 hours	30,7%
Channels used	Facebook	51,5%
	Instagram	100%
	Twitter (X)	14,9%
	LinkedIn	43,6%
	Snapchat	28,7%
	Reddit	7,9%
	TikTok	38,6%
	Other	7,9%

Reasons for use	Entertainment	90,1%
	Developing your personal and professional network	55,4%
	Sharing photos and videos	54,5%
	Learning	35,6%
	Staying connected with friends and family	71,3%
	Business	6,9%
	Other	5%
Posting frequency	Never	8,9%
	Often	54,5%
	Somewhat often	22,8%
	Very often	8,9%
	Always	5%
Privacy settings	Fully private	36,6%
	Only my basic information is public	52,5%
	Fully public	10,9%
Blockchain familiarity	Yes	30,7%
	No	69,3%
Blockchain privacy familiarity	Not at all familiar	27,7%
	Slightly familiar	25,7%
	Somewhat familiar	14,9%
	Moderately familiar	26,7%
	Extremely familiar	5%
Blockchain-based social media usage likelihood	Extremely unlikely	3%
	Somewhat unlikely	11,9%
	Neither likely nor unlikely	30,7%
	Somewhat likely	34,7%
	Extremely likely	19,7%

Table 3 Social media usage and blockchain familiarity statistics

2. Preliminary tests:

The first step before conducting our main data analysis and testing our hypotheses is to carry out preliminary tests on the database, aiming at testing our scales and the items used to measure them for reliability and validity. We start by testing the normality of the items used to measure each scale, which can be done by using a Shapiro-Wilk normality test. The results of the test revealed that most items have a statistic value close to 1, noting that the closer the value is to 1, the more normally distributed the data is. Also, the results show P-value inferior to the chosen significance level (0.05), which suggests that the data significantly deviates from a normal distribution. Diving into further detail about the distribution of the items, we investigate the skewness and kurtosis values revealed by the same test. The skewness of all items is negative expect for "BIN2", which means that all items have values that skew to the left, while "BIN2" skews to the right. Also, some items have a skewness that is outside the range -2 to 2, which reveals that they are not normally distributed. The kurtosis on the other hand reveals that most of our items have a heavy tail and sharper peak compared to a normal distribution. The detailed values of the Shapiro-Wilk test are revealed in appendix 2.

The next step involved the use of Cronbach's alpha test to assess the reliability of the items used in our study. While the instruments used were already tested and validated by the previous studies, we extracted them from, an additional test was done to make sure of their validity. Many researchers consider Cronbach's alpha value to be acceptable when it is above 0.7, but some researchers tend to

consider a lower or higher alpha measure depending on the characteristics of the study and the field of the research. The results of the test show that our “CFSMIP” and “Marketing Comfort” constructs, with respective results of 0.89 and 0.756, are above the 0.7 threshold used by most researchers, which makes them reliable constructs. The “Behavioral Intentions” construct on another hand exhibited a value of 0.641, which remains close to 0.7, it will therefore be accepted as well, since the construct and its items were already validated by the study it was extracted from (Osatuyi, 2015).

Following the Cronbach’s alpha test, an exploratory factor analysis (EFA) was done to measure the relationships between items measuring the same construct, with the aim to remove any items that may be correlated with each other. Based on the components matrices shown in Tables 4, 5, and 6 below, it was decided to remove all three items meant to measure the errors in the collection and storage of the data, as they had high values for more than one component. This manipulation resulted in an improvement of the construct’s reliability, as shown in table 7. The “Marketing Comfort” and “Behavioral Intentions” scales did not have any of their items removed.

Components matrix: CFSMIP			
	Component		
	1	2	3
UAC1	0,666	-0,465	-0,174
UAC2	0,810	-0,305	-0,321
UAC3	0,835	-0,235	-0,335
SUS1	0,733	-0,321	-0,129
SUS2	0,847	-0,038	0,016
SUS3	0,822	-0,024	-0,307
COL1	0,403	-0,201	0,711
COL2	0,447	-0,226	0,705
COL4	0,709	0,033	0,205
COL3	0,582	-0,177	0,347
ERR1	**0,536**	**0,741**	0,042
ERR2	**0,653**	**0,699**	0,028
ERR3	**0,679**	**0,659**	-0,001

Table 4 Components matrix: CFSMIP

Components matrix: Marketing Comfort	
	Component
	1
MCAD	0,750
MCOM	0,888
MCCR	0,825

Table 5 Components matrix: Marketing Comfort

Components matrix: Behavioral Intentions	
	Component
	1
BIN1	0,809
BIN2	0,782
BIN3	0,695

Table 6 Components matrix: Behavioral Intentions

Components matrix: CFSMIP		
	Component	
	1	2
UAC1	0,751	-0,155
UAC2	0,857	-0,311
UAC3	0,872	-0,330
SUS1	0,783	-0,116
SUS2	0,835	0,013
SUS3	0,813	-0,312
COL1	0,431	0,716
COL2	0,484	0,713
COL4	0,690	0,198
COL3	0,610	0,351

Table 7 Components matrix after cleanup: CFSMIP

Using the retained items from the exploratory factor analysis, the average scores of each scale were computed for an overview of the engagement received by the respondents for each construct. The results of the regrouped variables are showcased in the table 8 below:

Descriptive Statistics			
	N	Mean	Std. Deviation
CFSMIP	101	6,0931	0,90280
Marketing Comfort	101	3,9967	1,38844
Behavioral Intentions	101	3,0000	0,92616

Table 8 Descriptive Statistics

The statistics showcased above can be interpreted as the overall level of agreement respondents had with the statements representing each scale, without taking into account the "ERR1", "ERR2", and "ERR3" items which were removed after the EFA. Firstly, an average of 6.09 means that the respondents were highly concerned about the different aspects measured of their information privacy on social media, encompassing the unauthorized access, secondary use, and collection of their personal information. Second, a 3.99 average for the marketing comfort scale reveals moderate comfort with the usage of the respondents' information for marketing purposes. Finally, an average of 3 for the behavioral intentions reveals moderately positive intentions towards revealing personal information on social media, considering that the latter construct was measured on a scale of 1 to 5. On another hand, we notice that the standard deviations for the CFSMIP and behavioral intentions scales are rather

low, which indicates a low variability of the responses among participants, opposed to the standard deviation exhibited by the marketing comfort scale.

To finalize the preliminary tests, a Pearson correlation test is conducted to determine whether a relationship between the variables exists. The results as presented in table 9 below, show the existence of a weak negative correlation between “CFSMIP” and “Behavioral Intentions”, but the correlation is not statistically significant (p-value<0.05). As for the variables “CFSMIP” and “Marketing Comfort”, the results show the existence of an almost negligible positive correlation between them, but this correlation is also not statistically significant. Finally, the results also show a meaningful moderate positive correlation between “Marketing comfort” and “Behavioral Intentions”, this correlation is statistically significant at the 0.05 significance level.

Correlations				
		CFSMIP	Marketing comfort	Behavioral intentions
CFSMIP	Pearson correlation		0,005	-0,145
	Sig.		0,958	0,149
	N		101	101
Marketing comfort	Pearson correlation	0,005		0,238
	Sig.	0,958		0,017
	N	101		101
Behavioral intentions	Pearson correlation	-0,145	0,238	
	Sig.	0,149	0,017	
	N	101	101	

Table 9 Pearson correlation table

3. Hypothesis testing:

To test our previously formulated hypotheses, we will use linear regression (multiple and simple) to confirm the causal relationships between the previously identified components. Regression analysis answers queries regarding a response variable's reliance on one or more predictors. Simple regression analysis is used to assess the relative influence of a predictor variable on a specific outcome (Weisberg, 2014). The extension of simple linear regression to include multiple explanatory variables is known as multiple linear regression (Tranmer, Murphy, Elliot, & Pampaka, 2020). When analyzing the results of a linear regression, numerous important factors must be considered. It is necessary to establish the statistical significance of the coefficients. To determine significance, the p-value associated with each coefficient is considered. A p-value less than a preset significance level (0.05 in our case) implies that the coefficient is statistically significant and that the association between the predictor and outcome variables is not attributable to chance. This will allow us to establish a link between the analyzed variables, which will confirm each of our hypotheses. Additionally, R-squared is used to measure how well the regression model fits the data. It represents the percentage of the variance in the result variable that the predictor factors can explain. The value of the R-squared ranges from 0 to 1, with a higher value indicating a better fit.

H1: A higher concern for social media information privacy will lead to negative behavioral intentions from social media users.

To test the first hypothesis, A linear regression model was used to explore the relationship between the dependent variable "Behavioral Intentions" (BIN) and various independent variables measuring concerns for social media information privacy (CFSMIP). The analysis aimed to determine the extent to

which these variables collectively influence behavioral intentions. The R-squared value of 0.383 indicates that approximately 38.3% of the variability in behavioral intentions (BIN) can be accounted for by the combined effects of the independent variables.

Regarding the analysis of variance (ANOVA), which measures the overall significance of the regression model, the sum of squares for the regression was 32.875, with 10 degrees of freedom. The resulting mean square was 3.288. The calculated F-statistic of 5.593 had an extremely low p-value (0.000), highlighting the statistical significance of the regression model. This indicates that at least one of the independent variables significantly impacts the dependent variable.

Examining the coefficients, the constant term of 3.598 indicates the estimated value of M_BIN when all independent variables are zero. Each independent variable's coefficient captures its individual effect on M_BIN while considering the influence of the other variables in the model. Notably, "SUS1," "COL2," and "COL3" exhibited statistically significant coefficients with p-values of 0.004, 0.043, and 0.000, respectively. These variables emerge as meaningful predictors of behavioral intentions.

The standardized beta coefficients provide insights into the strength and direction of relationships. Specifically, "COL2" (0.234) and "COL3" (-0.636) stand out with significant standardized coefficients. However, variables such as "UAC1," "UAC2," "UAC3," "SUS2," "SUS3," "COL1," and "COL4" yielded coefficients that were not statistically significant (p-values > 0.05). The results of this first linear regression are shown in appendix 3.

In summary, the regression outcomes affirm that **concerns for social media information privacy indeed influence negative behavioral intentions concerning information disclosure on social media**. Notably, "SUS1," "COL2," and "COL3" hold significance as predictors within the model. Furthermore, the overall model demonstrated statistical significance. We therefore accept the first hypothesis H1.

H2: A higher concern for social media information privacy will lead to less marketing comfort from social media users.

To test the second hypothesis, we rely again on a linear regression model designed to explore the association between the dependent variable "Marketing Comfort" and the same independent variables previously used to measure concerns for social media information privacy (CFSMIP). The primary objective is to assess the cumulative influence of these variables on marketing comfort. The R-squared value of 0.224 reveals that approximately 22.4% of the variability in marketing comfort can be explained by the combined effects of the independent variables.

The assessment of variance (ANOVA) serves to evaluate the overall significance of the regression model. The sum of squares for the regression amounts to 43.276, with 10 degrees of freedom. The mean square is computed at 4.328. Remarkably, the calculated F-statistic of 2.605 possesses a p-value of 0.008, signifying the statistical significance of the regression model. This implies that at least one of the independent variables has a substantial impact on the dependent variable.

Looking at the coefficients, the constant term of 3.488 signifies the projected value of marketing comfort when all independent variables are set to zero. Noteworthy, "COL3" and "COL4" exhibit statistically significant coefficients with respective p-values of 0.007 and 0.001. These variables emerge as significant predictors of marketing comfort.

The standardized beta coefficients offer insights into the magnitude and direction of relationships. Specifically, "COL3" (0.329) and "COL4" (-0.499) hold substantial standardized coefficients. On the other

hand, independent variables such as "UAC1," "UAC2," "UAC3," "SUS1," "SUS2," "SUS3," "COL1," and "COL2" demonstrate coefficients that lack statistical significance (p-values > 0.05). The results of the second linear regression are shown in appendix 4.

Summarizing, the outcomes of the regression analysis propose that **concerns for social media information privacy (CFSMIP) indeed play a role in forecasting marketing comfort**. Specifically, variables "COL3" and "COL4" are statistically significant predictors within the model. Furthermore, the overall model attains statistical significance. We therefore accept the second hypothesis H2.

H3: Marketing comfort mediates the relationship between concern for social media information privacy and social media users' behavioral intentions.

Finally, the third analysis conducted involves a linear regression model that analyses the connection between the dependent variable "Behavioral Intentions" and the independent variable "Marketing Comfort", in order to test for the third hypothesis of our study. The R-squared value, computed at 0.056, denotes that approximately 5.6% of the variability in behavioral intentions can be elucidated by the variance in marketing comfort.

Shifting focus to the analysis of variance (ANOVA), which measures the overall significance of the regression model. The sum of squares for the regression equals 4.843, corresponding to 1 degree of freedom. This leads to a mean square of 4.843. Importantly, the derived F-statistic of 5.924, coupled with a p-value of 0.017, underscores the statistical significance of the regression model. This reveals that the variable "Marketing Comfort" indeed wields a substantial impact on the dependent variable "Behavioral Intentions".

Next looking at the coefficients, the constant term holds a value of 2.367, representing the projected value of "BIN" when "MC" is set to zero. The coefficient attributed to "Marketing Comfort" is 0.159, signifying that a unitary increase in marketing comfort corresponds to a 0.159-unit elevation in behavioral intentions.

The standardized beta coefficient for "MC" is estimated at 0.238, indicating a moderate positive standardized effect. Further substantiating this relationship, the calculated t-score of 2.434, coupled with a p-value of 0.017, underscores the statistical significance of the association between marketing comfort and behavioral intentions. The results of this third linear regression are shown in appendix 5.

To summarize, the linear regression results align with the third hypothesis, indicating that **marketing comfort operates as a mediator in the relationship between concerns for social media information privacy (CFSMIP) and Behavioral Intentions**. Furthermore, the statistically significant relationship between marketing comfort and behavioral intentions implies that heightened marketing comfort is linked to more favorable behavioral intentions regarding information disclosure on social media platforms. We therefore conclude the hypothesis testing by accepting the third hypothesis of our study.

VI. Discussion:

The privacy landscape of social media has witnessed an unprecedented surge in concerns, stemming from data breaches, unauthorized sharing, and a lack of control over personal information. Integrating blockchain-based social media platforms as a solution to these challenges offers multifaceted benefits that resonate with the evolving demands of privacy-conscious users.

Through our study, we gained a clearer picture of the factors influencing individuals' decisions when it comes to sharing personal information on social media platforms. Our investigation into the

relationship between privacy concerns and behavioral intentions sheds light on a crucial aspect of modern online behavior.

The results revealed a significant connection between privacy concerns and users' intentions to share personal information. When people express higher levels of worry about their information's privacy, they tend to exhibit greater caution and hesitation in disclosing private details on social media. This cautious behavior suggests that users who hold stronger privacy concerns prioritize safeguarding their personal data and are less inclined to share it casually.

Furthermore, the study highlighted the role of users' comfort with marketing practices as a mediator in this relationship. This means that users' attitudes toward how marketing activities are carried out on social media platforms can influence how their privacy concerns impact their behavioral intentions. Users who are comfortable with the way marketing operates seem to navigate their privacy concerns differently, potentially considering sharing their information more willingly.

Considering these findings in a practical context, there are several actionable insights that can benefit both social media platforms and marketers. For platforms, the results underline the importance of enhancing users' sense of control over their data. By implementing user-friendly privacy settings and transparent communication about data protection measures, platforms can empower users to feel more secure in their interactions. Also, user education initiatives that clarify how data is handled, stored, and utilized can alleviate privacy concerns. This educational approach can foster a sense of trust and informed decision-making among users, contributing to more responsible and mindful sharing practices.

Extrapolating the findings of our study to real-world implications, a potential shift in users' attitudes towards data privacy can be foreseen. While a significant portion of respondents exhibited limited familiarity with blockchain's privacy benefits, a substantial proportion also expressed a willingness to consider blockchain-based platforms in the future. This suggests that the concept of blockchain as a solution to privacy concerns is gaining traction, and user education and awareness efforts could further catalyze its adoption.

The idea of users embracing blockchain-based platforms as they become more popular and developed is promising. Our study's results highlight the substantial percentage of users open to such a transition. This indicates a latent demand for privacy-centric platforms that align with users' evolving expectations.

Blockchain's decentralized architecture is pivotal in empowering users with enhanced data ownership and control. Our results illuminate that users desire greater control over their personal data, and blockchain aligns with this demand. The immutable nature of blockchain ensures that user data remains unalterable, promoting transparency and accountability, and addressing concerns related to data manipulation.

Moreover, blockchain platforms leverage encryption techniques, thereby fortifying data security. This is a key aspect of user trust, and our findings underscore the importance of data security in influencing behavioral intentions. The implementation of smart contracts further facilitates users' consent management, as users can stipulate the terms under which their data can be utilized.

The elimination of centralized intermediaries through blockchain technology is a game-changer. As revealed by our study, users' privacy concerns are exacerbated by intermediaries' involvement. Blockchain's decentralized approach reduces the risk of data breaches and unauthorized access, and our respondents' concerns find alignment with this model.

For marketers, the study emphasizes the significance of aligning marketing strategies with users' comfort levels. Recognizing that users' comfort with marketing practices plays a role in how privacy concerns influence behavior, marketers can tailor their approaches accordingly. Personalized, transparent, and respectful marketing efforts are more likely to resonate positively with users, building stronger connections and reducing any potential apprehension.

In essence, this study not only enriches our understanding of the intricate dynamics between privacy concerns, user behavior, and marketing comfort but also provides practical insights that can enhance the digital landscape. By incorporating these insights, both social media platforms and marketers can create environments that prioritize user trust, empowerment, and informed decision-making.

VII. Conclusion:

The digital age has allowed for unprecedented opportunities for connection, communication, and information sharing through the medium of social media. However, this evolution has also brought to light significant concerns surrounding the privacy of personal information in the social media realm. This thesis dove into the intricate relationship between users' concerns for social media information privacy, their behavioral intentions when faced with such issues, and the mediating role of marketing comfort. With a focus on the potential application of blockchain technology as a privacy-enhancing solution, the study aimed to shed light on the complex interaction between user behaviors and the evolving landscape of digital privacy.

Through rigorous quantitative analysis, this research has unveiled key insights into the relationship between concerns for social media information privacy and behavioral intentions. The exploration of privacy calculus theory illuminated the intricate balance individuals strike between perceived benefits and potential privacy risks when engaging with social media platforms. The findings substantiated the hypothesis that a higher concern for social media information privacy indeed leads to negative behavioral intentions among users, reflecting a growing awareness of the trade-offs inherent in online information sharing.

Moreover, the investigation into the mediating role of marketing comfort added a novel layer of understanding to the privacy-behavior subject. The study demonstrated that marketing comfort plays a significant role in shaping users' behavioral intentions, suggesting that users' perceptions of marketing practices contribute to their willingness to share personal information online. This mediation effect plays a key role in highlighting the interconnectedness of privacy perceptions and marketing comfort in influencing users' decisions within the digital ecosystem.

Furthermore, the research also highlights the potential of blockchain technology in addressing privacy concerns. While a considerable percentage of respondents exhibited limited familiarity with the privacy benefits of blockchain, the willingness expressed by a significant proportion to switch to blockchain-based platforms reflects a growing openness to technological solutions that prioritize individual data control and security. As blockchain technology matures and gains prominence, its potential to reshape the privacy landscape cannot be ignored.

In conclusion, this study has illuminated the complex interplay between concerns for social media information privacy, behavioral intentions, and marketing comfort. Empirical evidence reinforces the importance of striking a balance between user benefits and privacy considerations in the digital era. Furthermore, the exploration of blockchain technology as a privacy-enhancing solution holds promise for shaping the future of online interactions.

As the digital landscape continues to evolve, there remains much to be explored and understood. The insights gained from this research not only contribute to academic scholarship but also hold practical implications for fostering a more secure and user-centric online environment. As users become more discerning and privacy-conscious, the collective efforts of researchers, practitioners, and technology developers are instrumental in shaping a digital future that prioritizes individual privacy while harnessing the potential of digital connectivity.

VIII. Limitations and Future Research:

While this study has undoubtedly provided us with valuable insights, it is crucial to acknowledge its limitations. One of the primary limitations is the size and diversity of the group we studied. Our research sample was relatively small and consisted of individuals with similar backgrounds, which

might raise questions about how broadly our findings can be applied. To truly understand the whole picture, it would be beneficial to include a more varied and extensive group of participants in future studies.

Another important aspect to consider is the approach we took in our study. We heavily relied on quantitative analysis, which provided us with solid numerical data to work with. However, this also means that we might have missed out on some more nuanced details that could only emerge from in-depth conversations with participants. Exploring the qualitative side of user perceptions and experiences could uncover layers of understanding that numbers alone cannot capture.

As we look ahead, there are exciting avenues for future research to explore. One possible direction is to further investigate the intricate nature of "marketing comfort." Diving into the specifics of different types of marketing practices and how they impact users' comfort levels could offer a more comprehensive view of this concept. Understanding what kinds of marketing tactics make users feel at ease or uneasy could provide valuable insights for businesses striving to create user-friendly environments.

Additionally, exploring the cultural and demographic factors that influence users' concerns and intentions can bring to light promising results. Different cultures and backgrounds can profoundly shape people's attitudes towards privacy and online behavior. Exploring these factors could unveil a richer understanding of why individuals from diverse backgrounds might react differently to privacy concerns on social media.

In conclusion, while this study is an essential stepping stone in uncovering the complex dynamics between privacy concerns, user behavior, and marketing comfort, it is important to recognize its limitations. By expanding the research sample and delving into qualitative dimensions, we could obtain a more holistic understanding of these dynamics. Future research exploring various types of marketing practices and considering cultural and demographic influences promises to shed further light on this intricate topic, benefiting both academic understanding and practical applications in the constantly evolving digital landscape of social media.

IX. Executive summary:

Our study focuses on the complex interplay between privacy concerns, user behavior, and marketing comfort in the current digital environment. A survey was conducted with 103 participants to gain a better understanding of how individuals' concerns about privacy influence their behavior and attitudes on the internet. The findings of the survey provided valuable insights into this relationship.

Our research indicates a strong correlation: individuals who express worries about their privacy on social media platforms are more likely to exercise caution when it comes to disclosing personal information. Moreover, these concerns have an impact on users' comfort levels regarding marketing practices on these platforms, which in turn affects their level of engagement.

Our study reveals that individuals have different levels of familiarity with the potential of blockchain technology to improve privacy. Although there are some individuals who are not familiar with the advantages of blockchain-based platforms, a significant number of respondents demonstrate a willingness to adopt these platforms as they become more popular and advanced.

This study highlights the significance of user-centered approaches for organizations that operate in the digital landscape of social media. Despite some limitations, this study offers valuable insights that have important implications for both theoretical understanding and practical application. Additionally, our

research highlights the importance of balancing technological progress with protecting user privacy. This work provides guidance for a safer and more user-focused digital future.

Word count: 15074

X. Appendix:

Appendix 1: Questionnaire of the study

Welcome to our survey on Social Media Use and Privacy Concerns!

This survey aims to explore your perspectives on the use of personal data for marketing purposes on social media platforms. Your input will provide valuable insights into how users perceive this practice and its impact on privacy.

Rest assured that your responses will remain completely anonymous. Your personal information will be kept separate from your answers, ensuring your privacy is protected. Your participation is voluntary, and your candid feedback will contribute to a better understanding of the evolving digital landscape.

Thank you for taking the time to share your opinions with us. Let's begin the survey!

Do you use social media?

- ○ Yes (1)
- ○ No (2)

How many hours do you spend on social media every day?

- ○ Less than 1 hour (1)
- ○ 1-2 hours (2)
- ○ 3-5 hours (3)
- ○ More than 5 hours (4)

Which social media channels are you most active on?
Choose all the options that apply

- [ ] Facebook (1)
- [ ] Instagram (2)
- [ ] Twitter (X) (3)
- [ ] LinkedIn (4)
- [ ] Snapchat (5)
- [ ] Reddit (6)
- [ ] TikTok (7)
- [ ] Other (please specify) (8)

_____

What do you use social media for?
Choose all options that apply

- [ ] Entertainment (1)
- [ ] Developing your personal and professional network (2)
- [ ] Sharing photos and videos (3)
- [ ] Learning (4)
- [ ] Staying connected with friends and family (5)
- [ ] Business (6)
- [ ] Other (please specify) (7)
_____

How often do you post on social media?

- ○ Never (1)
- ○ Rarely (2)
- ○ Somewhat often (3)
- ○ Very often (4)
- ○ Always (5)

How private do you keep your information on social media?

- ○ Fully private (1)
- ○ Only my basic information is public (2)
- ○ Fully public (3)

To what extent do you agree with the following statements:

	Strongly disagree (1)	Disagree (2)	Somewhat disagree (3)	Neither agree nor disagree (4)	Somewhat agree (5)	Agree (6)	Strongly agree (7)
Social media databases that contain personal information should be protected from unauthorized access – no matter how much it costs. (1)	○	○	○	○	○	○	○
Social media sites should take more steps to make sure that unauthorized people cannot access personal information on their site. (2)	○	○	○	○	○	○	○
Databases that contain personal information should be highly secured. (3)	○	○	○	○	○	○	○

	Strongly disagree (1)	Disagree (2)	Somewhat disagree (3)	Neither agree nor disagree (4)	Somewhat agree (5)	Agree (6)	Strongly agree (7)
Social media sites should not use personal information for any purpose unless it has been authorized by the individuals who provided the information. (1)	○	○	○	○	○	○	○
When people give personal information to social media sites for some reason, these sites should never use the information for any other purpose. (2)	○	○	○	○	○	○	○

Social media sites should never share personal information with third-party entities unless authorized by the individual who provided the information. (3)	○	○	○	○	○	○	○

Page Break

To what extent do you agree with the following statements:

	Strongly disagree (1)	Disagree (2)	Somewhat disagree (3)	Neither agree nor disagree (4)	Somewhat agree (5)	Agree (6)	Strongly agree (7)
Social media sites should take more steps to make sure that personal information in their database is accurate. (1)	○	○	○	○	○	○	○
Social media sites should have better procedures to correct errors in personal information. (2)	○	○	○	○	○	○	○
Social media sites should devote more time and effort to verifying the accuracy of the personal information in their databases before using it for recommendations. (3)	○	○	○	○	○	○	○

Page Break

To what extent do you agree with the following statements:

	Strongly disagree (1)	Disagree (2)	Somewhat disagree (3)	Neither agree nor disagree (4)	Somewhat agree (5)	Agree (6)	Strongly agree (7)
It usually bothers me when social media sites ask me for personal information. (1)	○	○	○	○	○	○	○
It usually bothers me when social media sites ask me for my current location information. (2)	○	○	○	○	○	○	○
It bothers me to give personal information to so many people on social media. (3)	○	○	○	○	○	○	○
I am concerned that social media sites are collecting too much personal information about me. (4)	○	○	○	○	○	○	○

How comfortable would you be if information about you or posted by you publicly on social media is used for:

	Extremely uncomfortable (1)	Uncomfortable (2)	Somewhat uncomfortable (3)	Neither comfortable nor uncomfortable (4)	Somewhat comfortable (5)	Comfortable (6)	Extremely comfortable (7)
Targeted advertising (1)	○	○	○	○	○	○	○
Gathering user opinion about products or services (2)	○	○	○	○	○	○	○
Building customer relationships (3)	○	○	○	○	○	○	○

How likely are you, within the next three years to:

	Very unlikely (1)	Unlikely (2)	Neither likely nor unlikely (3)	Likely (4)	Very likely (5)
Provide personal information on social media sites (1)	○	○	○	○	○
Make your social media account public for others to find easily (2)	○	○	○	○	○
Share information on social media sites with your friends (3)	○	○	○	○	○

Are you familiar with the blockchain technology?

- ○ No (1)
- ○ Yes (2)

How familiar are you with the benefits of blockchain in terms of privacy?

- ○ Not at all familiar (1)
- ○ Slightly familiar (2)
- ○ Somewhat familiar (3)
- ○ Moderately familiar (4)
- ○ Extremely familiar (5)

How likely are you to replace social media sites you currently use with ones that offer a higher control over the privacy of your data? (Using the blockchain technology)

- ○ Extremely unlikely (1)
- ○ Somewhat unlikely (2)
- ○ Neither likely nor unlikely (3)
- ○ Somewhat likely (4)
- ○ Extremely likely (5)

Page Break

What is your gender?

- ○ Male (1)
- ○ Female (2)
- ○ Non-binary / third gender (3)
- ○ Prefer not to say (4)

What is your country of residence?

- ○ Belgium (1)
- ○ Luxembourg (2)
- ○ Netherlands (3)
- ○ France (4)
- ○ Other (please specify) (5) _____

In which range does your age fall?

- ○ Under 18 (6)
- ○ Between 18 and 24 (1)
- ○ Between 25 and 34 (2)
- ○ Between 35 and 44 (3)
- ○ Between 45 and 55 (4)
- ○ Above 55 (5)

What is your current occupation

- ○ Employed (1)
- ○ Unemployed (2)
- ○ Retired (3)
- ○ Student (4)

Appendix 2: Shapiro-Wilk test of Normality

	Shapiro-Wilk test of normality				
	Statistic	**N**	**Sig. (p-value)**	**Skewness**	**Kurtosis**
Concerns For Social Media Information Privacy					
UAC1	0,565	101	0,000	-2,913	9,34
UAC2	0,507	101	0,000	-3,383	13,014
UAC3	0,436	101	0,000	-3,281	10,27
SUS1	0,613	101	0,000	-2,68	8,764
SUS2	0,652	101	0,000	-2,074	4,005
SUS3	0,470	101	0,000	-3,479	12,884
ERR1	0,816	101	0,000	-1,051	0,677
ERR2	0,856	101	0,000	-0,792	-0,24
ERR3	0,877	101	0,000	-0,812	0,47
COL1	0,866	101	0,000	-0,901	1,304
COL2	0,833	101	0,000	-1,023	0,486
COL3	0,841	101	0,000	-0,98	0,118
COL4	0,735	101	0,000	-1,835	3,134
Marketing Comfort					
MCAD	0,918	101	0,000	0,232	-0,963
MCOM	0,931	101	0,000	-0,17	-915
MCCR	0,931	101	0,000	-0,355	-0,726
Behavioral Intentions					
BIN1	0,902	101	0,000	-0,23	-0,755
BIN2	0,838	101	0,000	0,432	-1,205
BIN3	0,841	101	0,000	-0,861	-0,139

Appendix 3: Linear regression between CFSMIP and BIN

Récapitulatif des modèles

Modèle	R	R-deux	R-deux ajusté	Erreur standard de l'estimation
1	,619[a]	,383	,315	,76669

a. Prédicteurs : (Constante), COL4, COL1, UAC2, COL3, COL2, SUS1, SUS2, UAC1, SUS3, UAC3

ANOVA[a]

Modèle		Somme des carrés	ddl	Carré moyen	F	Sig.
1	Régression	32,875	10	3,288	5,593	,000[b]
	de Student	52,903	90	,588		
	Total	85,778	100			

a. Variable dépendante : M_BIN

b. Prédicteurs : (Constante), COL4, COL1, UAC2, COL3, COL2, SUS1, SUS2, UAC1, SUS3, UAC3

Coefficients[a]

Modèle		Coefficients non standardisés		Coefficients standardisés		
		B	Erreur standard	Bêta	t	Sig.
1	(Constante)	3,598	,593		6,064	,000
	UAC1	-,105	,121	-,137	-,872	,386
	UAC2	,172	,210	,206	,819	,415
	UAC3	-,033	,196	-,045	-,166	,868
	SUS1	,331	,113	,408	2,919	,004
	SUS2	-,142	,115	-,186	-1,231	,221
	SUS3	,071	,154	,091	,463	,644
	COL1	-,116	,088	-,143	-1,321	,190
	COL2	,142	,069	,234	2,053	,043
	COL3	-,433	,072	-,636	-6,012	,000
	COL4	-,038	,080	-,060	-,472	,638

a. Variable dépendante : M_BIN

Appendix 4: Linear regression between CFSMIP and MC

Récapitulatif des modèles

Modèle	R	R-deux	R-deux ajusté	Erreur standard de l'estimation
1	,474[a]	,224	,138	1,28884

a. Prédicteurs : (Constante), COL4, COL1, UAC2, COL3, COL2, SUS1, SUS2, UAC1, SUS3, UAC3

ANOVA[a]

Modèle		Somme des carrés	ddl	Carré moyen	F	Sig.
1	Régression	43,276	10	4,328	2,605	,008[b]
	de Student	149,501	90	1,661		
	Total	192,777	100			

a. Variable dépendante : M_MC

b. Prédicteurs : (Constante), COL4, COL1, UAC2, COL3, COL2, SUS1, SUS2, UAC1, SUS3, UAC3

Coefficients[a]

Modèle		Coefficients non standardisés		Coefficients standardisés		
		B	Erreur standard	Bêta	t	Sig.
1	(Constante)	3,488	,998		3,496	,001
	UAC1	-,139	,203	-,120	-,684	,496
	UAC2	,017	,353	,014	,049	,961
	UAC3	,260	,329	,241	,791	,431
	SUS1	,020	,190	,017	,106	,916
	SUS2	,301	,194	,263	1,556	,123
	SUS3	-,260	,258	-,222	-1,004	,318
	COL1	,170	,147	,140	1,158	,250
	COL2	-,162	,116	-,178	-1,395	,166
	COL3	,335	,121	,329	2,768	,007
	COL4	-,464	,134	-,499	-3,468	,001

a. Variable dépendante : M_MC

Appendix 5: Linear regression between MC and BIN

Récapitulatif des modèles

Modèle	R	R-deux	R-deux ajusté	Erreur standard de l'estimation
1	,238[a]	,056	,047	,90417

a. Prédicteurs : (Constante), M_MC

ANOVA[a]

Modèle		Somme des carrés	ddl	Carré moyen	F	Sig.
1	Régression	4,843	1	4,843	5,924	,017[b]
	de Student	80,935	99	,818		
	Total	85,778	100			

a. Variable dépendante : M_BIN

b. Prédicteurs : (Constante), M_MC

Coefficients[a]

Modèle		Coefficients non standardisés		Coefficients standardisés	t	Sig.
		B	Erreur standard	Bêta		
1	(Constante)	2,367	,275		8,594	,000
	M_MC	,159	,065	,238	2,434	,017

a. Variable dépendante : M_BIN

XI. References:

Boyd, D. M., & Ellison, N. B. (2007). Social Network Sites: Definition, History, and Scholarship. *Journal of Computer-Mediated Communication*, 210-230.

Carpentier, C. (2023, March 23). *Attention Economy*. Retrieved from United nations: https://www.un.org/sites/un2.un.org/files/attention_economy_feb.pdf

Carr, C. T., & Hayes, R. A. (2015). Social Media: Defining, Developing, and Divining. *Atlantic Journal of Communication*, 46-65.

Chatterjee, R., & Chatterjee, R. (2017). An Overview of the Emerging Technology: Blockchain. *2017 International Conference on Computational Intelligence and Networks*, 126-127.

Cryptopedia. (2022, February 23). Retrieved from Gemini: https://www.gemini.com/cryptopedia/steem-blockchain-social-media-steem-crypto-social-media

Datta, A., Buchegger, S., Vu, L.-H., Strufe, T., & Rzadca, K. (2010). Decentralized Online social Networks. In *Handbook of social network technologies and applications* (pp. 349-378).

Desai, V., & Vidyapeeth, B. (2019). Digital Markeing: A review. *International Journal of Trend in Scientifc Research and Development*, 196-200.

Dixon, S. (2023, February 14). *Most popular social networks worldwide as of January 2023, ranked by number of monthly active users*. Retrieved from Statista: https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/

Edmondson, D. (2005). Likert scales: A history. *In Proceedings of the conference on Historical analysis and research in marketing*, 127-133.

Frankenfield, J. (2023, April 16). *Decentralized Applications (dApps): Definition, uses, pros and cons*. Retrieved from Investopedia: https://www.investopedia.com/terms/d/decentralized-applications-dapps.asp#:~:text=Decentralized%20applications%2C%20or%20dApps%2C%20are,built%20on%20the%20Ethereum%20platform.

Guidi, B. (2020). When blockchain meets online social networks. *Pervasive and Mobile Computing*, 101131.

Guidi, B., Conti, M., Passarella, A., & Ricci, L. (2018). Managing social contents in Decentralized Online Social Networks: A survey. *Online social networks and media*, 12-29.

Hayes, A. (2023, April 23). *Blokchain facts: what is it, how it works, and how it can be used*. Retrieved from Investopedia: https://www.investopedia.com/terms/b/blockchain.asp

Heale, R., & Twycross, A. (2015). Validity and reliability in Quantitative studies. *Evidence Based Nursing*, 66-67.

History of Blockchain. (2023, February 27). Retrieved from Geeksforgeeks: https://www.geeksforgeeks.org/history-of-blockchain/

Huang, Z., & Benyoucef, M. (2013). User-centered investigation of social commerce design. *Online Communities and Social Computing*, 287-295.

Iansiti, M., & Lakhani, K. R. (2017). The truth about blockchain. *Harvard Business Review*.

Isaak, J., & Hanna, M. J. (2018). User data privacy: Facebook, Cambridge Analytica, and privacy protection. *Computer*, 56-59.

Ivey, A. (2023, July 3). *5 blockchain-based social media platforms to know*. Retrieved from Cointelegraph: https://cointelegraph.com/news/5-blockchain-based-social-media-platforms-to-know

Jacobson, J., Gruzd, A., & Hernandez-Garcia, A. A. (2020). Social media marketing: Who is watching the watchers? *Journal of Retailing and Consumer Services*, 53.

Joshi, A. (2015). Likert Scqle: Explored and Explained. *British Journal of applied science & technology*, 396-403.

Jungwoo, L., Cheong, K., & Kun Chang, L. (2022). Exploring the personalization-intrusiveness-intention framework to evaluate the effects of personalization in social media. *International Journal of Information Management*.

Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. *Business Horizons*, 59-68.

Kapoor, K. K., Tamilmani, K., Rana, N. P., Patil, P., Dwivedi, Y. K., & Nerur, S. (2017). Advances in Social Media Research: Past, Present and Future. *Information Systems Frontiers*, 531-558.

Kenton, W. (2022, December 29). *What is Web 2.0? Definition, Impact, and Examples*. Retrieved from Investopedia: https://www.investopedia.com/terms/w/web-20.asp

Kietzmann, J. H., Hermkens, K., McCarthy, I. P., & Silvestre, B. S. (2011). Social media? Get serious! Understanding the fucntional building blocks of social media. *Business Horizons*, 241-251.

Koch, R. (n.d.). *Cookies, the GDPR, and the ePrivacy Directive*. Retrieved from GDPR: https://gdpr.eu/cookies/

Krishnamurthy, B., & Wills, C. E. (2009). On the Leakage of Personally Identifiable Information Via Online Social Networks. *Proceedings of the 2nd ACM workshop on Online social networks*, 7-12.

Laufer, R. S., & Wolfe, M. (1977). Privacy as Concept and a Social Issue: A Multidimensional Developmental. *Journal of Social issues*, 22-42.

Lulandala, E. E. (2020). Facebook data breach: a systematic review of its consequences on consumers' behaviour towards advertising. In *Strategic System Assurance and Business Analytics* (pp. 45-68).

Malhotra, N., Nunan, D., & Birks, D. (2017). *Marketing research: Ana applied approach.* Pearson.

Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system. *Decentralized business review*.

Obar, J. A., & Wildman, S. (2015). Social media definition and the governance challenge: An introduction to the special isse. *Telecommunications Policy*, 745-750.

Osatuyi, B. (2015). Empirical examination of information privacy concerns instrument in the social media context. *Transactions on replication research*.

Petrosyan, A. (2023, May 22). *Number of internet and social media users worldwide as of April 2023*. Retrieved from Statista: https://www.statista.com/statistics/617136/digital-population-worldwide/

Quan-Haase, A., & Young, A. L. (2010). Uses and gratifications of social media: a comparison of facebook and instant messaging. *Bulletin of science, technology & society*, 350-361.

Rautela, S. (2021). Social Media for New Product Launch: A Study of Social Media Platforms Across the RACE Planning Framework. *International Journal of Interactive Mobile Technologies*, 15(5).

Rohit, C. (2019). *DISSENTING STATEMENT OF COMMISSIONER ROHIT CHOPRA.* Washington, D.C: Federal Trade Commission.

Royle, J., & Laing, A. (2014). The digital marketing skills gap: developing a Digital Marketer Model for the communication industries. *International Journal of Information Management*, 65-73.

Sajithra, K., & Rajindra, P. (2013). Social media - History and components. *Journal of Business and Management*, 69-74.

Smith, J. H., Milberg, S. J., & Burke, S. J. (1996). Information Privacy: Measuring individuals concerns about organizational practices. *MIS quarterly*, 167-196.

Srimani, P. K., & Srinivas, A. (2011). Behavioral Targeting- Consumer Tracking. *AIP conference proceedings*, 56-60.

Stewart, K. A., & Segars, A. H. (2002). An empirical examination of the concern for information privacy instrument. *Information systems research*, 36-49.

Tranmer, M., Murphy, J., Elliot, M., & Pampaka, M. (2020). Multiple Linear Regression (2nd Edition). *Cathie Marsh Institute Working Paper*.

Van Dijck, J. (2012). Facebook as a tool for producing sociality and connectivity. *Television and new media*, 160-176.

Weisberg, S. (2014). Applied linear regression. *Chichester: Wiley-Blackwell*.

Who Controls the Internet? (n.d.). Retrieved from internethealthreport: https://internethealthreport.org/v01/decentralization/

Wolf, M., Sims, J., & Yang, H. (2018). Social media? What social media? *UK Academy for Information Systems*.

Zyskind, G., Oz, N., & Pentland, A. (2015). Decentralizing privacy: Using blockchain to protect personal data. *2015 IEEE security and privacy workshops*, 180-184.