

Copyright
by
Soyoung Park
2021

**The Dissertation Committee for Soyoung Park Certifies that this is the approved
version of the following dissertation:**

**Data Rights in the 21st Century: Exploring the Boundaries of
Empowerment in Blockchain Social Media Settings**

Committee:

Sharon Stover, Supervisor

Kenneth Fleischmann

Joseph Straubhaar

Kathleen Tyner

**Data Rights in the 21st Century: Exploring the Boundaries of
Empowerment in Blockchain Social Media Settings**

by

Soyoung Park

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August 2021

Dedication

To my loving parents and sisters.

Acknowledgements

The completion of this work was only possible thanks to several people's guidance, encouragement, and support, to which I want to express my gratitude. I would first like to express my sincere appreciation to my committee members, Professor Sharon Strover, Professor Kenneth Fleischmann, Professor Joseph Straubhaar, and Professor Kathleen Tyner, for their valuable guidance and mentorship throughout the entire process of my dissertation research work.

I am extremely grateful to Professor Sharon Strover, my supervisor, for her valuable advice, extensive knowledge, and relentless support and care to accomplish this huge task. I am deeply indebted to her throughout the course of my academic journey at the University of Texas at Austin. Under her mentorship, I have been provided with thoughtful guidance and countless learning opportunities while working with her and many outstanding colleagues over the years at the Technology and Information Policy Institute, where she co-directs. She is incredibly insightful and logical and has an ingenious theoretical mind, and these qualities tremendously helped me to improve my dissertation research project further.

I would also like to thank Professor Kenneth Fleischmann, who inspired me to broaden my perspectives in research with his expertise in information studies specialized in social informatics. His invaluable insight into society and information technology helped me enrich my understanding of the social implications of emerging technologies. To Professor Joseph Straubhaar, I express my gratefulness for his deep theoretical insights and constructive discussions and suggestions. His advice was immensely helpful for me to

refine the discussion and interpretation of some findings presented in this dissertation. I also very much appreciate Professor Kathleen Tyner for encouraging me to incubate the early idea of this dissertation. She was always willing to provide constructive criticism and guidance that enabled me to overcome challenges in shaping the dissertation research.

It is my fortune to gratefully acknowledge the support and encouragement received from my fellow and senior colleagues at Austin. Big thanks to Ji Won Kim, Claire Shinhea Lee, Hogeun Seo, Jennifer Minsoo Kang, Jinsook Kim, Yoonmo Sang, Jaewon Royce Choi, Sooyeon Hong, and Hyun Jung Stephany Noh. I also had the great pleasure of working with Alexis Schrubbe, Richelle Crotty, Maria Skouras, and Melissa Santillana; my thanks are extended for their cooperation and support.

Finally, words would never say how grateful I am to my parents for their selfless love and unconditional support. They have always been there for me whenever I needed them. Their faith in me has encouraged me to pursue what I valued and desired. My deepest gratitude should go to them from the bottom of my heart. My special thanks also to my sisters for their wholehearted care, affection, and continued encouragement, which contributed a lot to the completion of my dissertation.

Abstract

Data Rights in the 21st Century: Exploring the Boundaries of Empowerment in Blockchain Social Media Settings

Soyoung Park, Ph.D.

The University of Texas at Austin, 2021

Supervisor: Sharon Strover

Personal data based on daily human activities in digital domains are becoming an asset accelerating the unprecedented growth of digital industries. Individual data subjects, tired of businesses using and selling user data for financial gain, are now more aware of the value of their data and are increasingly demanding a fair share of the value they create/contribute. Despite recent policy efforts in support of the public claim such as data dividends, data tax, and data fiduciaries, there are no easy or “perfect” solutions; multiple solutions have catalyzed multiple debates. The goal of this study is to broaden our horizons on this issue by investigating the potential of an incentivized blockchain-powered social networks platform to exercise our data rights. Drawing on Steemit, one of the first blockchain social networks as a site for research, this study explores how ordinary users’ ownership of data as well as data privacy are implemented and practiced in the form of user-generated contents (UGCs), how was their “investment” of personal information

actually rewarded, and the sociotechnical conditions of successful users on this type of the alternative platform to realize the value of our data.

Using a topic modeling approach, content analysis, and an online user survey, the study takes both macro and micro perspectives to examine UGC behaviors that employ personal information as a source of content, in addition to exploring individual users. The Steemit site essentially rewards people for sharing personal information, providing a unique platform with which to assess how people create value from UGC. For example, users' self-introduction posts created an extensive topic space that generally contains basic demographic information. However, a considerable proportion of topic space focuses on expressing who participants are by disclosing their personal traits, views of life, life goal or vision rather than conventional demographic profiles. This implies an increase of "less profile-able or less quantifiable" personal information. Comparing the topics over time periods when cryptocurrency values increased and then diminished illustrates the way that context may affect content creation, since self-disclosure dropped off when the value assigned to cryptocurrency dropped.

From a micro view of UGCs, a qualitative analysis of the user posts reveals how the posts weaving personal details earned more incentives than those without. It also shows that users have implemented privacy-aware identity strategies, characterized by the wide separation of personal and professional identities in online space as well as information balancing activities. Using selfie postings restricts self-disclosure in forms other than selfie itself.

Regarding the socio-technical context, the relationship between digital competency and using Steemit was analyzed using a square structural equation modeling based on user survey data. Results demonstrate the critical role of creative digital competency that is positively and significantly affected by web 1.0 operational skills and web 2.0 social skills as well as a sense of tech-dependence. Importantly, unemployed and people from less developed countries were likely to be marginalized on this platform, suggesting another form of digital inequality operating in this environment.

These findings highlight the potential of the decentralized rewarding platform that sheds light on the multifaceted role of users who embody the traits of natural private persons, content creators, and gig workers. When given the opportunities to invest not only skills, knowledge, and strengths, but also identity and experiences to yield rewards, participants were willing to share personal information but also exhibited privacy-aware behaviors. The study contributes to research on users' online privacy practices by looking at how people embed information about themselves in their self-creations beyond the profile page. It contributes to the field of digital literacy by demonstrating the structure of technological embeddedness intertwined with information privacy concerns for online creativity as a source of value. This study discusses the possibilities of this alternative technology-based solution to empower ordinary users to realize their own data rights in the contemporary data regime.

Table of Contents

List of Tables	xiv
List of Figures	xv
Chapter 1: Introduction	1
THE STATEMENT OF THE PROBLEM: DATA ECONOMY	1
PURPOSE, SCOPE, AND SIGNIFICANCE OF THE DISSERTATION	5
DISSERTATION ORGANIZATION.....	8
Chapter 2: Literature Review and Theoretical Framework.....	14
THEORY OF DATA OWNERSHIP	14
A Brief History of Data Rights for Individuals	14
<i>Personal Data Commodification</i>	18
<i>Gaps in the Public Law and Privacy Commons Perspectives</i>	20
Empowerment through Data Ownership Rights: Privacy as Property Revisited	22
CHANGING CONTEXT: DATA CONTRIBUTION AS RESOURCE INVESTMENT	28
The False Tradeoff in Data Transaction	28
Data Contribution in Context.....	31
SOCIOTECHNICAL CONDITION OF DIGITAL LITERACY: REALIZING DATA VALUE.....	34
Governing Architectures of Technology and Human Agency	34
Competence and Dependence as Embedded in Technology	36

Awareness of Privacy and Data Rights.....	39
Summary	41
Chapter 3: Methodology.....	43
RESEARCH SETTING: STEEMIT	46
The Corpus.....	49
Four Phases of STEEM.....	50
METHODS	54
Analysis Plan	54
Definitions and Operationalization.....	55
Chapter 4: Data as Investment: Macro View.....	58
STUDY 1 METHOD: TOPIC MODELING.....	58
Data Collection	60
Data Analysis Tools.....	61
Data Preprocessing	62
Implementing LDA Model	63
STUDY 1 RESULT: SPACE OF PERSONAL INFORMATION INVESTMENT	64
Comparison of Topics in User Information Investment Between Mania- and Blow-Off Phases	66
Value and Volume of Information Space During and Post Bubble.....	71
STUDY 1: CHAPTER DISCUSSION AND CONCLUSION	74

Chapter 5: Data as Investment: Micro View	81
STUDY 2 METHOD: CONTENT ANALYSIS	81
Data and Methods	81
STUDY 2 RESULT: STRATEGIC ASSEMBLAGE OF SELF IN USER-GENERATED CONTENT	85
Separation of Professional and Personal Identities.....	90
Identity Integration Practices	93
Visual Representations of Self: Information Balancing Practices.....	96
Various Identity Practices and Realized Values	97
STUDY 2: CHAPTER DISCUSSION AND CONCLUSION	100
 Chapter 6: Technological Condition for User Success	104
STUDY 3 METHOD: ONLINE USER SURVEY	105
Data Collection	105
Survey Instruments	106
Measuring User Performance	109
Data Analysis.....	110
STUDY 3 RESULT: TECHNOLOGICAL CONDITION FOR USER PERFORMANCE	112
Descriptive Statistics.....	112
Measurement Model Assessment	114
Structural Model Assessment	117
STUDY 3: CHAPTER DISCUSSION AND CONCLUSION	126

Chapter 7: Discussion and Conclusion	128
REFRAMING DATA OWNERSHIP AND CONTROL IN REVOLUTIONARY DIGITAL SPACE	128
Value Space for the “Qualified Self”	130
Agency in Action	132
Creative Competence as a Key Driver of Data Competence	133
Limitations and Suggestions	135
IMPLICATIONS OF THE STUDY	136
Social Media User Studies: Identity, Privacy, and Value	136
Creative Competence and Digital Divide	137
Methodological Implications	139
Practical Implications	141
CONCLUSION	144
Appendices.....	146
Appendix 1.1: Representative posts for topic spaces during Mania phase (Chapter 4)	146
Appendix 1.2: Representative posts for topic spaces during blow-off phase (Chapter 4)	153
Appendix 2: A sample of Steemit user recruitment document (Chapter 6)	157
Appendix 3: The table of Heterotrait-Monotrait ratio (HTMT) Results (Chapter 6)	158
References	159

List of Tables

Table 1: List of Python-based tools used	61
Table 2: Extracted topics, keywords, and description for M- and B-corpus	70
Table 3: Posting and curating rewards realized in M-corpus	73
Table 4: Posting and curating rewards realized in B-corpus.	74
Table 5: Coding scheme.....	85
Table 6: Coding results.	89
Table 7: Identity management practices and realized value in STEEM currency	98
Table 8: Measures of technological embeddedness.....	108
Table 9: CFIP scale (modified from Bellman et al., 2004).....	109
Table 10: Demographic information.....	113
Table 11: Measurement model results.	116
Table 12: Path results.....	118
Table 13: Effect of control variables.	121
Table 14: Path results for privacy constructs	124

List of Figures

Figure 1: Steemit home page that displays the categorized posts.....	50
Figure 2: Phases of a bubble. Jean-Paul Rodrigue (2008).....	52
Figure 3: Steemit price chart.....	53
Figure 4: Basic workflow of LDA topic modeling.....	60
Figure 5: Intertopic distance map for M-corpus (left) and B-corpus (right). PC: principal component.....	67
Figure 6: The composition of the most relevant terms for Topic 1 of M-corpus (left) and B-corpus (right).....	67
Figure 7: Change the information space from the Mania to the Blow-off phase	72
Figure 8: Conceptual path model.....	111
Figure 9: Base path model result.	120
Figure 10: Path model expanded with privacy construct of improper or unauthorized access.	125

Chapter 1: Introduction

THE STATEMENT OF THE PROBLEM: DATA ECONOMY

Collecting and using our data have grown at an unprecedented rate in recent years. OECD estimates that the global volume of digital data will multiply by a factor of 40 by the end of the current decade (Reimsbach-Kounatze, 2015). Supported by advancements in information and communication technology (ICT) and the Internet of Things (IoT) rich environments, people leave countless digital footprints with everything they do online, which then becomes converted into commodity values posited by average companies that produce, manipulate and sell that data (Benkler 1999; Samuelson, 2000; Zuboff, 2015, 2019).

In the era of ubiquitous connectivity where a wide range of daily human activities in digital domain are continuously datafied (Galanxhi & Nah, 2006), data itself is becoming an asset that accelerates the exponential growth of digital economy. This reifies the value of data as a base of information and knowledge, and positions data as a resource that needs to be managed not only for government and for-profit organizations but also for individual and non-profit entities.

With this growing value of data in mind, one logical question is: how do data subjects exercise ownership rights over the data they themselves generate? Indeed, it is not hard to trace an individual data subject's *contribution* across each byte of data; as sources, people contribute personally identifiable information (PII), web cookies, browsing, and location history through their many forms of online activities, and even contribute diverse types of shared or created content regardless of a range of intellectuality or creativity. A massive dataset reflecting one's own actions can be thought of as an expression of

intangible individual investment and intellect (Reichman & Samuelson, 1997) or as a product of “labor” (Couldry & Mejias, 2019; Terranova, 2000; van Dijck, 2009). Given this context, it is reasonable that we humans, as *producers*, should be entitled to ownership rights for our data, rights that encompass not just the “ability to access, create, modify, package, or remove data, but also that strongly assure the right to “*derive benefit from or sell* data, or assign these access privileges to others” (Loshin, 2002, p.31). This is the issue that animates this dissertation.

Individuals’ data ownership rights have not been sufficiently addressed by the current legal and regulatory frameworks to date. In fact, individual data ownership has not yet been explicitly legislated in the U.S. (César, Debussche, & Asbroeck, 2017; Glancy, 2010). According to U.S. law, personal data is not yet a category of intangible intellectual property (Glancy, 2010); personal rights to data have been subjected to an incomprehensible “patchwork of narrowly focused sectoral laws and self-regulatory rules” (Schwartz, 2012, p.1623), and in many cases, data produced online are collected through strong-armed end user license agreements that remove users’ control and rights. Similar problems are mirrored in EU law, in which the concept of “ownership” of data is never explicitly defined (César, Debussche, & Asbroeck, 2017; Janecek, 2018) – even in the General Data Protection Regulation (henceforth GDPR), the strongest data protection regulation in the world. Indeed, while copyrighted content has been widely scrutinized at the heart of intellectual property (IP) rights in a range of past media distribution networks and relevant legal discussions, whether “the law ought to grant individual property rights in their personal data” remains under-examined (Samuelson, 2000, p.6; Lessig, 1999, 2002). An exception can be found only in copyrightable databases or software using the

principle of “investment” in the *sui generis* regime (Boyle, 1996; Reichman & Samuelson, 1997).¹

Throughout most legal history, data for individual subjects is somehow relegated to being a matter of privacy and security (Wiebe, 2016; Wilks & Christie, 2013) rather than a matter of property and ownership. That aura of privacy protection remains precarious, as evidenced in several recent data breach crises and in the common “consent without consent” modes using clickthrough agreements that have become a preferred market solution to obtaining permission to use individual data (Samuelson, 2000). The repeal of the Broadband Consumer Privacy Proposal in the U.S. on March 28, 2017, that required Internet Service Providers (ISPs) to get consumers’ permission before collecting, sharing, and selling their data, implies even more severe threats to individuals (Naylor, 2017). Recently, there has been considerable focus on strengthening privacy and data rights legislation, evident in the EU GDPR, the California Consumer Privacy Act of 2018 (CCPA), or the California Privacy Rights Act of 2020 (CPRA), to facilitate users’ actual control over their data. While these actions have indeed stimulated other US states such as New York (2019), Virginia (2021), Colorado (2021), and Washington (2021) as well as other countries such as Brazil (2020) to consider and enact more comprehensive and stronger data privacy laws, we still need to wait and see how these actions will actually affect end users with intended/ unintended or beneficial/ damaging consequences.

Amidst these challenges, industry and government partnerships strengthened in order to fully exploit the economic value chain of privatized and commoditized information

¹ e.g., European Commission’s Directive on the Legal Protection of Databases in 1996 and the US’s H.R. 354, the Collections of Information Antipiracy Act of 1999, and H.R. 3872, Consumer Access to Information Act of 2004. The EU Directive was aimed at compensation for database creator’s “substantial investment” in either the obtaining, verification or presentation of data; US legislations protect any collection of information through the “investment of substantial monetary or other resources.”

(Schneier, 2015). The symbiotic cooperation between industrial and governmental sectors, driven by the long-ingrained spirit of market liberalism and national advantages, has always played a winning game with data. We even see individuals blamed for the harm suffered in data breaches because of their ‘careless technological practices’ (Lipton, 2010) or their inherent desire to share (Zuckerberg, 2009), instead of the careless or inadequate practices of platforms gathering personal data. As well, the fairness of the “deal” between data compilers and data subjects is easily dismissed by claiming the fairness of the tradeoff between data privacy and the potential benefits one receives by sharing it (Acquisti, 2004; Acquisti & Grossklags, 2005; Wright, Camp, Goldberg, Rivest, & Wood, 2002).

Consider the aforementioned irony in the current value chain of data: data becomes one of the most valuable resources in the present time and data are never depleted, creating endless possibilities for companies, while an individual data subject can rarely find a way to evaluate or claim her ownership and fair share of everyday data contributions. Recently, many policy proposals have surfaced to address the issue. Some argued that people should have a choice to get paid or “opt out” (Newsom, 2019; “Data Dividend Project”, Yang, 2021), while others suggest that big tech firms could be taxed on the data they collect and use (Thimmesch, 2016). Another approach is the “data fiduciary” model, in which a company that is considered fiduciary by law has obligations of confidentiality, care, and loyalty to its users (Balkin, 2016; 2020; Balkin & Zittrain, 2016; Whitt, 2019). However, these multiple solutions have been facing even more intricate issues and further debates that prevent them from actually being implemented. Facing this conundrum, this study explores the potential of a technology-based solution through an incentivized decentralized social networks platform in which one can realize the value of one’s private data as a means of self-empowerment. This inquiry will only become more important as the issue of data

rights becomes more controversial in many parts of the world: data contribute to a growing and powerful industry, an industry that sometimes commits injustices against unwitting data contributors.

PURPOSE, SCOPE, AND SIGNIFICANCE OF THE DISSERTATION

The quest of this dissertation is to investigate the potential of a technology-based solution to empower end users to easily exercise their data rights and realize the value of their data through dynamics of users' personal data practices. Its primary questions begin with the historical treatment of personal data that can be reconfigured in present digital spaces where one might actually use personal data in their self-creations online. After problematizing the commodification of personal data within commercial digital industries, it examines a blockchain-based social media platform, Steemit, as an alternative model to empower individuals' data ownership rights by realizing the value of their data – their contribution, creation, and commitment. Steemit is an experimental platform designed to reward every user with a type of cryptocurrency, Steem, for each person's contribution to the platform. Although a young and highly uncertain service, Steemit stands in as a system that yields some proxy data valuations to individuals. This study aims to grasp the potential of this incentivized social media networks to drive individuals' conscious personal data “investment” with respect to the exercise of data ownership and data privacy rights.

My lens focuses specifically on the situation of personal data ownership and privacy management. According to some internationally accepted guidelines, personal data or personal information refers to “any information relating to an identified or identifiable natural person” (‘data subject’) (The EU Directive, Art.2 (a); the GDPR, Rec.26; Art.4 (1))

and to “personally identifiable information” (PII) or “sensitive personal information” (SPI), which generally includes an indefinite list of individually identifiable information that can be used by itself or in combination with one or other elements of information to distinguish or trace the individual’s identity (Stevens, 2012, p.6; the CPRA, Cal. Civ. Code 1798.140(ae)(1), (2)). Other major developed countries share a similar understanding of personal information.

The scope and meaning covered by the above definition vary in context. This study is especially concerned with personal data as represented in individuals’ ‘digital identities’ reflecting their personal traits, preferences, experience, and daily life, not limited to the traditional set of Personally Identifiable Information (PII) such as name, address, or telephone number; date and place of birth; mother’s maiden name; Social Security number or other government-issued unique identification number; biometric data; or unique account identifiers among many others. This interpretation of personal data is referred to in the CCPA of 2018 as “personal information” with reference to “a broad list of characteristics and behaviors, personal and commercial, as well as inferences drawn from this (personal) information (e.g., biometric data, household purchase data, family information, geolocation, financial information and sleep habits)” (AB 375, 2018). As such, this study broadly considers personal information, in keeping with the spirit of recent legislation.

Situating Steemit as a site of research, this study references technology studies that address the politics in technology as imposed through data architectures appearing in routinized day to day lives. I argue that technologies can quietly and unobtrusively regulate one’s attitude, behavior, and actions, including creating and sharing UGCs online (Lessig, 2006) Various methodological approaches will be used here to investigate the multifaceted

value-making UGC practices using personal information. As well, my argument will address the socio-technical environment as a brake or alternatively as an accelerator for involvement in the data value chain.

The current study makes scholarly contributions in the following ways. First, it expands notions of individual sovereignty and ownership to personal information beyond the passive discussion of privacy and security risks. While this ethos of stronger protection that emphasizes one's safety and security is necessary to safeguard our data rights from unruly exploitation to a certain extent, one main argument of this study problematizes how this prevailing protectionism of the public law approach may blindside us from fully exercising our data rights. Second, this study investigates the implications of a decentralized incentive design as an alternative model for data value realization by exploring and advocating for a wider range of user contributions to reside in the sphere of personal data. Most research endeavors on the subject of the value chain of personal data have treated the end users as a mere "consumer" or "victim." Facing the reality of a data ecosystem that benefits only a few shareholders at the expense of almost everyone else, this study envisions the potential of an alternative model that attempts to democratically distribute the wealth of data for empowering users' rights; individuals can become their own center as independent *investors* who control their (private) data to invest for profit or refuse to sell or share. Relatedly, to date no study has been conducted to explore users' personal data management practices in the unique form of decentralized data economy, represented here by Steemit. Thanks to the blockchain technologies and the advent of cryptocurrency as one tangible currency asset, a user can obtain value for contributions based on other users' appreciation of her data offered, a kind of market. The findings examine how such context affects one's managing privacy and identifies in online space.

Taken together, the study discusses the hype and hope of emerging blockchain technologies. Although this might be provocative or premature, this framework could represent a new approach to reform the current status quo of the data economy in order to facilitate a fairer distribution of data wealth. Lastly, this study moves towards a more comprehensive understanding of individuals' personal data practices within the construct of *technological embeddedness*. Using the notion of embeddedness, the extent to which one assembles her life experiences at the intersection of social structures and technological infrastructures, this study not only advances a more systematic understanding of socio-technical influences on our perception towards personal data but also examines the condition of individuals' successful performance in this incentive system.

Overall, this dissertation seeks to show how individuals might recalibrate their rights to data by adjusting themselves in this novel system of data valuation. By exploring personal information practices and the economic benefits (and some noneconomic benefits) those users receive from a wide range of everyday contributions to data and by demonstrating situated socio-technological surroundings that might enhance or diminish our rewards for data contributions, this study is an opening shot in addressing the limitations of the current state of individual data ownership rights. It ignites more questions and discussions among “the data subjects” to challenge and reform the reigning discourses around data governance.

DISSERTATION ORGANIZATION

The dissertation comprises the following three sets of theoretical and empirical investigations around the power struggles over individuals' data rights: the first

conceptually sets the stage for Steemit, an alternative social media environment, as a revolutionary digital space that reframes data ownership rights; the second explores the practice of using the personal data of ordinary people at the boundary between alternative value-making/value realization and privacy protection; and the third looks at the role of socio-technical conditions that explain individuals' performance in the current new breed of a decentralized data economy. This series of inquiries is driven by the following research questions.

1. What are the problems in contemporary discourse around personal data rights? How have individuals' data rights been defined and addressed in the existing policy and regulatory arenas? How do they frame individuals' own powers with respect to personal data?

The literature review in Chapter 2 examines this question.

2. When compensation is assured, how do people invest in or get rewarded by information about themselves through self-created content?
 - 2.1 What types of personal information are voluntarily disclosed and shared? How does each type of the information relate to realized value?
 - 2.2 How does one strategically manage identity through divulging personal information? How do the characteristics of managed identities relate to the realized value?

Scrutinizing users' behaviors in terms of what content they share and how they share it in the context of a unique social media platform provides responses to these questions (Chapter 4 and Chapter 5).

3. How do people's technological embeddedness affect their value-making performance on Steemit?

3.1 How do technological competence and dependence affect value-making performance?

3.2 How do technological experiences and capabilities affect value-making performance in relation to awareness of data rights?

A survey of users provides information regarding relationships among competence, tech dependency, privacy concerns and sharing behaviors (Chapter 6).

Chapter 2 reviews major theories and literature to illuminate the elements of personal data rights. It introduces primary conceptual constructs around three key themes – revisiting property rights as a vital motivator of exercising data ownership, investing personal data and identity as resources for self-created content, and the role of socio-technical structures in data and privacy activities and outcomes. It leads to research questions around data ownership that explore the blurred boundaries of personal data as core resources of the decentralized data economy (empirically investigated in Chapter 4 and Chapter 5) and the mechanisms of competence and dependence within socio-technical structures associated with data practices (empirically investigated in Chapter 6).

To frame the state of data ownership, this study uses the lens of privacy theorists and intellectual property theorists, relying on work primarily published in the legal discipline. We examine the potential of a property approach for negotiating the power balance between data generators and data compilers. With the onset of the Internet and a more broadly networked environment as well as the digitization of many expressions of intellectual property, IP law and policy have profoundly shaped the debates around the value of data. Taking these perspectives together, this study criticizes “value gaps” between data aggregators and individual data contributors by problematizing the blind spots in ongoing data rights discourses where individual users (i.e., data contributors) might only be seen as vulnerable, innocent, and in need of protection. Additionally, the study briefly introduces recent efforts that support stronger personal data rights, even though they face controversy or challenges, or are still in the incubation stages. This provides a context that emphasizes the need and goals of full ownership of data to users, as well as the possibilities of new decentralized reward systems as a complement to bridge the value gap in the current data economy.

In line with the above, the research postulates personal data as a type of resource that users can mobilize. Here, the boundaries of personal data become a subject of inquiry in the view of privacy economics approaches. While a growing number of studies has raised the rationale of “privacy tradeoffs” and one’s negotiation between risks (or costs) and benefits in disclosing personal information, this research suggests one missing factor in current assumptions of the decision-making process: the direct and tangible economic reward value one receives for her contribution to a platform, thus positioning the user as an actual stakeholder in the data value chain who mobilizes data ownership, data privacy, and digital identities.

The last inquiry revolves around the notion of digital literacy within the architecture of technology in our everyday lives and its relationship to data ownership rights. This study deliberately chooses the term *embeddedness* to suggest a perpetual process in which individuals are becoming ingrained within systems of information and technology and conceptualizes our everyday experiences with technology as including both a sense of competence and also a dependence, whereby individuals utilize skills and information to meet life goals. The role of user awareness of privacy and data rights is also discussed in association with the notion of technological embeddedness, considering its impact on actual use behaviors and incentives for users.

With this theoretical backdrop and central questions, Chapter 3 establishes a research design focused on the critical case selection of Steemit, and details the data gathering methods, and analysis plans. In addition to offering more details of Steemit to justify its selection, it outlines three sets of studies. To examine the use of personal data in value-making practices and personal information and identity management strategies, topic modeling and content analysis approaches are applied based on the user-generated posts on Steemit in Chapter 4 and Chapter 5, respectively. To investigate the role of digital literacy based on the surrounding technological architecture, an online survey method is deployed in Chapter 6. The results analyze users' background information such as personal traits or socio-demographic attributes as well as the extent of their embeddedness within technologies and concerns for personal data. The collected online survey data is paired with the same user's activity record: the number of posts generated per month, number of followers, reputation score, and account value estimated in dollars. This approach allows an analysis of socio-technical conditions around users' performance based on data contributions.

Chapter 7 summarizes and discusses the results of the empirical findings. It aims to assess the potential of the decentralized incentive design of Steemit for personal data empowerment, including the strengths and weaknesses of the current system design in empowering people to exercise ownership and control over their personal information. To conclude, the study discusses practical and theoretical implications of this study aligned with policy recommendations towards empowering data subjects in their navigation of the data economy.

Chapter 2: Literature Review and Theoretical Framework

THEORY OF DATA OWNERSHIP

In this section, I review the major theories and concepts around individual rights to personal data, primarily from legal theories of privacy. I focus especially on those that are useful in the context of the changing information market and that recognize altered social and technological circumstances reflected in a data-intensive economy. After a brief description regarding how personal data rights have been constrained, I trace the process of information commodification linking the nature and transformation of privacy rights and IP rights to the new order of the digital. This review shows how commodification contributes to an uneven playing field for individual data subjects. In contrast to the public law and “privacy commons” approach that have been forwarded to grapple with the problems of the digital information market, I argue the limitations of public law and privacy commons approaches in exercising *real* property rights of individuals over their personal data; that approach has been stifled by potential social dilemmas and asymmetrical power dynamics within the personal information data market. A property rights approach for personal data is further discussed and elaborated as fundamental to this study. In doing so, the study also briefly outlines some recent policy proposals that share the same or different views, what they promised and why they are challenged, opening up room to discuss the technical means to enhance individuals’ data rights.

A Brief History of Data Rights for Individuals

“Data is the new oil” is now becoming a common phrase. Despite its economic potential, private data is ignored in the current legal regime and has not been framed in a way to grant individuals property rights. Although copyrighted content has been widely

scrutinized at the heart of IP rights in a range of past media distribution networks and relevant legal discussions, whether “the law ought to grant individual property rights in their personal data” is contested (Samuelson, 2000, p.6; Lessig, 1999, 2002). In fact, personal data is not yet a category of intangible intellectual property in United States law (Glancy, 2010). It is only databases that are entitled to constitute intellectual property with the acknowledgement of “investment” (Reichman & Samuelson, 1997) within both the European Commission’s Directive on the Legal Protection of Databases in 1996 and certain bills considered in the U.S. (e.g., the US’s H.R. 354, Collections of Information Antipiracy Act of 1999 and H.R. 3872, Consumer Access to Information Act of 2004). But this was only the *sui generis* case.

For individual data contributors, ownership rights to data have never been legislated. In response to the question of *ownership* to personal data, the US developed an uneven “patchwork system” in the same way that European Commission had reservations about the data ownership concept, even in the most recent GDPR regulation (César, Debussche, & Asbroeck, 2017; Janecek, 2018; Schwartz, 2012). Duch-Brown, Martens and Mueller-Langer (2017) state that “the GDPR deliberately does not consider full and transferable private ownership rights for personal data.... on the basis of human rights arguments” (p.16). That is, there is not yet a specific data-related law that explicitly recognizes ownership of personal data.

Rather than granting ownership rights that involve exclusive control over property, personal data rights have long been discussed only as a matter of *protection* in legal history (Wiebe, 2017; Wilks & Christie, 2013), particularly with respect to upholding claims for constitutional rights such as privacy and security. While being aware of the potential threats such as data loss or theft, identity theft, and phishing or hacking attacks among many other

cybercrimes, the dominant motives have been to protect one's data through the right to not be tracked, the right to remain anonymous, and the right to be silent at one's will. Even considering some historic milestones around human rights, data, and information,² it is striking that these decrees and discussions have emphasized protection of inalienable and non-tradable specific rights rather than addressing full and transferable individual ownership rights, i.e., enclosing property rights, over personal data (Duch-Brown et al., 2017).

Setting aside the argument over the effectiveness of a protectionist approach with regard to personal data, there is growing criticism over regulatory and technological architectures of personal data protection. From a regulatory point of view in the US, Sotro & Simpson (2015) point out the limits of the sector-by-sector legislative frameworks that often are narrowly tailored and address specific data uses. That sectoral approach also suffers from a dearth of strong empirical detail on the broader operation of data markets, and highlights the limits of industry self-regulation. Hirsch (2010) criticized the issue of a "privacy self-management" approach that corresponds to a "consent dilemma" upon customers (Solove, 2012). Contractual practice that requires companies to acquire consumers' permission or agreement to handle their personal data was initiated to resurrect the bargaining power of individuals over which personal data to reveal to which firm for what purpose (Samuelson, 1999), but now it arguably remains nothing more than a superficial gesture, with an element of "take-it-or-leave-it" in the privacy architecture of the system, a point criticized by Cohen (1998) and by Lessig (1999). The recent US repeal

² For example, the European convention on human rights (1970), the US Code of Fair Information Practices (1973), the US Family Educational Rights and Privacy Act (FERPA) of 1974, the OECD Privacy Framework (1980), 1995 EU Data Protection Directive, the US Children's Online Privacy Protection Act (COPPA) of 1998, the right to be forgotten debate since 2006, and the "Do Not Track" debate in 2010 and the heated discussion over the US Internet Privacy rules in 2017.

of Internet Privacy rules exempting Internet service providers (ISPs) from the obligation to get consumers' explicit consent before collecting, using, and selling users' information for their own commercial purposes in 2017 strongly repudiates even this established structure.

The technological mechanism of personal data protection also has revealed fragility and defects. So-called Privacy Enhancing Technologies (PETs), which embraces the broader range of privacy support and data protection technologies, were once thought to be a panacea for identity management, data anonymization, privacy preserving data processing techniques and so forth (Shen & Pearson, 2011). Past experiences, however, demonstrate how personal data, even that collected by highly regarded companies with advanced technologies, can be vulnerable to malicious attempts to breach the information system. Moreover, the power of the “default setting” in system and application design has remained problematic because it may induce users to accept a displayed ‘choice’ without knowing the vulnerabilities it conveys. Vaidhyathan, (2012) offers extensive documentation regarding the problems of default settings in software design and in broader platform operations. Worse, the growth of online social networks implies a more critical context where our friends could also disclose our personal information (Sarigol, Garcia, & Schweitzer, 2014).

As the commodity value of data accelerates, there is little individuals can do to realize the value their data contributes, irrespective of its origin. Grounded in this context, this section further considers how individuals have been losing ground through data commodification trends that intersect IP rights discourses, and how the value question of individuals' personal information has been left unaddressed in the dominant protectionist rhetoric. To this end, this chapter revisits the “privacy as property” debate in order to set

up a foundation for an empowerment framework of data ownership rights.

Personal Data Commodification

The process of personal data commodification dates from the earliest statements of privacy rights. To uncover the changing attributes of personal data, it is helpful to depart from the origins of privacy that appeared in Samuel Warren and Louis Brandeis' (1890) emblematic article, "*The Right to Privacy*." Their article attempted to highlight the fundamental conception of privacy rights as a link between the "right to be let alone" and the value in "preventing" publication (pp.193-215). They tried to secure a space for individuals to think and act freely without fear of being censored, monitored, or simply getting embarrassed in public. Warren & Brandeis (1890) are usually credited with the first formal pronouncement of an individual right to privacy, and subsequent law addressed personal information in terms of four torts of invasion of privacy: 1) intrusion into one's private life, 2) public disclosure of private facts, 3) false light and 4) appropriation of name or likeness (Prosser, 1960). Among others, the tort of appropriation is strongly associated with commodification of individuals in regard to commercial appropriation of images or aspects of the self (Epstein, 2018).

As such, this notion of privacy has been widely accepted not only for its acknowledgement of fundamental human rights of liberty, autonomy, and dignity (Schneier, 2015) but also for its recognition of the rights of "intangible" human products such as one's image in a photograph, which had been neglected at that time. In this early framework of privacy, personal information was hardly a subject of "transaction." In addition, every literary expression in any form was within the scope of protection, and a relative number of publishers made handling such rights fairly manageable.

The threat to this benign notion of a positive liberty theory of privacy (Boyle, 1996) became more evident after half a century, as reflected in Prosser's (1960) four torts of privacy, as noted earlier. While the first three torts of privacy still can be interpreted in the fundamental rights framework of privacy, the fourth tort of privacy, i.e., appropriation of name or likeness, was responding to the very moment where privacy and personal information could be subject to monetization such as through publishing one's photograph in a newspaper. The appropriation issues show that people of that time were aware of the possible harm from the use of an individual's personal information for commercial purposes without that person's permission.

Intellectual property laws transformed such rights as an “amalgam” between property and sovereignty, evoking the romanticized notion of the author to develop ideas of possessive individualism and original creativity (Aoki, 1996; Boyle, 1996, 2008). In the 20th century – particularly in the final decades – IP laws and theory evolved in the face of new elements of the digital economy when “constructed consent” and “manufactured scarcity” began to take shape in the very design of information products (Cohen, 1998). The growing intersection between “discrete and expressive” works of authors and “fungible and commensurable” bytes of underlying information appropriate to the new realm of software and databases extended the notion of data's value to a commercial setting (Reichman & Samuelson, 1997). For example, US copyright protection expanded to databases, framed as “compilations” for their originality (499 U.S. 340, 1991). Further, the copyright notion of authorship encompassed both property and sovereignty, complicating the boundary between contributing to the public domain (one fundamental idea in the creation of copyright) and the private domain of economic remuneration. The authorial regime privileged private owners of copyrighted works and pitted them against the rights

of ‘passive ‘consumers who would interact with these works. Individuals gradually were left powerless, as the property regime undervalues both raw and uncopyrighted sources for producing intellectual property (such as common cultural expressions) as well as the rights of audiences or citizens who might wish to interact with that expression on their own terms. With the overly romanticized figure of the author, the ownership of information itself became naturalized in the 19th and 20th centuries. The role of a single individual person as a reader/audience member or user was then relegated to an undervalued public domain.

The nature of both privacy rights and IP rights has mobilized the ideas of “relativism” and “contextuality” (Aoki, 1996; Boyle, 1996, 2008; Nissenbaum, 2004; Solove, 2002). Both terms signal how data have value only in certain contexts. That is to say that value depends on many other situational factors; in this view, data value is not absolute in any intrinsic way. This contextual quality allows powerful information intermediaries and data aggregators to easily assign value to certain forms of information products, e.g., collected or compiled data, in an existing commercial market while belittling the production of individual data generators. As Castells argues, a transference of power and rights to the “information capitalist” is the logical outcome (Castells, 1998; Halbert, 2000; Schneier, 2015; Vaidhynathan, 2012). It is analogous to the audience commodity theory, which argues that audience labor often can be unacknowledged and subsumed within the capitalistic media industries (van Dijck, 2009; Caraway, 2012). The emergence of information markets creates an uneven playing field for individual data subjects.

Gaps in the Public Law and Privacy Commons Perspectives

Building on the weak history of data rights for individuals, dominant discourses and policy endeavors of the 20th and 21st centuries have failed to safeguard data rights for

individuals. I first compare the regulatory frameworks of the US and EU by looking over some landmark legislative and other directives associated with personal data, pointing out the inherent foundations of the constitutional rights to personal data in privacy, security, and human dignity domains. This connects to the public law approach that upholds public benefits in shared information and the “privacy commons” position that conceptualizes privacy as a common pool of knowledge in opposition to the common law property and contract rules that highly favor protecting IP rights holders (Cohen, 1998; Sanfilippo, Frischmann, & Standburg, 2018). This impoverished ethos for data as public goods also is reflected in the fact that there has never been a very explicit recognition of the very concept of data “ownership” in the US and EU law for individuals (César, Debussche, & Asbroeck, 2017; Glancy, 2010; César, Debussche, & Asbroeck, 2017; Janecek, 2018). Some strong public law approaches in regulating data rights for individuals have appeared under the banners such as “human flourishing,” which pursues human progress through diversity and participative pluralism in order to achieve more abundant cultural production in the public sphere (Benkler, 2006; Cohen, 2012; Taylor, 2016) and using data for social justice (Kroll, Barocas, Felten, Reidenberg, Robinson, & Yu, 2017; Johnson, 2014, 2016; Newell & Marabelli, 2015; Heeks & Renken, 2016; Raymond, 2016). However, one might be skeptical of the sufficiency of those two positions in terms of their abilities to generate usable (and shareable) data and in terms of data rights. Several scholars have been hesitant to acknowledge the value of data as public goods because of the social dilemmas accruing to undertaking collective actions with unclear reward structures and the absence of an obvious incentive mechanism for producing social goods (Schwartz, 2004; Fairfield & Engel, 2015). While examples such as Wikipedia may be emblematic of the social goods perspective, it is striking that so few similar examples actually occur. My concern further

intersects with the asymmetrical power distribution that pits commercial interests on the side of the participating stakeholders and non-monetized social interests on the opposite side (see Soltani quoted by Feiner, 2018; Fairfield, 2017; Wilka, 2018). Lessig, an influential legal scholar, clearly captures this problematic mechanism of power in systems of information, calling attention to individual versus collective positions: “with copyright, the interests threatened are powerful and well organized; with privacy, the interests threatened are diffuse and disorganized” (2006, p.200).

The present snapshot of individual privacy rights in the United States, as we have seen in the defeated 2017 internet privacy rule discourse, illustrates how easily the “diffuse and disorganized” role and power of individuals can be challenged and compromised against consolidated institutional controls. A significant feature of the environment we analyze is precisely this disaggregated role for the individuals generating data.

Empowerment through Data Ownership Rights: Privacy as Property Revisited

The idea of property rights as a motivational basis for individual action deserves second thoughts. When we frame the issue of personal data only as a matter of privacy, it is difficult to see its economic/financial value. Personal data ownership cannot be assumed without granting property status to personal information that not only addresses one’s ability to access, create, modify, and package, but also confers the right to derive benefit from, *sell* or remove data, and even assign these access privileges to others (Loshin, 2002).

In fact, there is nothing new in this way of framing. Giving individuals property rights in their personal data has been considered one promising option in order to locate control in the end user and away from the entity that creates, aggregates and circulates data.

However, there has also been a conundrum that led many privacy theorists to doubt and criticize the feasibility and desirability of a property viewpoint.

According to renowned IP and privacy law scholar Pamela Samuelson (2000), the key mechanism of property law, “free alienability,” i.e., transferability, will not be aligned easily with personal information as it favors the context that allows buyers to freely transfer whatever they acquired from the initial seller. Samuelson further argues that the allocation of scarce resources, another justification for property rights, also seems less applicable to the case of personal data since such data is already abundant and ubiquitous. Professor Jessica Litman (2000) expressed similar concerns that a property approach might open up a possibility in which facts can be privately owned and be subject to restricted use, thus ending up causing friction with the First Amendment and free speech interests. In a similar vein, such data handling might facilitate the personal data market rather than constrain it, and thereby reinforce and perpetuate economic inequality and exploitation of personal data (Cohen, 2012). Therefore, many theorists conclude that a property approach is likely to conflict with information privacy goals.

From another point of view, however, there also has been a continual defense of privacy as property rights, for property is the most effective concept to encapsulate the “control of information concerning an individual's person” (Murphy, 1996, p.2381). Daniel Solove, a legal scholar well known for his work on privacy, also comments that John Locke and Alan Westin, known pioneers of privacy, defined the notion of privacy as an extension of property – one can have property rights “in their person and the fruits of their labor” or “as the right of decision over one's private personality” (2002, p.1112). Solove himself also justified personal information as property as an extension of personality and further emphasized the “instrumental” value of privacy:

Several privacy scholars who claim that privacy is valuable in itself locate the source of the value in a form of respect that must be provided to all rational beings.... However...., I contend that privacy has an instrumental value – namely, that it is valued as a means for achieving certain other ends that are valuable (ibid. p. 1145).

One body of research attempts to relieve these concerns. In response to the potential for a violation of free speech, Lessig made clear that even if facts can be subject to copyright regulation as in compiled databases, that does not mean one cannot control the use or dissemination of that fact, as seen in the examples of trade law and contract law. Lessig went on to state the somewhat exaggerated fear that financial value might encourage more efficient trading in the information market, remarking that the very same notion of property rights would empower individuals to refuse whichever deal they are unwilling to make. Not only that, under this approach individuals benefited from their different valuation of privacy. Indeed, “the advantage of a property system is that both of our wishes (e.g., never sell, or willing to sell access to the specific segment of personal information) get respected, even though the wishers are so different” (Lessig, 2002, p.262).

The proponents of a property approach in personal data do not completely deny the downsides that might be caused by the private information trade nor the unique “privacy boundary” of individuals. Back to Murphy’s economic defense for privacy, the core basis acknowledges dynamic benefits to privacy beyond individual “taste” for privacy and considerations of the nature of the transactions and the nature of the information in applying a disclosure default rule (1996, p. 2383). Most importantly, the proponents of a property approach make a point for unlocking possibilities. Paul Schwartz (2004), a leading

information privacy law scholar, for example argues that it is still uncertain that the presumed harms will surpass the benefits from greater individual participation in personal data trade. Schwartz references the transition of English society from the feudal period to the modern times, and remarks on the way in which the passing of the power and control of traditional aristocracy occurred with the marketization of lands.

Within rights-based approaches, there are several models to fashion a propertied personal information market. Continuing from Schwartz (2004), a data property market could be stabilized with some critical background and regulatory conditions consisting of limiting individuals' right to alienate personal information; allowing a right of exit in the market; default disclosure of the terms of trade; setting the standard to assess caused damage; and establishing the policing institutions for market. Law and media scholar Lauren Scholz (2016) invites us to consider privacy as "quasi-property" that determines the extent and scope of corresponding rights with reference to a relationship between individual; this view is in line with Solove (2012), who underscored the need to value privacy not in a general and abstract way, but contextually, again echoing Nissenbaum (2004).

In the regulatory sector, a handful of legislative acts have moved toward articulating methods for controlling data. The recent progressive arguments and legislations include the EU GDPR, known for the strongest data protection provision ever created, that stipulates the data rights of the data subject as a natural person, not a 'consumer,' across dimensions of access, use, and control of data. These rights include breach notification, right to access, the right to be forgotten, data portability, privacy by design and the designation of a data protection officer. Similarly, the CCPA and CPRA (2018, 2020) affirm Californian consumers' right to opt out of businesses selling or sharing their data as

well as of automated algorithmic profiling; the right to delete/correct their collected data; the right not to be discriminated against on the basis of personal data; the right to data portability, in addition to being fully informed of collected data, including the logic for data processing and decision-making.

Accordingly, recent actions have focused on implementation – how to achieve desirable outcomes in line with these strengthened data rights. Some highlight the need for wealth distribution, considering the basis of personal data in property. In this vein, the initiative of a “data dividend” argues that big tech companies should pay a dividend for their data-driven revenue, either to residents suffering from growing income inequality (Daniels, 2019) or to all Americans who have not received a “fair share” as stated by former presidential candidate Andrew Yang (Yang, 2021). Other ideas focus on mechanisms to better regulate the data market to protect individual’s data rights and to create public benefits. One initial approach is taxation: at a national or international level, tax instruments on businesses’ monetization of personal data can encourage or penalize the market to further protect individual privacy interests (Thimmesch, 2016). Legal scholars Balkin and Zittrain (2016) proposed the possibility of “data fiduciaries” between data collecting companies and end users. This aims to give a legal incentive to companies in handling user data within obligations of confidentiality, care, and loyalty by emphasizing their relationships of trust and confidence with their clients. In comparison, there is a “data trusts” model that puts third-party intermediaries in the position of data trustees, like a trade union, which is considered to be better able to negotiate data rights of groups of individuals as beneficiaries (Ruhaak, 2021).

Despite the promising outlooks each of these proposals illustrates, there have been criticisms around the uncertainty of the value of personal data, lack of uniform standards

in expanding the scope of data for regulation, and the limits of various theories' capacity to resolve practical problems. Many of these ideas are in the conceptual stage awaiting further experimentation. Hence, rather than jumping to judgment on each proposal, this study seeks to draw attention to one missing piece of the puzzle: the role of individuals: what should, or can the individual end users, the data commoners, do to safeguard and enhance our rights? While many assume the right to control personal data is getting much stronger, most expect only bigger, pre-organized hands; in that vision, individuals can become a force only by joining a group of "trustees." Currently, that perspective is a fairly realistic account of the reality. Given the countless ways that our data may be collected and processed beyond our expectations, or the limited time, resources, or rationality to research and assess our rights and risks, we need collective strength. Yet if we cannot be convinced that we ourselves can make meaningful change on an individual level, how can this effectively motivate us?

We indeed need a force to represent us in this fight for data that is in the firm grip of powerful tech titans, but that does not mean we, as individuals, do not need to know how to decide or control our own data issues in our everyday encounters. In this regard, the emergence of decentralized reward platforms provides us opportunities to directly recognize economic and instrumental value of our data by earning some tangible rewards for what we post. It could be less than a dime for a post; or it could be a dollar or even a hundred in rare cases. No matter how small (or big), these incremental experiences of achievements can boost our motivation. This can motivate us to continue to create and share new things and be willing to experiment with the extent of personal data we can or cannot invest in so that we have a greater sense of self-efficacy in consciously managing our digital presence. What starts here could have the power to make a difference.

CHANGING CONTEXT: DATA CONTRIBUTION AS RESOURCE INVESTMENT

This section challenges the dominant economic arguments and empirical market setting over “privacy tradeoffs” by examining research on the viability of a personal data calculus with the inherent limits of asymmetric personal market conditions. It reviews how personal data, discrete bits of information, became mobilized for value creation and how units of data can be extracted from an extensive range of contextualized user-generation in digital space. Contemplating the value of a fundamental human contribution to a vast pool of personal data, this study argues the growing role of individual agency balancing benefits and risks in the data value chain, especially in the context of an alternative data economy based on decentralized rewarding systems.

The False Tradeoff in Data Transaction

There are benefits as well as risks when we share our data. This “tradeoff” argument frequently arises when people become annoyed by businesses’ excessive data collection and use. There is no free lunch; when we use the “free” digital services we still have to pay something. However, individuals do not always have to be on the losing side.

Needless to say, whether intended or not, trading one’s personal data for products and services can bring benefits to the engaged individual subjects. People can enjoy various services that promise much more convenience in their lives in exchange for giving their private information. For instance, once they allow companies to track and record their previous browsing history or shopping list, they can be provided with more specialized shopping recommendations with discount offers. This logic can spread into other service areas beyond mere shopping, thereby ensuring a more convenient and efficient user experience.

Legal and cultural studies examining this tradeoff of personal data are rooted in ideas of privacy, security or the collective public. Research argues that resulting benefits of the data tradeoff for individuals are not restricted only to the personal sphere; it can benefit the public sphere as well. Most studies have employed normative reasoning, archival research and in-depth ethnographic investigation of some exemplar cases to document how certain aggregated data can be mobilized for useful public purposes or social benefits. For example, the aggregation of personal data through countless trades can lead to the dissemination and distribution of intellectual work linked to innovation (Reichman & Samuelson, 1997), social welfare (Cohen, 1998), and the exchange of ideas and collaboration (Benkler, 2006; Cohen, 2012); it also becomes a precursor to “human flourishing” (Benkler, 2006; Cohen, 2012; Taylor, 2016) that calls for nurturing individual potential to the fullest for a more abundant cultural production in the public sphere.

From the economics and business disciplines, research on privacy tradeoffs generally examines the economic and social benefits from personal data and privacy exchanges based on one’s risk-benefit calculation (Culnan & Armstrong, 1999; Donnenwerth & Foa, 1974; Li, 2012; Barth & De Jong, 2017; Awad & Krishnan, 2006; Acquisti, 2004; Acquisti, Taylor, & Wagman, 2016). The basic premise of the privacy calculus assumes one’s intention to disclose or exchange personal information by weighing perceived benefits against risk probability. If perceived benefits outweigh risks, one is likely to disclose personal information in exchange for social or economic benefit. Under this rationale, many studies on the value of personal data are subject to mathematical modeling, experiments, panel studies, or consumer surveys (Egelman, Felt, & Wagner, 2013; Nolte, 2015; Rayna, Darlington, & Striukova, 2015; Spiekermann et al., 2015).

It is evident that there are tradeoffs between privacy and (public) utility in using online services, yet it is hard to say whether this tradeoff is fair. Most studies addressing the privacy tradeoffs confess the difficulty of measuring the value of personal data due to context-dependence, contingencies, and human heuristics and bias (Acquisti, 2004; Berthold & Böhme 2009; Spiekermann, et al., 2015). One inherent gap in research may be caused by the actual market setting, where information control is inherently asymmetric between platforms, data brokers, and consumers in the first place. That asymmetry – the uncertain value of personal data and how it is collected, processed, used, and shared – hinders the individual’s ability to have full bargaining power in data trades with service providers. In this context, “information buyers” and “information sellers” can never have an equal chance to succeed in the personal data market, not even remotely close.

Moreover, the value of personal data is contextual as much as is the value of privacy, meaning that each study provides only a partial account of the problem. More often than not, experimental economic research on privacy tradeoffs tends to overlook the context, for example, the initial intention behind individuals’ desire to be involved in the trade, or the extent of one’s cumulative investment in certain digital spaces; in a controlled setting, some designate a predetermined value, e.g., a dollar value, or set of choices for privacy and personal data and ask their research subjects to accept these or not, or to choose one thing over another (Grossklags & Acquisti, 2007; Lesk, 2012). Such assessments are not suitable for addressing the condition that may facilitate one’s deliberate data behaviors on the basis of free will.

The current study’s examination of a decentralized rewarding platform has the merit of constituting this specific niche. By examining autonomous individual agency represented by people actively engaging in directly exchanging (personal) information and

receiving returns on it, the case of Steemit provides an opportunity to observe how one's personal information and identities as resources can actually realize value through ongoing data trade processes.

Data Contribution in Context

Despite all the concerns and cautions around the commodification of personal data, the data market has been growing rapidly in recent years (Spiekermann, Acquisti, Böhme, & Hui, 2015). The business implications of using personal data are indeed enormous – in terms of both relevant players and the types of business strategies used (Christl, 2017). Oracle, one of the worlds' largest data cloud and platform service companies, claims that it will provide “the richest understanding of consumers [across] the entire consumer experience” (Oracle, 2015, p.9). Here that consumer experience includes behavioral data (what consumers do), social data (what consumers say), and purchase data (what consumers buy) based on one unified addressable consumer profile across all devices, screens, and channels. Shoshana Zuboff (2019) also clearly detailed how Google exploits the “behavioral surplus” of consumers not only to improve their service but also to “feed” their machine intelligence in order to build momentum in new behavioral future markets. Zuboff even coined this “surplus” as “the new means of production.”

The data-driven marketing endeavor occurs not only through gathering traditional, personally identifying information. It also includes the data by-products grounded on individual online activities. Those data are collected by automatic means, including our online social interactions with others and even multiple forms of user-generated content such as photos and messages. Since each interested party now seeks to derive benefits from almost every single data unit out of one's life, contextualization becomes the key: Alissa

Lorentz, VP of a Cloud-based Data Platform Augify, states how crucial the context is to transform meaningless data into real information for actionable insights and intelligent decision-making; according to her, the key is a holistic and interpretive lens (Lorentz, 2013). It seems obvious we can better understand one person not by having just her purchase record only but also by linking that to an accompanying dataset including her behaviors and details of social interaction and other background information.

As more and more personal information contributes to contextualization, what becomes the subject of tradable assets is self, echoing Locke's notion of "one's person," and Westin's "private personality" (Solove, 2002). The value of data is rooted in both its intrinsic and added value (Loshin 2002; Zuboff, 2019). Personal data has been upgraded mostly through "added values," a phrase noted in the EU legislation as a database creator's "substantial investment" in obtaining data, verifying or presenting it, or it is recognized as "creativity" in the U.S. copyright framework.

Acknowledging the significance of 'added value' for appreciating the worth of personal data, this study also argues for the contribution of the *intrinsic* value of data; this deserves to be more carefully considered as it is represented in current forms of personal-branding, self-promotion, and identity and lifestyle constructions in digital spaces. The intrinsic value of personal data engages a broader process of self and identity that is always complex and laborious on the part of the actor (Davis, 2014). Defining the self as both a performance and an exhibition, and content as reflecting and affecting processes of performances and exhibitions (Hogan, 2010), the intrinsic value of personal data goes beyond one's personality; it also entails her intellectuality, creativity as well as labor to perform, exhibit and manage her own identity through various content creation and continuous social interactions (Terranova, 2000; van Dijck, 2009).

A multivariate concept of user agency has been observed in dimensions of content production and data generation in addition to consuming content (van Dijck, 2009). Consider UGCs on social media that display identity and daily life: with self-taken photos, videos, other multimedia elements, or lengthy accounts, many people strive to construct identity in a form of storytelling or narrative (Davis & Weinshenker, 2012; Briganti, Varriale, & Mele, 2021; Panahi, Watson, & Partridge, 2012). More and more individual contributors exhibit or enmesh a variety of “personal” or “professional” self by using information regarding their taste, status, lifestyle, experience, attitude, and group membership in addition to conventional demographic details in their created content (Batenburg, & Bartels, 2017). Data creation and value include expression to further enhance (sharing positive aspects only) or verify oneself (showing negative side of self in addition to positive aspects). Enacting identities online becomes a matter of *how* people post it, more than what types of information they post (Livingstone, 2008). These efforts for various self-expression, combined with unique originality, clearly showcase the creator’s investment and creativity. The whole display and achievement of selfhood can be the factor that other community members value. (In the Steemit context, they even vote on it.)

Additionally, all human beings care about at least a certain amount about privacy. Given this, the working logic of decentralized rewarding systems — direct beneficial rewards based on other users’ votes for user contribution — questions how value intersects one’s privacy practices in identity expressions. On one hand, sharing personal information can be a form of excessive or inappropriate self-disclosure. On the other hand, it constitutes another aspect of “creativity” self-representation by involving strategic decision-making to manage and share certain details of private information (Jeong, & Coyle, 2014; Jeong, &

Kim, 2017; Jin, 2013) while making a post interesting and engaging. The current study encapsulates the extensive performance of self in online space at the intersection of managing identities, privacy, and value-making practices.

SOCIOTECHNICAL CONDITION OF DIGITAL LITERACY: REALIZING DATA VALUE

This section uses the construct of competence and dependence in order to expand on how the architecture of contemporary technology interweaves and penetrates one's life. Against this backdrop, personal agency theory and the media and information literacy literature aid in hypothesizing the effect of technological embeddedness on various individual data practices and data valuation processes. It draws on the broad socio-technical premise of embeddedness, "the mutual constitution of people and technologies" (Sawyer & Jarrahi, 2014, p.3), as a grounding theory. This allows us to consider the role of information privacy perception not only as it relates to privacy concerns but also as it reflects important perceptions of data rights.

Governing Architectures of Technology and Human Agency

Although data ownership by individuals is only loosely regulated under the law and policy, there are several real-world factors affecting our relationships with technology-based systems. The central idea of Lessig (1998), "*code* is law," highlights the role of technology as "regulating" or conditioning one's use of data or services in addition to the other three factors he discusses – actual law, markets and social norms. He uses the term code, i.e., the hardware and software that comprise cyberspace, for the architecture of technology that functions as a main but often invisible constraint in our environment (Lessig, 1998, 2006; Scharf, 2012). For instance, in the context of this study, decentralized

social media systems using blockchain technology can be a “medium” (McLuhan, 1964) that shapes and controls the scale and structure of user networks and activities or an “affordance” that refers to design aspects or properties of artifacts that suggest how an artifact might or should be used (Moreno & D'Angelo, 2019).

Aligned with this framework, technology in this study can be interpreted as artifacts and systems as well as “processes that bring them into being.” This definition embraces creativity involving human ingenuity, individual and collective, to create and control a human-built world (Huges & Huges, pp.3-4). While Lessig’s argument tends to reify the notion of code by claiming the effects of technology can *determine* freedom and rights (Hosein, Tsiavos & Whitley, 2003), I prefer to emphasize the “co-production” of knowledge and social order, based on the relationship between technology and society (Jasanoff, 1996). Similarly, Hood (1994, p.2) offers a complementary reading of code by examining the regulatory changes enacted by humans such as the power of interest groups, the power of ideas, and transformative social development, alongside the technology. On top of technological impact, the conception of social forces that shape individuals’ uses of technology and the perception and valuation of their personal data also are important to understand the operation of human agency (Bandura, 2002).

This does not argue the superiority of one factor over another, but rather strikes a balance between the role of an individual’s autonomous agency and the assumed impact of technological constraints. Technologies may disrupt human action by impairing certain choices, but regulatory change is not automatic since it requires human interpretation, deliberation and action (Hood, 1994, p.12). Indeed, the technological modalities of regulation are constructed mutually by human reasoning and technological architecture, given that the objective or subjective condition of the modalities of code is dependent on

how humans perceive it (Scharf, 2012). To be more specific, “a constraint is subjective when a subject, whether or not consciously, recognizes it as a constraint. It is objective when, whether or not subjectively recognized, it actually functions as a constraint” (Lessig, 1998). In this way, technological architectures can be imposed upon people as an objective or subjective condition on which, to an extent, their own autonomous and deliberate action is predicated.

One’s perception of *code* is shaped through an *internalization* process, a long-term, incremental and cumulative experience of being exposed to and making use of technologies in social reality, a process by which we internalize our life patterns, including the way we scope our data rights, as embedded in technologies. I use the term embeddedness over immersion because, as noted by Haraway (1985), it implies a continuous process of engraving individuals within the architecture of technological systems rather than experiencing technology in a relatively ephemeral and instantaneous way. The root of embeddedness also follows the conception of the sociologist Mark Granovetter (1985)’s use of the term to explicate “the extent to which economic action is embedded in structures of social relations” (p.481). The notion of technological embeddedness can be *the extent to which data practices are embedded in structures of technologies*. Embeddedness, the broad context that this study explores, is how the architecture of technologies affects (or regulates) human perception within one’s own realm of digital information.

Competence and Dependence as Embedded in Technology

With the ubiquity of technology (Quan-Haase & Wellman, 2005), being digitally capable becomes an essential means of amplifying human capacity. An optimistic view of the electronic era has been proposed by Bandura (2002, 2004) in that it fosters the

expansion of self- and collective efficacy “to organize and execute the courses of action required producing given levels of attainment” (Bandura, 1998, p.3).

With an inevitable surge of digital technology uses, many researchers have recognized the role of digital capabilities as lifelong competence in the 21st century. Correspondingly, a growing body of digital divide studies have been conducted that link access gaps to Internet and devices as a first-level divide, Internet skills and use gaps as the second-level divide, and gaps in achieving beneficial outcomes of Internet use as the third-level divide (Scheerder, van Deursen, & van Dijk, 2017). This perspective contributes to a comprehensive understanding of the ways in which the compound construct of skills, uses, and outcomes explains how digital skill building leads to certain usage patterns with varying kinds of outcomes, reducing or reinforcing existing social inequalities. Here, various disciplines tend to juxtapose the notion of ICT-related “skills” with competencies, capabilities and literacy, guiding this study’s conceptualization of technological competence as a set of digital and Internet skills. Although further investigations are required to adequately address the relational complexity of multifaceted tech-competence constructs, there are several empirical studies that attempted to delineate the possible relationship between different components of technological competence (van Deursen, Helsper, Eynon, & van Dijk, 2017; van Deursen, Helsper, & Eynon, 2016; Choi, Straubhaar, Skouras, Park, Santillana, & Strover, 2020). For example, van Deursen et al. (2017) conceptualized an Internet skills framework as comprising four skill dimensions, namely 1) operational skills for basic Internet use, 2) information-navigation skills for online information literacy, 3) social skills facilitating online communications and interactions, and 4) creative skills for online content creation from basic uploading of materials to generating textual, video, photo, and other multimedia and remixed contents.

In light of a sequential view of digital skill building for performing within a digital environment, this framework postulated that operational and information-navigation skills are fundamental bases preceding social and creative skills. Choi et al. (2020) suggested three dimensions of capabilities to undertake (1) basic media and information tasks, (2) intermediate work/information related capability that are used at workplaces or school, and (3) advanced privacy/security and creative capability such as content production and computer coding skills. They demonstrated that the relationship between these sets of competencies can be consecutive, where lower level capabilities significantly influence one's acquisition of advanced capabilities. Acquiring additional skills can be understood as a cumulative advantage that helps achieve results.

However, the potential outcome of the reciprocal interaction between technological architecture and humans may not always be in the direction of expanding human freedom and capacity. Being technologically embedded and internalized in networked digital environments may also effectively establish parameters for human capacity as they determine what is possible or not. This in turn may force us to renegotiate our paths and goals. For example, one's level of digital capability may suggest both the ability to maintain a desirable degree of control over digital information or a sense of helplessness or cynicism as one fully realizes that there will be no easy way out of surrounding architecture of technology (Turow, Hennessy, & Draper, 2015; Zuboff, 2015, 2019).

This problem suggests the interesting possibility that the level of competence can be limited by a barely distinguishable notion of *dependence* – as people overly rely on and value a certain role of technologies in their daily lives, it might negatively influence competence or a critical autonomy and awareness (Park, Straubhaar, & Strover, 2019); the latter can be defined as interpreting media texts, systems, and constructed social reality as a

means of “empowerment as a liberating idea, a form of self-determinism, and an extension of agency” (UNESCO, 2013; Livingstone, 2004, 2008). The perception of dependence may be useful in the beginning to motivate people to learn and use technologies, but once a certain point is reached, it can flatten the curve of learning for new things, and one may cling to patterned use and behaviors. If this is the case, user skill may also imply the obedient position of individuals who are blindly acquainted with the instrumental efficiency and convenience of using the conventional technologies which are less likely to allow a critical awareness towards technology or a generative and productive use of technology. This interplay of competence and dependence factors is questioned as a primary concern of the study: how do technological competence and dependence affect value-making performance? I investigate the technological embeddedness of self, what it means, and how the interrelatedness of different aspects plays out in regard to one’s productive and generative practices and outcomes in an online environment.

Awareness of Privacy and Data Rights

Competence skills and dependence included in this study are related to the perception of privacy and data rights to address productive platform use and outcomes. In comparison with research on privacy skills and their relationship to digital capabilities in digital literacy study from the early 2010s, the quest to establish a theoretical conception around data rights has only recently begun (Pangrazio & Sefton-Green, 2020).

Research has found a link between the sense of privacy, privacy skills and digital skills. Byrne et al. (2016) recognized the fact that “access and skills are linked to opportunities and risks” in digital use. Unlike children who lack knowledge of digital safety technologies and managing privacy setting, regardless of their technology skill levels

(Byrne et al., 2016), young adults with high internet skills are shown to be aware of privacy risks and engage in privacy-related behaviors in using social media (boyd & Hargittai, 2010; Hargittai& Marwick, 2016). In this respect, it can be reasonable to assume that technological competence is positively associated with privacy perception and skills.

However, privacy contains many facets and is expanding to respond to multiple dimensions of data rights in today's social media environments. In this regard, the framework of concern for information privacy (CFIP) (Bellman, Johnson, Kobrin, & Lohse, 2004; Stewart & Segars, 2002; Smith, Milberg, & Burke, 1996) consisting of concern for information collection, improper access, potential errors in collected information, and unauthorized secondary use resonates with important aspects of data rights in legislations — using the example of CPRA, right to opt-out, right to know, right to delete/correct, and right to limit use and disclosure and the like. Taking the above rationales into account, a positive relationship can still be considered between competence and each dimension of information privacy concerns. In terms of the sense of dependence, however, it can have no or negative impact on these concerns; for them, their dependence on social media may lead them to underestimate or even disregard the perceived risks. Bellman et al. (2004), in their study researching international Internet users from 38 countries found that users with more Internet experience are likely to exhibit lower levels of privacy concern about their personal information. Considering these concerns for privacy as potential risks perceived by people, this will mediate the association between competence and actual use or directly affect the use; the greater the perceived risk becomes, it will negatively influence actual use of social media.

SUMMARY

This chapter examined how the existing modalities of formal regulation have only vaguely inscribed the rights to actually empower individuals to exercise control over their data. Not only are data policies weak, but also the exploitative nature of the data market highlights the absence of safeguarding personal data rights within the dominant “human right-based approach to data” or “data protection” discourses. Recently proposed solutions have been facing challenges and criticisms. Note that, however, this does not intend to completely deny the essential imperatives rooted in personal data as basic human rights and public goods. The argument here is to expand the field of fundamental rights by envisioning an individual who is entitled to control and own her personal data so that she can decide on what, how and with whom she discloses information about herself. In other words, an intact form of data ownership in its entirety, one that encompasses the state of possession and control in addition to being responsible, is an alternative. Full data ownership includes not only the ability to access, create, modify, package, and remove; it also should contain the right to derive benefit from, sell, or assign such privileges to others. This approach attempts not only to redistribute the wealth of data economy, but also to locate consumers, natural individuals, to a position with *special, direct* interests.

The current position of individuals in the value chain of personal data has been neglected or degraded even though the intrinsic value of an individual person is more and more essential in the process of contextualization that transforms meaningless data into real information for actionable insights and intelligent decision-making. How one manages, maintains and renegotiates the self is a laborious investment, which involves identity, intellectuality, and creativity, in order to interpret and utilize data in the right context. The rationale of privacy tradeoffs research, which assumes the benefits from one’s

exchange of personal information, leaves room to be further queried with respect to the questions of data valuation within an inherently asymmetric data market conditions and lack of a natural setting in designing research.

In addition to the current limits of policy and law and the existing data market, in order to fully exercise and realize one's ownership rights to data and its value, some acknowledgment of socio-technical conditions is essential. Empowering ourselves through rights to data is intertwined with the continuous structuring of our lives by technology. Invoked by the notion of technological embeddedness with a sense of competence and dependence in technology use, individuals can build a system of self-governance to perceive, manage and value their own personal information.

Within the aforementioned frameworks and concerns, the current study proposes and elaborates one alternative direction. Drawing on the case of Steemit, an incentivized blockchain-powered social media platform, as a site for research, the present study investigates a natural environment in which the actual information behaviors of users involved in data trades can be observed to unveil the complexity in managing and valuing personal information. Examining information transformations within this environment can identify the "comfort zone" of information to be disclosed or shared across the vast range of personal information, as well as how diverse one's strategies are to express digital identities. It can discover what and how some personal aspects are shared more often, and how they are valued.

Chapter 3: Methodology

Using the incentivized blockchain-based social media platform Steemit, I examine empirical data from self-generated posts of Steemit users that reflect on their personal information practices premised on direct and tangible values. This study builds on prior work around conventions of self-disclosure in social media settings, a limited set of studies examining valuation processes with respect to personal data, and matters of technological dependency and competence. The current chapter first introduces the research setting, and then outlines the three studies that comprise empirical assessments into aspects of handling personal data. It includes how the conceptual constructs for each study are defined and operationalized. Each study has its own methodology section.

Despite extensive research efforts, how an individual invests and values private information in reality is still far from clear. As discussed earlier, research on the relationship between one's efforts in social spaces and attaining value from one's contributions is limited; the value of personal data is context-dependent, and most research suffers from the lack of a natural setting that confines their participants as well as the subjects of their investigation to a narrow range of predetermined set of choices (Grossklags & Acquisti, 2007; Lesk, 2012; Spiekermann et al., 2015). Further, the existing formula of data value as derived from classic economics may be restricted by data market conditions (e.g., an inherent limitation of information failure or status quo conditions like terms and conditions contracts).

To sum up, in an attempt to determine personal data values, previous research has failed to address one's deliberate data practices free from platform capitalism and other predetermined constraints. These gaps could not only reduce the explanatory power of

human data behavior in natural circumstances, e.g., free market, but also could deepen the imbalance in market information, making consumers powerless.

The present study sees this gap in scholarly endeavors as an opportunity to reconstruct the breadth of individual practices around everyday data transactions. Using the Steemit platform as a backdrop, this study applies a mixed-methods approach consisting of topic modeling, content analysis, and survey data analysis combined with digital trace data to get a better picture of users' investment of their personal data and realized values. Steemit is one of only a few websites that monetarily rewards users for what they contribute. Three empirical studies as reported in Chapter 4-6 examine users' behaviors on Steemit to respond to our core research questions and hypotheses:

- Topic modeling approach

2. When compensation is assured, how do people invest in or get rewarded by information about themselves through self-created content? (Cont'd in Chapter 5)

2.1 What types of personal information are voluntarily disclosed and shared?

How does each type of the information relate to realized value?

- Content analysis

2. When compensation is assured, how do people invest in or get rewarded by information about themselves through self-created content?

2.2 How does one strategically manage identity through divulging personal information? How do the characteristics of managed identities relate to the realized value?

- Online user survey analysis

3. How do people's technological embeddedness affect their value-making performance on Steemit?

3.1 How do technological competence and dependence affect value-making performance?

3.2 How do technological experiences and capabilities affect value-making performance in relation to awareness of data rights?

From the topic modeling approach in the first study, I investigate how people invest in or get rewarded by volunteering information about themselves. By extracting underlying topic spaces from the corpus of user-generated posts, I examine from a macro view which personal information topics are voluntarily disclosed and shared on the platform and how much value was realized in each topic space. From a micro point of view, the second study uses content analysis methods to undertake an in-depth look at randomly selected user posts to explore varying strategies of enacting identities on the platform not only in terms of the type of information posted but by how it is posted. In a third study, an online user survey augmented with extracted user records about peoples' performance on the platform explores the structure of technological conditions for accomplishing identity – and hence making value – on the platform.

Details regarding the rationales for the selection of Steemit as the main site of research as well as more specific research designs for each research question are expanded below.

RESEARCH SETTING: STEEMIT

Steemit (Steemit.com) is a decentralized social networking blogging platform built upon the Steem blockchain, which was launched in March 2016. Inspired by Reddit's provocative hypothesis in 2014,³ financial analyst Ned Scott and developer Dan Larimer envisioned a social network of mutually supportive communities where people can help each other with their "subjective contributions" (p.3), which resulted in their co-founding of Steemit, Inc. Larimer left Steemit after laying out the technical foundation, and Scott sold the company to Justin Sun, a Chinese billionaire who founded the Tron platform, as of February 2020. This recent takeover by Sun has sparked a "war" between Sun and Steemit community members over the control of decentralized community, which has many interesting aspects that can bring valuable lessons across blockchain governance for the coming decades, but is beyond the scope of the present study.

Setting the above controversies aside, Steemit's social network media platform had embarked upon a revolutionary vision, which has attracted more than 1 million registered users worldwide as of 2019 (Guidi, Michienzi, & Ricci, 2020). Like many other conventional social media platforms such as Reddit, Facebook, Twitter, or Tumblr, it allows users to create and share content, upvote or downvote, and network with other users.⁴ What differentiates Steemit from other social media platforms is that Steemit users

³ In 2014, Reddit considered creating its cryptocurrency arguing that "its platform would be improved if everyone who contributed to reddit.com by posting stories, adding comments or voting were rewarded with a fair share in Reddit, Inc" (Steem Whitepaper, 2017, p.5).

⁴ In terms of Steemit's affordance – the design aspects or properties of artifacts that suggest how the artifact might or should be used – consists of 1) identity affordance, 2) social affordance, 3) cognitive affordance, 4) emotional affordance, and 5) functional affordance (Moreno & D'Angelo, 2019): Although the interface of Steemit has been similar to that of Reddit, it's important to note that the detailed affordance features below are from August 2021 and may be more refined or different than they were in the early days of the community (*Note*. As an active member of Steemit, these affordances were observed and elaborated by the researcher during the period 2019-2021).

1. Identity affordances: Steemit can be classified as a low-identity social media platform where no identity requirement exists; users are typically identified by a self-selected username, focused on

can be rewarded by the community with STEEM cryptocurrency by posting, commenting, sharing, or voting on content, unlike general social media users whose data is extracted in exchange for “free-of-charge” services from the platform. STEEM, like any other digital currencies, can be sold, traded, and exchanged in open cryptocurrency market. The supply of new STEEM coins bases its protocol on Delegated Proof of Stake (DPoS), which uses community-elected “witnesses” instead of “miners” to produce blocks; 21 Witness accounts elected by the community produce a new block every 3 seconds that is mined to be allocated to the “reward pool” for content publishers and curators (75%), Steem Power (SP) Holders (15%), as well as Witnesses (10%) (Steemit Bluepaper, 2017). This study’s selection of the Steemit platform benefits from the “uniqueness” of its in-built incentive mechanism that involves users in the value-making and value realization processes based on their data activities.

Upholding the concept “everyone’s meaningful contribution to the community should be recognized for the value it adds” (Steem Whitepaper, p.6), Steemit users can be

-
- sharing content rather than expressing personal identity. Notably, however, the user account is liked with a STEEM wallet page that displays how much a user earned based on his/her activities.
2. Social affordances: to facilitate group-based user networking and a sense of belonging around a particular interest, experience, or social group, Steemit provides tagging and community board features (in addition to following and mentioning). Note that each community can have leadership that allows the executive committee to set its own rules allowing membership registration and posting. It does not provide direct message function between users.
 3. Cognitive affordances: defined as tools to expand one’s learning, Steemit provides an opportunity for increasing awareness and knowledge of digital currency, digital currency communities and market, and international affairs. It also enhances creativity by allowing customizable content. Additionally, Steemit users can rejoin and reengage web-based discussions with an automated alert feature.
 4. Emotional affordances: with “upvoting” or “downvoting” function, Steemit allows users to express emotion. As it also customizable content, users can generate sympathy by providing photos and personal stories or sharing personal information.
 5. Functional affordances: like many other platforms, Steemit content is replicable i.e., “reesteem-able”; scalable (with the number of votes and the amount paid for a post); searchable (with hashtags); permanent (posts can’t be deleted but edited); no limit on composition time at the maximum size limit of 64 kb; and without Steem Power, posting, transaction, voting activities, etc. may be limited (referred to as “Bandwidth limit”).

paid based on votes from community members on their posting or curating⁵ for other's contributions as well as exercising vesting ownership⁶ as a direct shareholder of the platform. Along with this, Steemit employs a Steem Power (SP) and Reputation mechanism as one way to properly adjust the amount of value content has brought to the platform. Simply put, SP is STEEM that has been committed to a certain vesting period; Reputation increases from upvotes on posts or comments by users with higher reputation than the writer. Although both SP and Reputation scores do not directly affect the reward, it affects the impact of your upvotes (SP) and the post visibility and trustworthiness (Reputation); this mechanism helps reduce abusing activities as well as avoids "get rich quick" attempts. Steemit made clear that while high rewards on a single post is not an impossible occasion, in most cases the rewards are usually dependent on long-term, consistent contributions to the platform that include making connections with others, network building, and developing a reputation for bringing high quality content. Accordingly, Steemit users have diversified the strategy to "win" in the system (i.e., getting more votes without penalties) in both machine- or human-driven ways using voting bots or building user communities based on shared interest, culture, or location to support each other.

Such features of Steemit make it unique, providing the closest proxy for a free-market system in which content (data) consumers directly "pay" the producers with votes. With its design that enables effective micropayments for each person's contribution, Steemit can arguably be the most suitable place to observe one's deliberate, motivated, and

⁵ If a user discovers a post and upvotes it before it becomes popular, s/he can earn a curation reward.

⁶ Vesting ownership makes a long-term commitment and cannot be sold for a minimum period of time. It not only allows users to be paid interest on the balance remaining vested, but also to have more influence on the distribution of rewards (Steem Whitepaper, pp.7-8).

active data behaviors consisting of self-disclosure, self-expressions as well as social interactions.

The Corpus

Steemit organizes every user-generated content by two primary features, namely, “tags” (sorting by subject) and four categories of “trending”, “hot”, “new”, and “promoted” (sorting by popularity, by the time generated, and by the author’s payment for the promotion).⁷ Steemit feeds are much like the typical social media feeds, where the preview of users’ self-generated posts appears on your feed from people you follow as well as from the two features above. Information associated with each post includes a title, author ID, author’s reputation score, attached tags, time posted, payout value earned, the number of votes obtained, response comments in addition to the main body of the post consisting of text, image, and video and the like. Figure 1 below provides a snapshot that helps readers navigate the Steemit user interface.

“Introduceyourself” and “blog” are tags or categories that organize posts by specific interests or purposes. As the name indicates, “introduceyourself” is where people introduce themselves mostly when they first join the platform. Topics of posts usually contain information such as name, age, occupation, hobbies, family, nationality, selfies, and their other SNS accounts or blogs and so on, and how, why, or by whom they joined, soliciting help or advice to do well on Steemit, and how they would contribute to the platform. With the “blog” tag attached, people post anything they want to write or share. This is the most general category where people contribute a vast range of topics from food recipes,

⁷ This interface was effective until the time of data collection around the end of November, 2019. After that, while the basic operation is similar, the organization of the website has gone through some changes.

gardening, and traveling to the future of blockchain and cryptocurrency and to political issues like the presidency of Donald Trump or international conflicts. There is basically no specific character or content limit in posting. There are a variety of tags beyond "introduceyourself" and "blog," but this study chose these two categories for analysis: the former reflects the scope of personal information shared at an entry-level of community participation; the latter will illustrate the subtle context in which people use their personal information, without being too much slanted to the uniqueness of a topic, to supplement their posts or themselves.

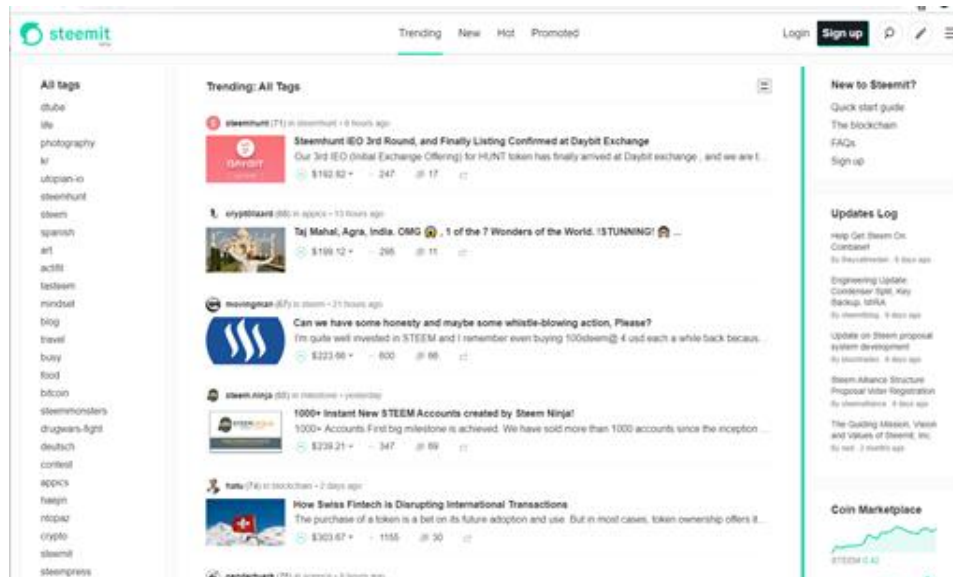


Figure 1: Steemit home page that displays the categorized posts.

Four Phases of STEEM

The fundamentals of the Steemit ecosystem are rooted in STEEM, a unit of cryptocurrency based on the Steem blockchain. Every user activity on Steemit is closely related to tangible STEEM rewards. This means that for most users, if not all, the STEEM

value—the price of one STEEM unit—has been a crucial factor shaping their Steemit activity: the higher the STEEM price, the more motivated and active the user can be. This rationale makes it reasonable to assume the association between the focus of this study – voluntary and motivated information sharing behavior of people – and the price change of STEEM over time. This means that our examination of the value of research data must be cognizant of the market conditions and fluctuations.

Indeed, the value of cryptocurrency has been fluctuating in recent years. In particular, several studies have reported the hype around cryptocurrency represented as a Cryptocurrency Bubble, evident in the sharp rise of value and sudden collapse of Bitcoin during the period from the end of 2017 to the beginning of 2018 where it lost more than half of its value (Kyriazis, Papadamou, & Corbet, 2020; Fruehwirt, Hochfilzer, Weydemann, & Roberts, 2020). STEEM and many other cryptocurrencies went through similar bubble phases around the same time period. This contextual feature may influence the value of STEEMIT and consequently the types of user engagement and privacy behaviors.

To choose the best time frame for research between these phases, the study draws on Canadian economist Jean-Paul Rodrigue’s model of an economic bubble in Figure 9 (2008). Rodrigue explains bubbles unfold in four stages, from the stealth phase where “smart money” is invested taking a risk, to the awareness phase where fast-followers begin to invest as they notice the momentum. These phases are later more sophisticated early-stage investors with better information and a better understanding of economic contexts. The last phase is called the “Blow-off” phase, signifying some regret and denial.

The present study targets the next, “Mania” phase in which a steep rise of value occurs through investments from the general public. There are two reasons why the mania

phase satisfies the purpose of this study. First, this phase attracts the most ordinary people, and the study seeks to investigate data practices of average users who are not classified as celebrities or innovators (who may bias data). Second, the mania phase, even slightly over to the blow-off phase, displays peak values. This high value, combined with spreading positive media coverage, provides conditions that motivate people to act the most. As well, the change in value from the “mania” phase to the “blow-off” phase can provide an indication of the sensitivity of users to the value of their personal data, a temporal factor that contributes to the topic modeling analysis as well, as described in later pages.

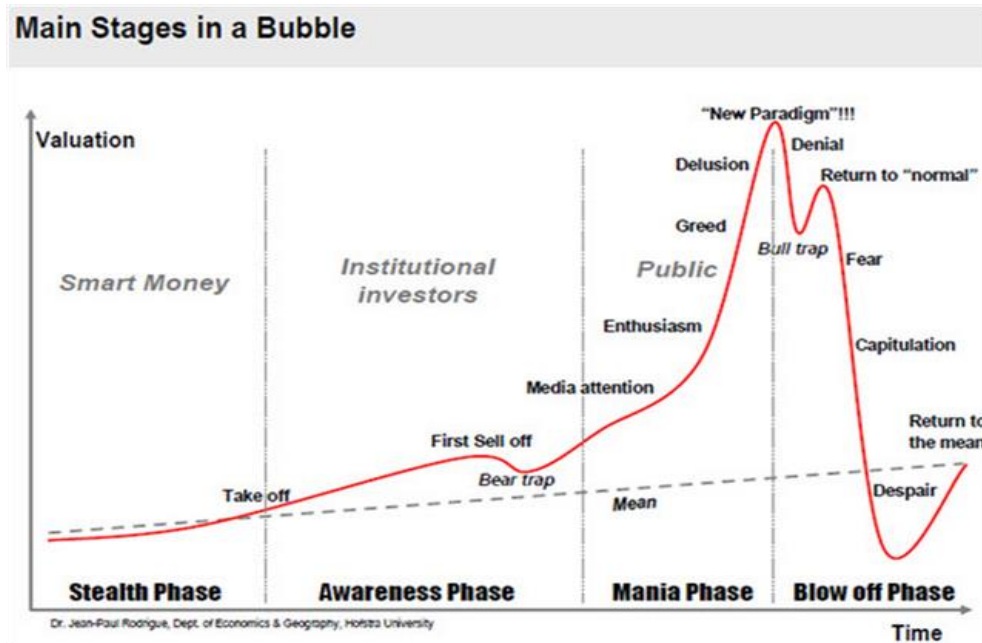


Figure 2: Phases of a bubble. Jean-Paul Rodrigue (2008).

Comparing the phases of bubble with the price change of STEEM over time in Figure 3, the full two-month period from December 1, 2017 to January 31, 2018 (mania phase) constitutes the appropriate phase for data gathering in this study. In this period the STEEM price showed the highest record, and thus one can assume the general public is

engaged the most (around the price-tagged period in Figure 3). The topic modeling study also uses another two-month period from September 2019 to November 2019, in which the price of STEEM drastically declined (blow-off phase), in order to compare the possible impacts of each market condition on the public engagement on the platform.



Figure 3: Steemit price chart.⁸

⁸ Accessed June 30, 2021 at: https://coinranking.com/coin/fXGcu_EzDgP25+steem-steem

METHODS

Analysis Plan

Steemit transactions describe a value space of personal data that represents the values and concerns in our everyday digital surroundings. To investigate how people negotiate their privacy and self-disclosure practices in this unique social medium, three separate studies are conducted using data obtained from Steemit. The first study employs natural language processing (NLP) and topic modeling approaches to analyze components of personal information shared across the large-scale collection of users' self-created posts. The topics shared on Steemit can illustrate the collective user consensus on a "comfort zone" in terms of privacy, and can also illustrate how people might respond textually to valuing their data. The advantage of this NLP computational approach is that it is possible to sample all user posts generated during the study period and is mostly unsupervised, so it is not only highly representative but also less likely to be biased by the researcher's preconceived notions. However, it inevitably lacks depth and may overlook the contextual meaning that each post can express, either explicitly or implicitly.

The second study, using content analysis, aims to provide a complementary picture to further explore the findings of the earlier research design by taking a closer look at people's utilization and expression of personal information in content creation. Based on criteria established through previous studies, it analyzes randomly selected user posts in order to explore the conditions and context of full voluntary disclosure of personal information: under what conditions and when are people actively willing to disclose and share information about themselves by creating their own content?

Finally, I chose a survey of Steemit participants to examine some other possible endogenous or exogenous conditions affecting users' value-making activities. Specifically,

the survey investigates how technological competence and dependence, existing perceptions of privacy and data rights, and socioeconomic characteristics relate to activities on Steemit. An online survey targeting Steemit users is augmented with the respondents' Steemit activity record to unobtrusively observe the relation between user's reported technological modalities in everyday life and value-making performance on the platform.

Definitions and Operationalization

Each chapter tackles different aspects regarding how people generate value and engage the social sphere of Steemit.

When thinking about personal data space (Chapter 4), this study refers to a fixed set of topics that best describes a set of self-introduction content posted during the peak and decline periods of the STEEM cryptocurrency market. Considering the exploratory nature of the topic modeling methodology, this analysis is data-driven without pre-defining topics.

A content analysis of actual posts (Chapter 5) examines users' identity management practices; identity management refers to the strategies for representing the self by deciding what and how people reveal information about themselves (Livingstone, 2008). Assuming opportunities (i.e., value-creation) and risks (i.e., privacy) in the identity-expression process, the study applies five different constructs to analyze the posts: 1) elements of self-display (mode of communication) (Panahi, Watson, & Partridge, 2012), 2) authorship (Jin, 2013; Flath, Friesike, Wirth, & Thiesse, 2017), 3) online boundary management behaviors (Batenburg, & Bartels, 2017; Livingstone, 2008), 4) implicit self-presentational information (Jeong, & Coyle, 2014; Jeong, & Kim, 2017), and 5) explicit self-presentational information (Jin, 2013).

The elements of self-display assess how diverse types of multimedia elements were used in a post (e.g., photos, texts, links, etc.). The authorship is coded in terms of the main source of the post – whether it is original work, remixed or copied. Online boundaries consider four aspects: whether the author’s personal and professional identities appeared together or not, and whether the self-disclosure is fully positive or mixed positive-negative. For implicit/ explicit personal information, five dimensions of information are assessed. In each construct, items do not always uniformly function to increase or diminish the opportunities of values or privacy risks. Further explanations and details shall be elaborated in Chapter 5.

A user survey (Chapter 6), probes the roles of technological competence and dependence in users' online behavior. These two qualities are the main pillars of technological embeddedness. Technological competence, defined as capabilities and know-how in utilizing technologies to live, learn, work, create and interact in a digital society, consists of four types of skill domains: operational, information-navigation, social, and creative (Van Deursen et al., 2017). Each domain was measured by four to seven-item instruments based on 5-point Likert scale. Technological dependence is assessed by the extent to which people rely on a certain technology, comprising two different dimensions of task-based and social-centered dependence; it uses four-item instruments for each dimension (Park et al., 2019). Awareness of privacy and data rights in a social media setting is measured by modifying the concern for information privacy scale that is used to measure consumer privacy concerns in the context of e-commerce (Bellman, et al., 2004; Stewart & Segars, 2002; Smith et al., 1996). Comprised of four different dimensions – information collection, inappropriate access, information errors, and unauthorized secondary use, each information privacy dimension has three to four items in a 5-point Likert scale instrument.

As a mediating variable, the study uses the actual use of Internet and the Steemit platform. This is measured by how often you use the Internet a day, how long you use it, and how many social media platforms you use when you use the Internet. The same measurements apply to the use of the Steemit platform, but in terms of other social media usage, we measure the number of other social media accounts that are disclosed on Steemit. For a dependent variable on user performance, the study looks at how “successfully” the user has performed on the platform in terms of the number of followers, number of posts posted per month, reputation score, and the estimated account value in USD.

Chapter 4: Data as Investment: Macro View

STUDY 1 METHOD: TOPIC MODELING

Computational topic modeling approaches using natural language processing (NLP) techniques were used to extract the primary dimensions of users' information investment in general, and features of personal information reflected in the users' semantic data contributions.

NLP, in a broad sense, refers to any kind of computational understanding and manipulation of unstructured textual data in order to detect patterns and discover the latent meanings of it (Bird, Klein, & Loper, 2009). In response to the availability of a huge volume of data, NLP applies highly scalable statistics-based techniques using machine learning models to analyze a large corpus of texts. NLP is versatile. Its major techniques include text summarization, information extraction, and information retrieval that can be used for many different domain-specific applications (Chowdhury, 2003). It is one promising way to leverage traditional communication research by allowing the analysis of actual "conversations" or transactions between people in a digital setting (Choi, 2018). Although research must be cautious in claiming construct validity and making inferences based on collected data and the theoretical framework (the collected data may not necessarily represent complete populations nor true human behavior, according to Howison et al., 2011 and Jungherr et al., 2017), it is useful in situating knowledge in a certain context by exploring the sheer richness of data (Malik, 2018). From this perspective, this analysis examines the categories of personal information people were willing to invest, i.e., disclose or share, through their self-generated contents and how each topic cluster of personal information was valued. By looking at topics in the two time phases, the study considers a temporal dimension related to the cryptocurrency market

condition to see if it affects the type of information invested and values realized from it. This speaks directly to how value can be expressed and used by individuals.

This study applies Latent Dirichlet Allocation (LDA) topic modeling using Gensim (<https://radimrehurek.com/gensim/index.html>), an NLP-related Python package for building topic models. LDA is a statistical model for finding the latent structure of a vast collection of documents (Blei, Ng, & Jordan, 2003; Hoffman, Blei, & Bach, 2010). LDA is a generative probability model that involves three-level hierarchical Bayesian inference (document-topic-word), where the model assumes that each document, as random mixtures, can be represented as a finite set of latent topics and that each of these topics is attributed to a multinomial distribution of words (Blei, Ng, & Jordan, 2003; Wang, 2018). It is an unsupervised learning algorithm that automatically and efficiently derives hidden thematic structure of large corpora based on patterns of word co-occurrence (Jacobi, Atteveldt, & Welbers, 2016; Porter, 2018). As it readily overcomes the limits of hand annotation, there has been wide application of LDA topic modeling in a variety of sciences as well as social science research, including the field of communication studies (Maier, Waldherr, Miltner, Wiedermann, Niekler, Keinert, Pfetsch, Heyer, Reber, Häussler, Schmid-Petri, & Adam, 2018).

Users' information investment, whether from self-disclosure or self-expression, in Steemit is the object of the current analysis. We develop topic models with posts attached to the Introduceyourself tag as input. In doing so, we compare the input that occurred in the mania phase and late blow-off phase respectively in order to see how the prospect of cryptocurrency has influenced user's data investment. The analysis process is summarized in Figure 4.

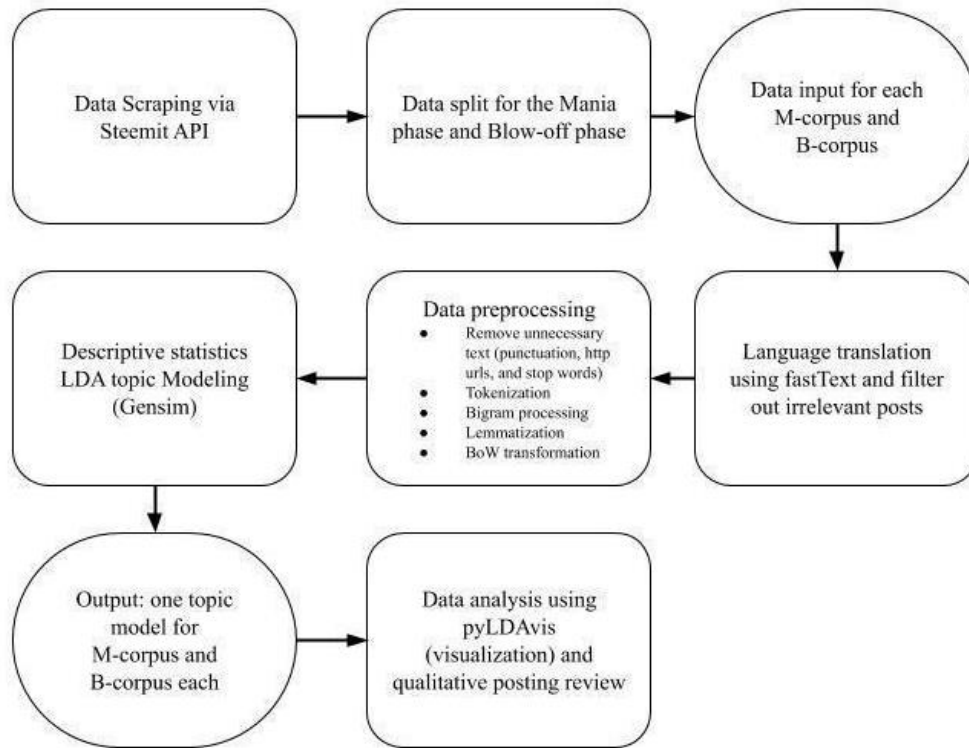


Figure 4: Basic workflow of LDA topic modeling.

Data Collection

The study collected the posts tagging the “*Introduceyourself*” category from 00:00 December 1st, 2017 to 00:00 January 31st, 2018 (mania phase: M-corpus hereafter) and from 00:00 September 11th, 2019 to 00:00 November 11th, 2019 (late blow-off phase: B-corpus hereafter) for the two-month period each through “*steem-python*,” Steemit API-based python library. Each corpus contains English language posts only, but some non-English messages are also included as some users have posted in multiple languages. To filter messages that are not exactly intended for introductions, the study only included posts that have sub-categories of “introduce”, “introduction”, “introducemyself”, and

“introduceyourself” from the metadata. Out of 54,253 and 16,710 posts each scraped for the study, a total of 26,413 and 7,039 posts remained after the filtering for the corpus of the mania phase (M-corpus) and the late blow-off phase (B-corpus), respectively. The shortest post per corpus had 1 and 2 characters, and the lengthiest post per corpus had 35,208 and 57,286 letters for the M-corpus and B-corpus each. On average, the mean length of a post was 1,766 and 1,833 letters and the standard deviation of the length per post per each M- and B-corpus was 1,866 and 2,380.

Data Analysis Tools

The Table 1 below details the Python-based tools that are used for the study.

Data processing flow	Tools
Data scraping	steem, json, pick, pprint
Data extraction	os, ast, csv, pandas, re, fastText
Text preprocessing	nlTK, re, string, spacy, pandas
Topic modeling	gensim, pandas, pyLDAvis
Visualization and NLP-based analysis	pyLDAvis, numpy, pandas, seaborn, matplotlib, WordCloud, Counter

Table 1: List of Python-based tools used

The main analysis tools include Gensim for computational thematic analysis of a large collection of texts and NLTK (<https://www.nltk.org/>) for natural language processing that provides a suite of text processing libraries for text classification, tokenization, stemming, tagging, parsing, and semantic reasoning and so forth. As Steemit is a platform composed of global users who use various languages from different parts of the world, their posts also contain a wide variety of languages, including Spanish, Chinese, Korean, German, French, and so on in addition to English. To detect English language in the post,

the study used ‘fastText (<https://pypi.org/project/fasttext/>),’ a text classification library that recognizes more than 170 languages. For the visualization purpose, libraries like pyLDAvis, matplotlib, seaborn, and WordCloud were used.

Data Preprocessing

M- and B-corpus were preprocessed separately, as each corpus constitutes a different topic model. The specific procedure, however, was almost the same except the application of several stop words. First, as aforementioned, the scraped data retained only the posts of which the “category” of the metadata corresponds to “introduce,” “introduction,” “introducemyself,” and “introduceyourself,” and filtered by the output of fastText detection of English language. After then, since the post is unstructured text, the standard method of preprocessing was performed: lowercase text; replace line breaks, trailing whitespaces, or special characters (e.g., slash, hash, parentheses, square or angle brackets and so on) in the post with a single space; removal of http URL links; removal of punctuation; tokenization; filtering stop words; text lemmatization; tagging each token and keeping only tokens in the form of noun, adjective, verb, adverb in the corpus.

The list of stop words refers to a set of commonly used words such as “a,” “the,” and “and” that is basically filtered out in processing natural language data because it adds little meaning for analysis. To avoid having the generated topics revolve only around these common terms, this study’s initial analysis filtered a NLTK-provided set of 179 stop words. Then the basic stop words list was needed to be extended when the extracted topics contain words that have little semantic value or make the topic uninterpretable. The process of detecting such useless terms entails repetitive testing of several candidate models and interpreting the ten most relevant terms for each topic. When the model achieved sufficient

coherence without holding significantly irrelevant terms, the number of stop words for M-corpus was 194 words and B-corpus applied a total 217 words.

Implementing LDA Model

The preprocessed text turns into a Gensim dictionary that encapsulates the mapping between terms and their integer IDs, which classifies word groups so that researchers can find abstract topics that best characterize a collection of documents.⁹ According to Maier et al. (2018), the number of topics is determined by the researcher by adopting a two-step approach: first, researchers can use the mean intrinsic coherence of a specified LDA model. Second, thereafter, they qualitatively examine candidate models that best fits the theoretical framework of the study.

Topic coherence provides a convenient measure to evaluate the consistency of a single topic that is assessed by the semantic similarity between highly relevant words in a topic (Porter, 2018; Prabhakaran, 2020). While there is no strict range for an acceptable coherence score (1 at its maximum), the model with the highest value for a given data usually offers meaningful and interpretable topics (Maier et al., 2018; Prabhakaran, 2020). The study calculated the coherence score by using the Coherence Model from Gensim. Multiple testing results of candidate LDA models with different values of numbers of topics (K) from 1 to 20 indicate that the optimal number of topics for each of M-corpus and B-corpus is 10 ($c = .440$) and 8 ($c = .513$), respectively, where the coherence value is

⁹ The processed dictionary is converted into the bag-of-words (BOW) format and the corpus of all unique words is vectorized based on the term-document count matrix for all integer word IDs. In addition to the specified corpus and the dictionary, there are three model parameters called α , β (referred as “eta” in Gensim), and K (the number of topics to be generated) that must be determined by the researcher. To appropriately select two prior parameters (i.e., α , β), this study set an “auto” option for the values of both parameters in an expectation of optimization since it automatically learns the asymmetric prior value from the corpus being analyzed.

highest. The qualitative interpretation of the dominant keywords and the most representative documents for each topic also support that the selected models can adequately reflect the data and be meaningfully interpreted.

To further analyze the topics retrieved from each M- and B-corpus, the Python extension (pyLDAvis) for interactive topic model visualization tool was used. The study set the value of lambda (λ), a weight parameter that shows the relevance of a term to a topic (Sievert & Shirley, 2014), to 1, with 10 and 8 optimal topics for M- and B-corpus each. pyLDAvis displays an intertopic distance map via multidimensional scaling based on the inferred LDA model results. The layout of pyLDAvis consists of two charts shown side-by-side: the left panel shows visualized topic bubbles scattered or clustered throughout the chart representing the topic prevalence while the right-hand side panel contains the salient keywords with estimated term frequency within the selective topic (Katre, 2019).

Through its interactive visualization feature, pyLDAvis allows one to adjust the term ranking by changing the value of λ from 0 to 1, where a small value of λ highlights more unique and exclusive terms for the selected topic and large λ value emphasizes frequently appearing terms for the chosen topic (Sievert & Shirley, 2014).¹⁰

STUDY 1 RESULT: SPACE OF PERSONAL INFORMATION INVESTMENT

This section provides the full results of two specified topic models showing the topics extracted respectively from 1) the Mania period from December 1, 2017 to January 31, 2018 and 2) the Blow-off period from September 11, 2019 to November 11, 2019 out

¹⁰ Note these terms selected for each topic may not be completely exclusive; a term chosen for a topic may also appear in another topic.

of the Steemit users' self-introduction posts. Again, the two topic spaces will hereafter be referred to as M-corpus and B-corpus for the sake of brevity.

Our analysis reveals that the optimal number of topics for M-corpus is ten; for B-corpus, it was eight. In Figure 5 below, each topic is plotted as a circle and the size of circles denotes the prevalence of given topics, meaning that the percentage of a document in the corpus that can be explained by a topic. The topic numbers are listed in ascending order from the most prevalent to the least prevalent. The position of circles on the two-dimensional chart is drawn from the computed distance between topics; this shows how topics relate to each other, meaning some subject overlap may occur between some topic areas. Most topic circles are fairly dispersed throughout the chart, providing a justification for this model choice (Sievert & Shirley, 2014). In association with the Steemit user's invested/shared personal information, the study labeled topics generated from each M- and B-corpus based on the top 30 most relevant terms as well as the most representative posts, five to fifteen, for each topic. These 30 terms are shown in Figure 6, where each bar displays overall corpus-wide term frequency in light blue and the estimated topic-specific term frequency in red.

We first delve into the topic space during the Mania phase (M-corpus) at coin bubble peak, and then move into the late Blow-Off phase (B-corpus) after the bubble burst. Finally, the study examines whether, and how, the characteristics of topic areas, such as the types of information, topic prevalence, post volume, and the value realized in each topic space as a whole and in each post contribution on average, have changed between both periods. Topics that account for less than 1% of the corpus are excluded from further analyses.

Comparison of Topics in User Information Investment Between Mania- and Blow-Off Phases

Figure 5 shows the specified topic model from the LDA algorithm during the Mania phase (left) and Blow-off phase (right). While the majority of the general public becomes enthusiastic during the Mania phase as they observe the soaring value of cryptocurrency, including STEEM, this sentiment was dampened during the Blow-off phase when a sharp downward signal was detected after the burst of hype and bubble around cryptocurrency. Accordingly, the volume of user-generated posts decreased in the Blow-off phase compared to the same 2-month period in the earlier market peak phase.

Overall, labels for 10 topics of M-corpus and eight topics of B-corpus are determined according to the most relevant keywords as well as the most representative posts in Table 2 (see Appendix 1 and 2 for examples of representative posts). For the Mania phase, the most dominant of all topics was “self-expressions”, accounting for the largest circle in Figure 5; this topic is followed by “self-disclosure.” In comparison, in the case of the Blow-off-corpus, “self-disclosure” expressed the largest proportion of all topics, followed by the topic on using the Steemit platform.

Clear topic shifts are evident between the Mania and the Blow-off phases. The M-corpus focuses on self-expression, self-disclosure, personal narratives, personal relationships, the use of Steemit platform, work/education history, creative work, cryptocurrency and coin market, and lifestyle/leisure topics. Most topics appear to remain the same for both the M- and B-corpus, with similar keywords for each topic overlapping. However, the topics “self-expression” and “work/education history” observed in the M-corpus disappeared from the B-corpus. The topic proportions also differ in the phases.

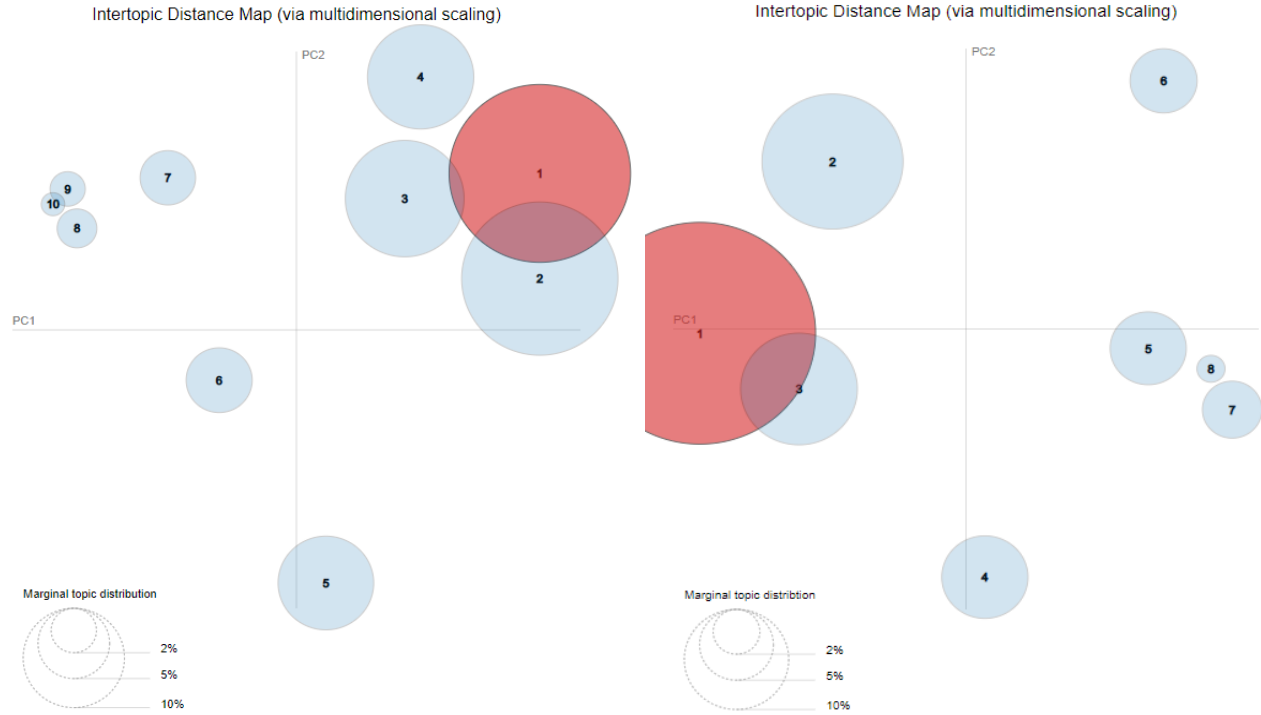


Figure 5: Intertopic distance map for M-corpus (left) and B-corpus (right). PC: principal component.

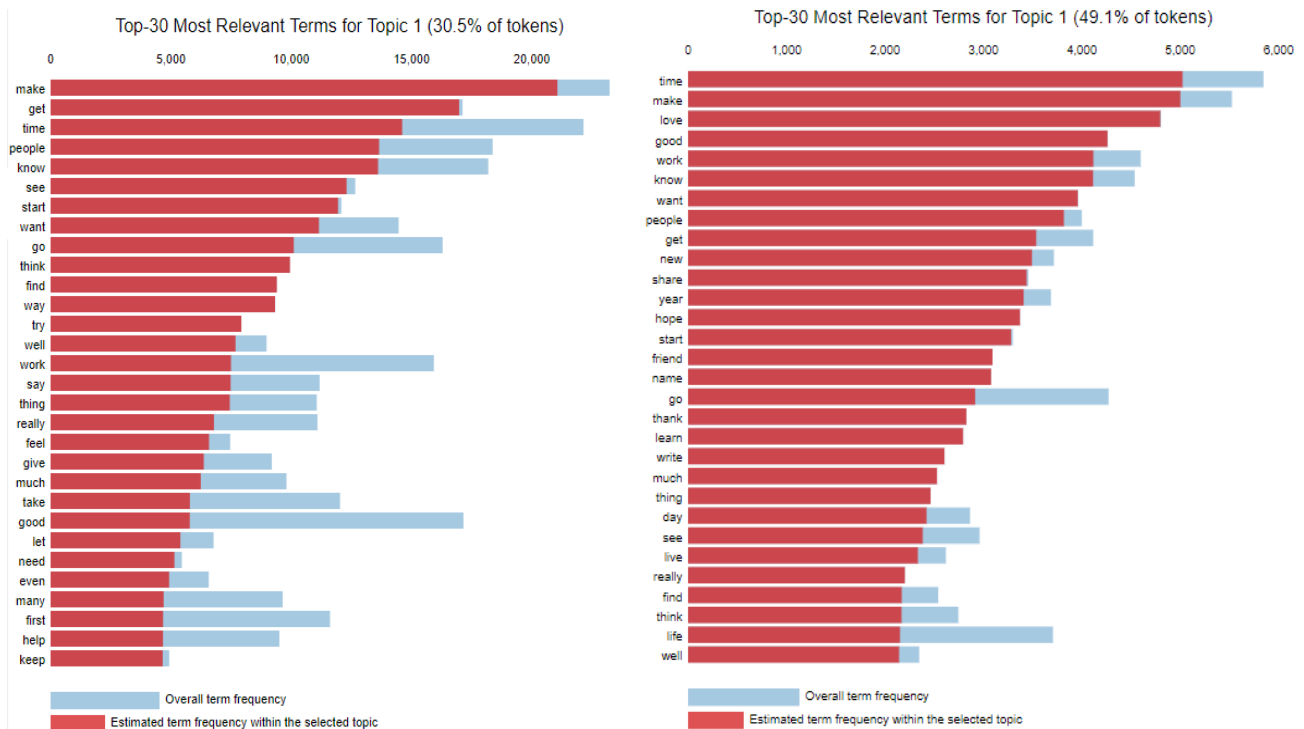


Figure 6: The composition of the most relevant terms for Topic 1 of M-corpus (left) and B-corpus (right)

Table 2 details the M- and B-corpus topics associated with how individual users contributed information about themselves on Steemit. In terms of the M-corpus Topic 1 (self-expression) and Topic 2 (self-disclosure), people mainly focused on introducing themselves effectively, while adopting a slightly different strategy of self-display: the posts in Topic 1 give lengthy explanations of people – what they believe in life, their life vision or goals, rather than (or in addition to) basic facts, for example: “.....*I am quick to say how I feel and people find it offending well.....I never worry about things especially things I can’t change.*” or “... *I am a regular guy, I do not use fancy words or try to tell people how to keep up with the Joneses....*” In Topic 2 self-disclosure, many posts provide short and descriptive basic facts such as: “...*So my name is XX, I was born in Peureulak on 11 March 1999 and I live in Aceh Indonesia in YY district of Peureulak City. And my profession is still in learning or learning about other sciences both religious knowledge and other general knowledge. My hobby is playing football and badminton. I have one brother and two sisters and I love them so much....*” People frequently mentioned, in entirety or in part, information such as their name, gender, where they were born and live, birth date, job and hobbies, and their family. The expressive work of individual users went beyond mere self-profiling.

Further, a considerable proportion of the M-corpus posts was intertwined with narratives on diverse subjects in which people share anecdotes based on their own direct or indirect experiences, or others’ experiences (Topic 3, personal narratives: 13.7%), faith and love for family, neighbors, and God (Topic 4, personal relationships: 11/0%), detailed history of their education and work (Topic 6: 4.2%), creative original work or appreciating of art (Topic 7: 3.0%) or leisure or living tips such as cooking and gardening (Topic 9: 1.2%). Along with this, more instrumental purposes (beyond personal information) were

observed in Topic 5 and Topic 8, in which people share experiences or solicit help or advice to thrive on Steemit, or to better understand what cryptocurrency is, how to earn it, and how the cryptocurrency market works.

M-Phase Topics (% of the corpus)	B-Phase Topics (% of the corpus)	Topic Description
Theme 1 (32.1) Self-expression make, get, time, people, know, see, start, want, go, think		Self-introduction focusing on: - Expressions of who they are (e.g., personalities, life philosophies, life goal or visions, etc.) Motivation to join
Theme 2 (23.8) Self-disclosure love, friend, name, thank, share, good, new, year, hope, learn	Theme 1 (49.1) Self-disclosure time, make, love, good, work, know, want, people, get, new	Self-introduction -focusing on: - Brief demographic description of self - Friend who recommended to join What to share or exchange
Theme 3 (13.7) Personal narratives year, day, go, take, time, back, leave, come, home, move	Theme 3 (12.5) Personal narratives go, say, time, take, end, look, never, back, come, move	Anecdotal stories that they have experienced or heard; confessions
Theme 4 (11.0) Personal relationships life, music, love, family, child, believe, man, always, person, dream	Theme 4 (6.8) Personal relationships life, child, family, feel, mother, man, parent, girl, young, teacher	Love and faith in family, god, cultural assets and humanity
Theme 5 (8.9) Instrumental focus: Steemit platform use, create, content, post, video, account, project, comment, follow, website	Theme 2 (18.2) Instrumental focus: Steemit platform use, content, blockchain, project, create, account, user, support, platform, need	Strategies and tactics to do well on Steemit or to succeed using Steemit From the B-corpus - Strategies to game the system The advent of different user communities and different projects
Theme 6 (4.2) Personal history of education & work school, business, company, be, work, teacher, education, study, computer, student		Education and work experiences or introduction to a company, institution, or academic discipline

Theme 7 (3.0) Creative work travel, photo, art, picture, photography, draw, place, movie, artist, adventure	Theme 5 (5.3) Creative work music, art, draw, design, image, film, artist, style, paint, song	Showing the output of their own artistic or creative work (e.g., photos and drawing) or introduce with an artist identity
Theme 8 (1.5) Instrumental focus: coin and other cryptocurrencies crypto, invest, bitcoin, cryptocurrency, coin, exchange, trade, dollar, buy, money	Theme 6 (4.1) Instrumental focus: coin and coin market business, coin, market, bitcoin, crypto, cryptocurrency, company, money, service, invest	Cryptocurrency investment strategies and coin market prospects
Theme 9 (1.2) Lifestyle & leisure food, eat, cook, cat, recipe, colleague, fully, water, garden, meal	Theme 7 (3.2) Health & lifestyle die, water, eat, body, health, healthy, exist, happiness, videogame, food	From the M-corpus Living and lifestyle information such as food recipes with cooking tips, gardening, information about pets and so on. From the B-corpus Nutrition and health information, healthy eating, and food recipes
Theme 10 (0.5) *	Theme 8 (0.7) *	N/A

Table 2: Extracted topics, keywords, and description for M- and B-corpus

Note. M-corpus (N=26,413) and B-corpus (N=7,039).

Note. The number label of topics (1, 2, 3...) is allocated according to their proportion in the corpus; it corresponds to the topics numbered in Figure 5

* The theme that takes up less than 1% of the corpus was excluded from the analysis.

As for the B-corpus “self-disclosure” topic, users’ self-introduction posts tend to focus just on basic accounts of demographic information such as name, place of birth and living, where they study or what they do for living in a pretty concise manner, rather than contextually expressing themselves (see Appendix 2 for the examples of representative posts in details). As such, considering the keywords and classified posts according to each topic, the major changes that were made between the “same” topic of the Mania and Blow-off time phases have to do with the particular characteristics of the content that appear in each topic space. For example, although labeled similarly, there are qualitative differences between “health & lifestyle” of B-corpus (Topic 7) and “lifestyle & leisure” of M-corpus

(Topic 9). While the topic of M-corpus focused on what they had for a meal or cooking recipes, the B-corpus topic put more emphasis on health information by including healthy eating tips, nutrition, and general health information.

The variations in the types of personal information decreased in the case of B-corpus, probably being less excited given the decline in reward value. However, posting about the instrumental use of the platform (e.g., how to get more votes, how to trade STEEM coin, etc.) had greatly increased during the same period. Furthermore, as mentioned earlier, it was not only the content but the volume and distribution of topics that differ between the Mania phase and the Blow-off phase. Why might strategies vary across the two time periods? This will be expanded in the following section examining the STEEM currency value realized within the information space of the two phases.

Value and Volume of Information Space During and Post Bubble

Figure 8 shows how the proportion of each topic space has changed between the Mania phase and the Blow-off phase. Notably, during the Blow-off phase, incoming users focused more on the strategic and innovative use of Steemit platform, along with other instrumental and functional purposes such as the cryptocurrency and coin market, and health and lifestyle information. Compared to the Mania phase, the topics emphasizing “self” diminished in the Blow-off phase, as can be seen in the case of diminished topics regarding “self-expression” and “personal history of education and work.”

The reward value for publishing and curation activities dropped between the two phases. Overall, during the Mania phase, the average value paid for each post was 2.895 STEEM Dollars (SBD) with the standard deviation (SD) of 14.806; curation activities i.e., voting or commenting, specifically had an average value of 0.751 SBD with SD of 4.123.

For the Blow-off phase, the average value realized decreased, resulting in 1.312 SBD for each post with SD of 6.788 and 0.493 SBD for each curation activity with SD of 2.348.

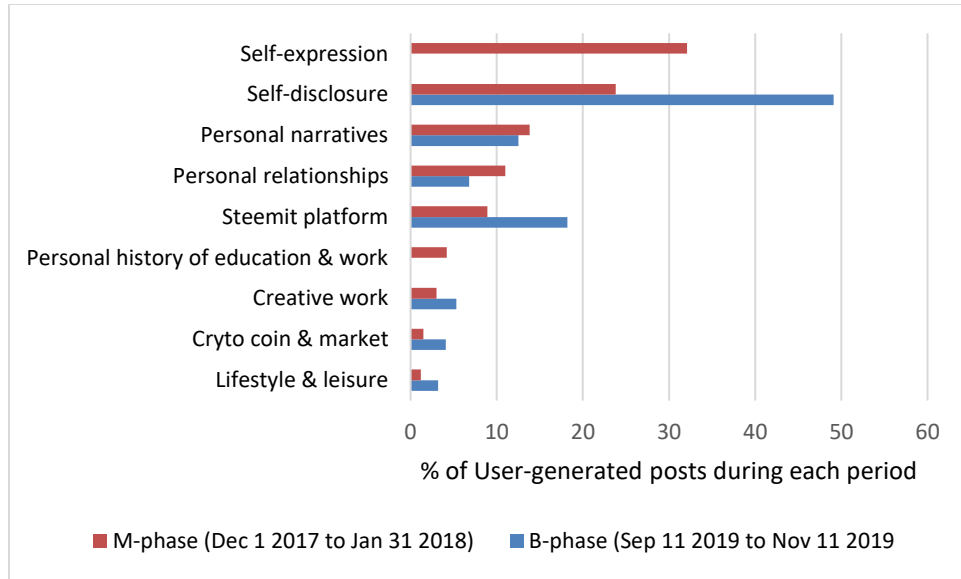


Figure 7: Change the information space from the Mania to the Blow-off phase

Tables 3 and 4 display how these posting and curating rewards are distributed across the two topic spaces, respectively. For M-corpus (Table 3), the topics that received the highest summed value were “self-expression”, “self-disclosure,” “the use of Steemit platform,” and “personal narratives” (in order). However, in regard to median reward value, posting related to the “self-disclosure” topic was highest, followed by “self-expression” and “personal relationships” (posts of “creative work” topic earned the highest reward, but the total number of posts was only 25). During the Blow-off phase, the trend of high value realization associated with “self-disclosure” continued as reported in Table 4. However, there were changes in rankings of value realization between topic spaces of the Mania phase and the Blow-off phase. For example, the rank of summed value for the

topic “use of Steemit platform” rose from fourth in the Mania phase to second place in the Blow-off phase (Table 3, 4). Except for the topic “self-disclosure,” the value realized through other topics entangled with information about oneself, such as “self-expression,” “personal relationships” and “work and job history,” completely disappeared or declined in the later Blow-off phase.

Dominant Topics					Unit: SBD*
	<i>N</i>	<i>M</i>	<i>SD</i>	50% percentile (median)	max
	<i>Sum</i>				
1. Self-expressions	12,267				
Post	41,823.107	3.409	15.249	0.196	668.408
Curation	11063.997	0.902	4.305	0.017	155.065
2. Self-disclosure	12,523				
Post	27,756.473	2.216	7.036	0.233	176.402
Curation	7007.523	0.560	2.034	0.020	73.277
3. Personal narratives	1,018				
Post	2,485.872	2.441	8.673	0.099	91.512
Curation	622.676	0.612	2.241	0.003	29.670
4. Personal relationships	215				
Post	550.485	2.560	8.521	0.174	87.066
Curation	140.378	0.653	2.167	0.021	19.932
5. Steemit platform	329				
Post	3,802.442	11.558	81.745	0.094	1192.040
Curation	1006.135	3.058	22.172	0.002	333.047
6. Personal history of education & work	24				
Post	10.756	0.448	0.721	0.088	2.922
Curation	2.011	0.084	0.173	0.004	0.683
7. Creative work	7				
Post	11.349	1.612	2.969	0.287	8.111
Curation	1.579	0.226	0.334	0.045	0.763
8. Coin and coin market	5				
Post	3.771	0.754	1.585	0.032	3.589
Curation	0.980	0.196	0.438	0	0.980
9. Lifestyle & leisure	25				
Post	9.765	0.391	1.027	0.048	4.928
Curation	2.461	0.098	0.296	0	1.418

Table 3: Posting and curating rewards realized in M-corpus

*STEEM Dollars (SBD) is a type of cryptocurrency which has a circulating supply of 7,050,644 SBD (as of April 20, 2021, \$7.97 USD for 1 SBD). At the time of posting, the price of 1 SBD ranged from \$2.4 to \$13.15 USD.

Dominant Topics					Unit: SBD
	<i>N</i> <i>Sum</i>	<i>M</i>	<i>SD</i>	50% percentile (median)	max
1. Self-disclosure	6,595				
Post	7,263.326	1.101	4.305	0.093	95.826
Curation	2,789.959	0.423	1.693	0.022	56.674
2. Steemit platform	325				
Post	1,834.565	5.645	24.462	0.092	251.037
Curation	631.943	1.944	7.643	0.020	74.566
3. Personal narratives	75				
Post	82.667	1.102	3.330	0.034	19.355
Curation	31.470	0.420	1.419	0.009	8.843
4. Personal relationships	6				
Post	1.615	0.269	0.384	0.043	0.770
Curation	0.500	0.083	0.120	0.012	0.246
5. Creative work	9				
Post	0.137	0.015	0.025	0	0.062
Curation	0.031	0.003	0.006	0	0.015
6. Coin and coin market	16				
Post	46.004	2.875	7.452	0.030	26.989
Curation	16.840	1.053	2.484	0.008	8.906
7. Health & lifestyle	13				
Post	5.036	0.387	1.130	0.031	4.128
Curation	1.539	0.118	0.347	0.010	1.265

Table 4: Posting and curating rewards realized in B-corpus.

Note. At the time of posting the price of 1 SBD ranged from \$ 0.61 to \$ 0.83 USD.

STUDY 1: CHAPTER DISCUSSION AND CONCLUSION

This first analysis explores Steemit users' value-making practices, focusing on how participants use information about themselves in their posts. The self-generated post, as a piece of composite information that often combines various aspects of oneself or one's daily life in an expectation of rewards, transforms into a special form of commodity that is self-governed within the incentive design of a blockchain social media platform. The results reveal not only the boundary of personal information that is voluntarily disclosed or shared when reward is assured, but also the characteristics of personal information topic

spaces. As many economists note, the inherent difficulty of studying personal data in the economic context has been bounded by the complexity of privacy itself, related to its susceptibility to context, human heuristics or bias (Acquisti, 2004; Spiekermann et al., 2015).

The research results demonstrate four primary conclusions. First, the extracted personal information topic spaces illustrate the popular subjects of self-representation consisting of abstract personal values and beliefs in life, memorable personal experiences or events, and personal relationships in addition to conventional demographic profile sharing. Second, comparing the emerging Mania-stage market and the late Blow-off stage market illustrates differences in strategies of self-display as well as the type, composition, and value of personal information invested/shared on the platform, presumably in part due to the lower reward value accruing to the Blow-off phase. Third, empirical evidence demonstrates the realized value for each topic. Fourth, the platform affords the possibility of unique practices of providing a relatively deep and idiosyncratic display of self, which I call the rise of the “unquantifiable self” or “qualified self,” echoing Humphries (2018).

To reiterate, the Steemit user is closer to a direct stakeholder who has the control to make autonomous and deliberate decisions over personal information, compared to users in typical social media environments. In this context, the identified two topic models for two different phases of Steemit demonstrate which types of personal information were widely shared as a result of collective information behaviors. Considering both the Mania and Blow-off phases’ topic spaces, we observe several common subjects of personal information shared on Steemit, including basic demographic information such as name, age, gender, place or country of birth or residence, family or friends, and personal stories based on their or their close ones’ memorable experiences or events, in whole or in part. A

small share of personal information space is taken up by creative or artistic works, hobbies, leisure activities, or health issues. It did not appear in both phases, but a considerable number of people showed a tendency to express who they are in depth by disclosing their personal traits, view of life, life goal and vision; information about personal relationships and educational or work history are invested as well, though less often. The findings appear to be fairly consistent with previous research in terms of the layers of private information disclosed on Twitter; namely, the layers comprise descriptions of daily life (least private), social identity, competence such as intelligence or social skills, socioeconomic status, and sensitive health information (most private) (Jin, 2013). However, such studies do not account for one's personal values and beliefs or personal narratives, which were a crucial part of identity representation in this study. As well, one more context is present here because sharing of personal information can be negotiated between value realization and privacy risks. According to this study's analysis of information shared, personal values and philosophies constitute the largest proportion of personal information, which was subsequently followed by social identities, memorable life experiences or events, information about personal relationships, and detailed work or educational backgrounds. This collective order of shared information may be the result of one's negotiation between risk and opportunity, balancing the most valuable part of self without sacrificing too much of one's sensitive personal attributes.

Furthermore, the study suggests a link between the prospects toward cryptocurrency market, including STEEM, and the type, composition, and value of publicly shared personal information space by focusing on the two contrasting time frames, namely the "Mania" phase, the crypto-bubble period when the aspirations of people reached their peak, and the "Blow-off" phase in which the bubble burst dampened people's enthusiasm

for cryptocurrency. Altogether, during the Blow-off phase, the influx of new users decreased, which significantly reduced the amount of invested personal information and realized value. With respect to the topics of the Blow-off phase, the important point to note is that there was a significant decrease in the amount as well as diversity of personal information contributed to the platform. The topics regarding self-expression and education and work history disappeared from the corpus of the Blow-off phase. This implies that human propensity to disclose personal information might be subject to external environmental factors, in turn demonstrating the flexibility in disclosing, sharing, and using information about themselves when information owners have actual control.

Additionally, the current study provides some evidence on the value realized and distributed in the topic spaces through the process of collective decision-making. Although not an exact equivalent of official currencies, the current analysis ranks the values of different areas of personal information, taking as an example the group of topics that realized the highest total value: self-expression, self-disclosure, and personal narrative. By demonstrating the economic value of less-marketable data such as personal values or life philosophies, the analysis shows the possibility of adding value to aspects of the self that are rarely datafied in the market. Importantly, the value realization of personal data occurred differently when people had rosy outlooks for the market (Mania phase) or when their perspective was bleak (Blow-off phase). Value realization for various areas of personal information increased when people saw an upturn in the cryptocurrency price, but as the price fell, the outcome of value realization behaviors decreased significantly. Interpreting the process of valuation is likely to be limited with respect to the unique system architecture of Steemit, but the current approach offers a benchmark value which is determined by free and democratic votes from community members.

Most importantly, a key finding from the topic modeling analysis is that there were a great number of “unquantifiable” illustrations of self (32.1%, the largest portion of the corpus, of “self-expressions” and 13.7% of “personal narratives” during the bubble period; 12.5% of “personal narratives” after the bubble burst) that portray one’s view or philosophy of life, life visions or goals, detailed personality trait(s), or memorable moments or experiences in life. Although there are still general categories of personally identifiable information such as demographic information, personal relationships, education and work experiences, and hobbies or leisure activities, many users appear to be selectively choosing what to show to others to express who they are and use different methods of delivery (e.g., narrative techniques). This practice is in line with Livingstone’s observation (2008), “This suggests a definition of privacy not tied to the disclosure of certain types of information, rather...having control over managing this disclosure” (pp.404-405), addressing the privacy paradox of youth who are concerned about privacy yet readily disclose personal information. Given the current research’s self-expression topic area in which expressed information was mainly about life philosophy, life vision or goal, personalities, or memorable life experiences, the results reveal a synergistic combination of “the types of information to disclose” and “control over how to manage this disclosure,” to sustain intimacy, construct stylistically-elaborated identity, and draw attention (Livingston, 2008) while protecting privacy.

These observed user practices of self-expression gain greater importance as the mission of global platforms and data collecting companies is closely interlaced with the collection of user data and “quantifying” the users as a basis of data-driven digital marketing and targeted advertising. Let’s go back to the earlier examples representing self-disclosure and self-expression, respectively: for the former, “...*So my name is XX, I was*

born in YY on 11 March 1999 and I live in.... And my profession is...My hobby is.... I have one brother and two sisters...” and for the latter, *“I never worry about things especially things I can’t change... I am a regular guy, I do not use fancy words or try to tell people how to keep up with the Joneses.”* Supposing you try to train a machine-learning model based on both types of information above to develop or advertise a product, which information would be more readily datafied, fed into the machine, and profiled as the target audience? In this sense, the considerable amount of “unquantifiable” self-display documented in the current topic model implies a multifaceted outlook for exercising personal data ownership under the incentive design of the blockchain-powered system.

The results of this study can be limited in several ways. First, the data have limitations. This study only analyzed the posts with self-introduction tags, a space usually recommended for newcomers once they signed up for the platform. Therefore, the current analysis might not be able to explain information disclosure practices as these users master the platform. Second, there are limits in tuning the topic model. Unlike other topic models that use personal pronouns (e.g., I, my, me) as a basic signal of self-disclosure, this study relied on the nature of the dataset itself, where the main subject is supposed to be information about oneself. Lastly, there is an inherent hardship in separating the effect of Steemit’s incentive design from the “default rule” set by existing social media platforms. Because users already may have been accustomed to the platform-centered culture in sharing and using their personal information, their practices on Steemit may be obfuscated by their previous experiences.

In spite of these limits, the significant implication of this study is a novel conception of data empowerment to strengthen ordinary individuals’ rights to control their data beyond “controlling” it through informed consent in the conventional platform architecture. By

setting Steemit's decentralized incentive design as an alternative framework of personal data governance, as well as a site of a natural experiment in terms of comparing time phases, the study provides empirical evidence on the topic spaces of private information, the benchmark value that "ranks" the types of certain personal information, and the notion of "unquantifiable self" whose strategies can defy the predatory marketing practices using algorithms. The experimental framework of a decentralized incentive design can be further questioned, investigated, and tested for ordinary citizens, policymakers, and digital entrepreneurs with a vision of a redistributive data economy.

Chapter 5: Data as Investment: Micro View

STUDY 2 METHOD: CONTENT ANALYSIS

Data and Methods

In researching everyday user activities such as posts, comments, or votes on Steemit as mechanisms to realize value, this phase of the investigation focuses on reward-seeking users' deliberate decision making around selective disclosure of personal information. It seeks to understand the strategic management of identities as embodied in their shared content. Content analysis investigates “privacy balancing strategies” —the balance between privacy, information sharing, and reward—of the users based on user-generated posts. The content analysis applied in this study is by and large quantitative, but it still has some qualitative elements in that the coding scheme involves latent content categories that demand interpretation on the side of the researcher.

The user-generated posts are drawn from the *blog* corpus and include those generated with the tag “blog” from the two-month period between December 1, 2017, to January 31, 2018. Given the presence of various tags on Steemit, the intentional selection of the “blog” tag begs further discussion. First, it is less context-specific, meaning that users can post whatever they want without having to tailor their content to be compatible with a certain topic. In other words, the generality of the tag is less likely to affect the user's preconceptions regarding disclosing their private information. Second, as this can reflect a wide range of users' interests, it offers a chance to observe shared or distinctive information practices in broader contexts. Last, posts attached to this tag are usually drawn from the user's sustained activities, which are relatively persistent beyond the initial introduction of

oneself to the platform. Thus, the blog content can be thought of as close to normal everyday environments in which people manage their information.

From a total of 155,988 user posts, I randomly selected 1,000 sample posts for in-depth analysis. Using Dedoose, a cross-platform application for analyzing qualitative and mixed-methods research data in various forms, all components that constitute a post from text to image, sound, video, hyperlinks, etc. were collected for analysis. Drawing on a thematic analysis approach (Braun & Clarke, 2006), the study framed and characterized the types of personal information used in user-generated content and how users managed it. The premise of content coding sampled posts rests in how they may reflect strategic choices in generating value.

Conceptualizing the enactment of identities as opportunities, i.e., value realization, and risks, i.e., privacy, the coding scheme (see Table 5) applies the following five different criteria: 1) mode of communication: the use of certain or multiple media elements in a post aimed to “stylize” the post or the statement of identities; this includes a category of “selfie” and “v-logging” for posts, 2) type of authorship, depending on whether the post is fully self-generated, remixed, or copied, 3) online behavior management boundaries (OBMB): the way people enhance or verify their personal identities and professional identities. In terms of the use of multimedia elements and authorship, this may be the factors related to one’s creativity, originality, and investments, which help to improve value realization opportunities to better elaborate oneself or to craft more compelling content. For the OBMB construct, the measures address both opportunities and risks in managing one’s identities online; by disclosing both personal and professional aspects of life online, people can have more degrees of freedom in portraying themselves compared to revealing only one side of themselves, but it also means that they are at a higher risk of losing privacy.

Likewise, depending on whether one seeks to embellish oneself (self-enhancement with positive self-disclosure) or make oneself appear more authentic (self-verification by expressing both positive and negative aspects of self), it can also be linked to opportunities and risks, but in more complicated ways.

The study also includes two different types of personal information that, imbued with other information, help to identify a particular person in SNS context: 4) implicit self-presentational information and 5) explicit self-presentational information. As stated earlier, while both implicit and explicit information about oneself presented in the form of UGCs can be used to identify a particular person, they reveal different levels of richness in self-disclosure. Implicit forms of information, primarily based on user's immediate actions, offer hints to *infer* the users' identities by showing places they've been or visited, people they've met, activities they've been involved in, and issues they are interested in. On the other hand, explicit forms of information construct identity information that is *directly provided users to show* who they are. In general, as an auxiliary or main component of content, direct identity information such as name, nationality, membership, education, family, or work information is elaborated in much greater depth than appears in user profiles. Given that rationale, it may be more rewarding and riskier to express identities with explicit information than implicit information, though both still constitute an aspect of identity. Items for each construct are detailed below.

Criteria	Description
Mode of Communication (Panahi, Watson, & Partridge, 2012)	<p>When a post applies single or multiple content forms, media elements, or different modes of communication, such as:</p> <ul style="list-style-type: none"> • Texts (only): when only text was used in a post • Links • Audios • Photos or images • Videos

	<ul style="list-style-type: none"> • Other <p>When a post includes visualized self-representation, such as selfie or self-hosted videos featuring self:</p> <ul style="list-style-type: none"> • Selfie • Vlog
Posting type by authorship (Jin, 2013; Flath, Friesike, Wirth, & Thiesse, 2017)	<p>Classify a post by the source of information input, primarily based on its authorship:</p> <ul style="list-style-type: none"> • Self-generated content - created or produced by the author • Remixed content - remix, any transformation or appropriation observed • Other-generated content (OGC), i.e., copied or “resteemed” • Unknown - when the source is not clear
Online behavior management boundaries (Batenburg, & Bartels, 2017; Livingstone, 2008)	<p>Rooted in the theorization of Online Behavior Management Boundaries (OBMB), the code observes how one’s personal and professional identities are integrated or separated in a post and the ways these identities are expressed.</p> <ul style="list-style-type: none"> • Integrated: when personal-professional identities appeared combined • Separated: when personal-professional identities appeared separated • Verified: when one discloses negative information about self as well as positive information • Enhanced: when only positive information is disclosed <p><i>Note. Here “professional identity” is an identity that is closely related to one’s primary professional job or business, work for a living, or other income activities.</i></p>
Implicit self-presentational information (Jeong, & Coyle, 2014; Jeong, & Kim, 2017)	<p>Personal information presented in postings that provides implicit cues to infer an individual</p> <ul style="list-style-type: none"> • Location-based posts (i.e., where you are) • Action-oriented posts (i.e., what you are engaged in) • Social activity posts (i.e., who you spend time with) • General information posts (i.e., general posts about news, sports, other public affairs, etc.) • Personal information posts (i.e., personal relationships or personal events)
Explicit self-presentational information (Jin, 2013)	<p>Personal information disclosed in the context of life presented in postings that provides explicit cues to identify an individual in depth</p> <ul style="list-style-type: none"> • (Layer 1 Daily life and entertainment) favorite foods, restaurants, music, movies, and travel • (Layer 2 Social identity) school, occupation, group memberships, and gender; education, political affiliation, socio-economic status

-
- (Layer 3 Competence) intelligence, motivation, strengths, and social skills, including *creative or artistic activities*
 - (Layer 4 Personal relationship) family, friends, and other close ones.
 - (Layer 5 Health) mental health and physical health
-

Table 5: Coding scheme.

For the sampled 1,000 posts, two coders examined and classified all elements of each post using the coding scheme displayed here. To ensure the reliability of the coding of the posts, intercoder reliability assessment was based on selecting and checking 25 randomly selected posts. The study used Cohen's kappa (κ) as a measure of intercoder agreement as it addresses the limits of a simple percent agreement that does not account for agreement that could occur by chance (McHugh, 2012). Most criteria for the coding frame for this study indicate an acceptable level of consistency between the coders with a Cohen's Kappa score range from moderate to substantial levels (from 0.46 to 0.84) (McHugh, 2012). Implicit categories for codes have lower reliabilities than explicit categories.

STUDY 2 RESULT: STRATEGIC ASSEMBLAGE OF SELF IN USER-GENERATED CONTENT

After the initial coding, a total number of 825 posts are included for the analysis, excluding posts that are not written in the English language (175 posts). Of the 825 research samples, 335 posts contained personal information, whereas 470 posts did not. The latter posts generally include anything people want to share: what's happening on Steemit, local or international news, coin or stock market reports, food, plants, animals or insects, best destinations for travels, or promotion of new cryptocurrency-based projects or other digital businesses among many others. For this portion of the sample posts lacking personal

information, the study only coded for mode of communication and posting types by authorship. Finally, when coding and reporting the results, we anonymized the authors of the posts, turning their user IDs into newly assigned study IDs to protect their privacy.

For a considerable proportion of Steemit posts (335 posts, 43%), intentional presentation of personal information as a part of their postings was a common practice. In addition to analyzing the posts according to the pre-established criteria of online identity management behaviors including identity separation (250 posts), integration (64 posts), enhancement (13 posts) or verification (26 posts) practices, the study slightly expanded the coding frame by using a “competence” criterion of the explicit personal information category; it labels “armature creation” for the posts primarily generated for fun or simple sharing (65 posts), separated from the posts labeled as “professionals,” which are driven by self-branding or promoting purposes or other forms of profit-making activities based on the author’s occupation or expertise (11 posts). Moreover, in terms of visualized self-representation, there were 55 posts with selfies and 20 posts by v-loggers, respectively. Selfie and v-log posts are separately coded because some of v-loggers were anonymously blogging (e.g., using voice only, or using costumes that completely cover face and body).

Table 6 shows the comprehensive characteristics of the posts with personal information that can be differentiated not only from each other but from the posts that contain no personal information.

The posts without personal information account for more than half (470 posts, 57%) of the randomly sampled posts. For example, some of them observe and report issues of the Steemit community: *“This article utilizes publicly available data to capture suspicious users in Steemit, where writers and subscribers are rewarded through posting an article and/or voting an article....”* There are also some other posts focusing on a soft topic such

as the health benefits of fruits: *“Pomegranate is not only tasty, but extremely useful fruit. It can help in the battle with various diseases, improve your overall health and raise your mood. For example...”* Given this, a substantial number of Steemit users have generated content in a way that does not risk their privacy while still pursuing possible incentives or rewards.

A common characteristic of both posts with or without information about oneself is found in the mode of communication. Both types of the posts were actively using multimedia elements in their creation comprising visual components such as photos, images, moving animated pictures (e.g., gif) or video clips, hyperlinks, texts, or audio, features the most prevalent being photo or other images (approximately 88% on average) followed by the use of hyperlinks to other sites, web sources, or social media accounts and video clips (about 34.6% and 24.2% for each). Although the texts were still an integral part of a number of posts, text-only content was very rare. Above all, the most distinctive characteristic of the posts with personal information in this study was the type of authorship; beyond remixing, when a user in any way enacts her identity in a post, it is very likely that a significant portion of the post was originated by the user herself (range from 78.5% to 100%). On the other hand, when no personal detail appears in the post, the user tends less to the very creator of that post (25.5%); it is likely that a part of the post (28.9%) or the entire post (16%) is from other sources, or even the source of the content is unknown (22.3%).

Apart from the non-personal posts, the study examined the three conspicuous practices that represent unique identity management strategies of the posts presenting personal information: 1) identity separation, 2) identity integration, and 3) information-balancing. First, the mode of identity separation was the most frequently implemented

practice in the posts in which the users intended to separate personal and professional identities; they tend to selectively disclose only one aspect of identity rather than representing both. Second, when these personal-professional identities are displayed as an integrated form, these are not likely to occur by chance - there were specific goal-oriented motivations and purposes that the user aimed to achieve through the posting. In doing so, it differently engages with the practices of identity-enhancement and identity-verification. Lastly, the study found “information balancing” efforts among the posts with visual information about oneself such as selfie or self-featured videos: when the visualized personal content is the key component of the posting, these posts were less likely to disclose personal information using other means of communication - for example, less textual or linked description of the self, presumably out of concern about revealing too much.

Organizational Boundary Management Behaviors					Competence-Resourced		Viz-Centered		Not Personal		
(%)					(%)		UGCs (%)				
		Separator	Integrator	Enhancer	Verifier	Amateur	Pro	Selfie	V-logging	κ	
Mode of Communication											
• Audios		1 (0.4)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	4 (0.9)	*
• Links		82 (32.8)	27 (42.2)	5 (38.5)	6 (23.1)	25 (38.5)	6 (54.5)	13 (23.6)	7 (35)	107 (22.8)	0.6888
• Photos or images		230 (92)	59 (92.2)	12 (92.3)	26 (100)	59 (90.8)	9 (81.8)	55 (100)	14 (70)	376 (80)	0.6479
• Texts		10 (4)	4 (6.3)	0 (0)	3 (11.5)	3 (4.6)	0 (0)	0 (0)	0 (0)	26 (5.5)	*
• Videos		22 (8.8)	9 (14.1)	2 (15.4)	1 (3.8)	7 (10.8)	3 (27.3)	3 (5.5)	20 (100)	58 (12.3)	*
Authorship											
• Self-generated		210 (84)	59 (92.2)	12 (92.3)	25 (96.2)	51 (78.5)	9 (81.8)	53 (96.4)	20 (100)	120 (25.5)	0.6377
• Remixed		46 (18.4)	8 (12.5)	3 (23.1)	4 (15.4)	17 (26.2)	2 (18.2)	2 (3.6)	0 (0)	136 (28.9)	
• By others		0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	75 (16)	
• Unknown		1 (0.4)	0 (0)	0 (0)	0 (0)	0 (0)	1 (9.1)	0 (0)	0 (0)	105 (22.3)	
OBMB											
• Integrated		0 (0)		5 (38.5)	11 (42.3)	19 (29.2)	8 (72.7)	15 (27.3)	7 (35)	0.6137	
• Separated			0 (0)	8 (61.5)	15 (57.7)	40 (61.5)	3 (27.3)	38 (69.1)	12 (60)		
• Enhancing		8 (3.2)	5 (7.8)		0 (0)	10 (15.4)	0 (0)	1 (1.8)	1 (5)		
• Verifying		15 (6)	11 (17.2)	0 (0)		8 (12.3)	2 (18.2)	2 (3.6)	0 (0)		
Implicit											
• Location-based		53 (21.2)	19 (29.7)	2 (15.4)	4 (15.4)	11 (16.9)	0 (0)	16 (29.1)	8 (40)	N/A	0.6512
• Action-oriented		192 (76.8)	49 (76.6)	11 (84.6)	20 (76.9)	58 (89.2)	11 (100)	31 (56.4)	16 (80)		0.5955
• Social activity		32 (12.8)	8 (12.5)	1 (7.7)	3 (11.5)	6 (9.2)	1 (9.1)	9 (16.4)	3 (15)		0.6479
• General info		21 (8.4)	2 (3.1)	1 (7.7)	0 (0)	4 (6.2)	0 (0)	1 (1.8)	1 (5)		0.6575
• Personal info		71 (28.4)	44 (68.8)	5 (38.5)	19 (73.1)	26 (40)	7 (63.6)	29 (52.7)	6 (30)		0.5902
Explicit											
• Daily life and ent.		182 (72.8)	28 (43.8)	3 (23.1)	12 (46.2)	23 (35.4)	2 (18.2)	33 (60)	13 (65)	0.8377	0.8377
• Social identity		44 (17.6)	53 (82.8)	5 (38.5)	11 (42.3)	23 (35.4)	8 (72.7)	30 (54.5)	7 (35)		0.6888
• Competence		42 (16.8)	25 (39.1)	10 (76.9)	8 (30.8)			5 (9.1)	6 (30)		0.5614
• Relationship		48 (19.2)	19 (29.1)	3 (23.1)	7 (26.9)	8 (12.3)	1 (9.1)	19 (34.5)	3 (15)		0.4565
• Health		11 (4.4)	4 (6.3)	0 (0)	6 (23.1)	3 (4.6)	0 (0)	2 (3.6)	1 (5)		0.6479
N		335 (Each post can apply to more than one category)								470	
		250	64	13	26	65	11	55	20		

Table 6: Coding results.

Note. Kappa was not computed for audio, text, and video components due to its very low usage.

Separation of Professional and Personal Identities

People construct or dissolve boundaries between work and private life by managing professional and personal identities (Batenburg & Bartels, 2017). The increasingly digital world facilitates blurring these work-private boundaries, and more and more people enact their identity by combining professional and personal self on social media platforms. Unlike Facebook or LinkedIn, which are heavily based on existing offline personal/professional contacts, however, a considerable number of Steemit posts (250 posts) tend not to disclose the author's private and professional identities simultaneously. Most of the user-generated posts are more likely to illustrate clear boundaries between private and professional self.

When only professional identities are revealed, the main purpose of posting is to provide work-related information, advice, or suggestions or self-promotion; these often come with professional contact information or a sample of previous work as below. For example, this post also offered a link to a law firm in India where the author's professional contact is released.

"... (the sample advice) is the general question asked by people to which I have answered.... It is just a legal advice and would recommend checking from your source too. Feel Free to Contact me for any legal advice irrespective of any country."

The primary objective of this type of identity representation is to show expertise and capability and to promote what individuals can contribute to the community so that they can increase followers, votes, and reputation. or possibly, reach out to potential clients.

Alternatively, some workers provide their work-related backstory that is often sensitive in terms of the type of work itself or subject matter. For example, a sex worker's complains about the ignorance and carelessness of the "clients:" "...*You have no idea how infuriating it is as a sex worker, to have clients that have no idea what they're doing or what they want...*" or a front desk receptionist's experience about strange guests. While this gives a little glimpse of the industry with great details about their work experiences, their private identity, or actual name and other demographic information are concealed. In such contexts, one's active management of boundaries allows more purposeful refining of identity and prevents unnecessary (or inappropriate) information disclosure. The exclusion of "personal" aspects can be a strategic decision not only to secure some personal space but to look more professional.

As shown in Table 6, however, it can be assumed that the posts in which users express their identities separately tend to more heavily focus on personal rather than professional identities considering the prevalent sharing of daily life observations (72.8%) rather than social identities (17.6%) or competence (16.8%). This type of posting also relies on things or activities the writer had or has engaged in (i.e., action-based, 76.8%). Such posts are less likely to rely on content creators' demographic information, social identities, affiliations, or other personal information than on activities, hobbies, or other actions-based information.

Closely looking at the content of the posts that contain only personal identities, many posts with implicit personal information rely on self-expression that takes a "snapshot" of an aspect of personal life; such posts tend to focus on exhibiting a selected part of one's daily life to others such as what they had for a meal, a picture of scenery while taking a daily walk, small daily routines, or a memorable moment from past experiences

and the like. Even on the premise of sharing, this type of content generation was often similar to a daily journal or record to just talk to oneself. It usually contains less text, one or two pictures, and a short description or reflection; so much of the context is omitted or veiled that it may almost seem like a code that only the creator can read like the below example. Based on this example, it is hard to clearly picture the writer's identity. A representative post might be as follows:

(With a picture of a mountain somewhere on the way to the moon lake, known as, Chandra taal, Spiti, HP, India) *"I so miss these mountains, they gave me peace which I need the most in the current phase of my life. To Spiti, Till we meet again..."*

By comparison, the posts explicitly presenting personal information tend to embed direct personal information in the context of the content. They usually expand on topics similar to the above with "implicit" information but in greater detail and context associated with direct personal information, and attempt to provide further recommendations, suggestions, or other tips based on the represented experiences and events. Trip or activity reviews, house remodeling, food recipes, and other do-it-yourself (DIY) activities are some representative examples of this type of posting. These are structured by synthesizing heavy texts, several photos or images, and links that support and complement the content. Many of these multimedia materials were original content created by the author. In doing this, intentionally or unwittingly, people blend some personal details in context; this conveys their experience more vividly, leads others to familiarize and sympathize with the content as well as the writer, and readily adds a touch of originality to make their story unique. The

following illustrative post showcases an exciting life experience with a very thorough review and lots of pictures capturing the process, including herself.

“...as I publish you will be able to realize where I am, what I like to do, what I dedicate myself, etc. And that's why I chose one of the best moments I experienced during 2017 and I have not detailed it so much to anyone until now, that it was “Fly in Paragliding”.”

The practices of identity enhancement (3.2%) and identity verification (6%) were insignificant in the posts separating identities. Rather than carefully managing the positivity and negativity of the information about oneself in the content, the posts of this category mainly attempted to describe what has happened or what one experienced. Overall, identity separation practices reflect a standard form of day-to-day UGC that can be created by the majority of Steemit users who are not celebrities or influencers, with no special knowledge or experience, and without any advanced technical background or skills. Compared to users who disclose more about themselves – those who are open to represent their personal as well as professional self - they were less likely to risk their privacy for a greater valuation opportunity. Table 7 later in this chapter shows how different degrees of value realization have resulted from such practices.

Identity Integration Practices

Another strategy of enacting identity, though to a lesser extent (64 posts), integrates personal and professional identities in a post. This way of identity representation generally involves more self-disclosure than an identity separation strategy, including some personal

detail consisting of social identities and personal relationships among many others. While the use of multimedia components such as photos, images, hyperlinks, or videos are similar to posts that maintain boundaries between personal and professional identities, these posts tend to contain more original content. In other words, when the lines between personal and professional life are blurred in a post, this is more likely to be generated entirely by the author herself (92.2%).

What drives more disclosure? Relationships people forge on Steemit are intrinsically less dependent on their offline work or personal contacts, so people can more freely unleash their self-expression. Based on the analysis of user-generated posts, there were two major conditions of identity integration.

One approach pursues verifying identities through posting by using both negative and positive information about oneself (26 posts) - when the author wants to be considered favorable by others, despite the fact that her pre-existing self-conceptions contain personal flaws or other negative information about herself. This contains some sensitive topics such as personal histories, family histories, or health - drug or drinking - problems.

“I did not see my daughters for about 7 years...The judge then promptly ruled that we were unfit parents and that my Mother would have permanent custody of my daughters. Our rights were terminated as parents... I want to be clear...I was drinking and using drugs when this whole thing happened, but I did not use in front of the children. They were not physically or sexually abused...”

(a long story about “a hard-core drug addict that has been redeemed by Jesus”)

“My Bosses have given me permission to mention the name of our Ministry in my

blogs. I work at Endeavor House Ministries Inc. in Lansing, Michigan. If you look us up you can see us on Facebook, and at <https://endeavorhouseministries.com>.”

Such disclosures were most likely to use text as the main element of a post, along with decorative uses of existing images or pictures found on the web. This sort of “testimony” seeks out supportive feedback or reaction from other users; it is often to defend their position in life and in what they have done or to recharge and regain strength to keep moving forward to challenge or overcome their current difficult situations.

A second identity integration practice can be found in postings for self-marketing and personal branding (11 posts), especially for postings related to a career in areas where self-disclosure is encouraged or becomes an unavoidable part of the content. Vlogging or indie artists’ self-promotional posts, for example, tend to be more open to show themselves to others. This type of posting is characterized not only by its emphasis on creative, artistic, or intelligent competence or expertise but also by its relation to paid occupations regardless of whether it is the main source of revenue: “...*I am a California Singer-Songwriter and multi-genre collaborative music artist. I run a collaborative artist music house, where I promote and create original music with producers and artists...*” From disclosing instantaneous daily routines or events to multifaceted social identities in a network of social connections with family, friends, or other close personal and professional contacts, this type of post brings a sense of “authenticity” in utilizing pieces of personal information. That is, these posts tend to share not only positive information about the author herself but also negative information (self-verification) rather than only sharing their positives (self-enhancement): from an indie music artist: “...*There was a time the music stopped... my sister’s funeral changed everything for me..(and some struggles)..a new chapter in my life*”

began, it feels good to sing my own song (and a link to an album).” It can be a strategy to increase intimacy, seek legitimacy, and build trust between the author and her readers so that she can grow followers, subscribers, or readership.

Interestingly enough, such disclosure of identity information decreases as these posts, especially rooted in the authors’ artistic competence or skill, are created mainly for fun but not tied to self-marketing or self-promotion efforts. This “amateur” style of posting (65 posts) reveals only positive aspects of oneself while actively managing the boundaries between personal and professional lives. For these posts, an online platform is a space for freedom of artistic expressions with less fear of being criticized. Such a context could be an explanation for these amateur posts, which tend to present either personal or professional aspects of self (61.5%) rather than both personal and professional aspects together (29.2%) compared to the “professional” posts with an apparent motivation to ‘sell oneself.’

Visual Representations of Self: Information Balancing Practices

The results indicate that a fair share of user posts in this study (55 posts) use the selfie as a major part of a typical post, and 20 posts contain self-generated videos or v-log channels hosted by the users themselves. The self-taken photo or UGC videos about oneself can be more revealing than a textual description, conveying more manifest or nuanced information such as physical appearance, personal/social affiliation or group membership, or other background objects or circumstances. Based on this, other users can see who the author is (e.g., gender, age, race/ethnicity, or other biometric information), who she is affiliated with, her location information, and/or what activities or events she is engaged in.

The revealing nature of visual self-representations invokes several noteworthy practices commonly associated with content creation tactics as well as with personal information management, which this study calls “information balancing practices.” Simply put, when selfies are posted here, other forms of identity expressions tend to be somewhat limited. Selfies or v-logging posts were more likely to be presented with one or two sentences of short descriptions or reflections without rich text, audio, or other linked materials associated with self-disclosure information. This type of post maintains boundaries between personal identities and professional identities. Besides, rather than disclosing information accumulated by life experience, these posts rely more on immediate actions or simple displays of an aspect of daily life such as what one does, where one is, or who with whom one spends time. Considering the nature of visual self-representation, this practice can be understood to be an extension of the effort to avoid excessive exposure of oneself by striking a balance between visuals or other communicative elements in a post.

Various Identity Practices and Realized Values

	<i>N</i>	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>	<i>Implicit</i>	<i>Explicit</i>
Personal info**							
Not included	470						
Curation		0.91	5.49	0	94.64		
Author		3.63	18.72	0	294.12		
Included	335						
Curation		2.04	5.73	0	55.81		
Author		8.33	20.64	0	177.26	1.50	1.47
By authorship**							
Self-generated	400						
Curation		1.77	5.91	0	61.23		
Author		7.05	21.25	0	240.00	1.51	1.45
Mixed	189						
Curation		1.70	7.57	0	94.64		
Author		7.16	25.47	0	294.14	1.49	1.55

Copied/unknown	181						
Curation		0.41	1.30	0	11.94		
Author		1.67	4.48	0	34.66		
OBMB*							
Separated	250						
Curation		1.79	5.30	0	55.81		
Author		7.66	19.63	0	177.26	1.42	1.34
Integrated	45						
Curation		3.76	8.12	0	43.75		
Author		13.61	28.32	0	143.25	1.93	2.00
Verification	26						
Curation		1.76	5.55	0	27.91		
Author		6.13	18.37	0	92.14	1.77	1.77
Enhancement	13						
Curation		2.31	4.83	0	17.90		
Author		8.91	18.07	0	66.63	1.58	1.62
Viz	73						
Curation		2.99	7.85	0	55.81		
Author		10.96	26.55	0	177.26	1.67	1.65
Competence							
Amateur	65						
Curation		3.75	9.27	0	55.81		
Author		15.59	32.56	0	177.26	1.63	1.88
Professional	11						
Curation		2.74	5.70	0	18.68		
Author		10.96	24.95	0	84.11	1.73	2.00

Table 7: Identity management practices and realized value in STEEM currency

Note. Unit: SBD.

Note. The *implicit* personal information does not include the “general information” item in the analysis.

Note. * categories show a significant difference between mean curation rewards. ** categories show significant difference in both mean curation rewards and mean author payout rewards.

Table 7 shows how values are realized according to the identity managing practices on Steemit. To reiterate, curation reward is designed to allocate the rewards given to a post to users who voted for this post; the earlier you upvote for a popular post, the higher your reward. Author reward is the reward for the author of posts. Table 7 also includes the number of facets of implicit- and explicit self-presentational information that were shared

through authoring activities. It indicates that the posts with personal information earned significantly higher rewards than those without (curation rewards: mean difference = 1.13, $t(803) = 2.83$, $p < 0.01$; author rewards: mean difference = 4.70, $t(803) = 3.37$, $p < 0.01$). On average, posts containing personal information shared 1.5 implicit and 1.47 explicit facets of personal information.

As presented in the conceptual framework, the posts with original touches are likely to receive a higher reward than copied work. A one-way analysis of variance is used to determine any statistical evidence that the associated mean rewards for each posting type are significantly different. There was a significant difference in realized values depending on whether it is created by the author or not (curation rewards: $F(2, 767) = 3.78$, $p < 0.05$; author rewards: $F(2, 767) = 5.10$, $p < 0.01$). Interestingly enough, the realized values did not differ whether it is entirely self-generated or remixed.

In managing the boundaries of identity between the personal and professional self, the curation values realized became significantly greater when both sides of identity are expressed together than when separated (mean difference = 1.97, $t(293) = 2.10$, $p < 0.05$). Although only marginally significant, the same logic applies to the author rewards, with higher rewards for personal and professional self-integration (mean difference = 5.95, $t(293) = 1.74$, $p < 0.1$). Additionally, when both aspects of the self are integrated, the posts tend to be more rewarded as well as revealing more personal information (1.93 implicit and 2 explicit aspects of personal information) than any other type of expressing identities. Similarly, amateur posting focused on the authors' artistic abilities or other skills received the highest average rewards. Although there were no statistically significant differences, the rewards realized for amateur postings were numerically higher than those of professionals. Taken together, among many identity management practices, competence-

driven amateur posting, posts that combine personal and professional identities, posts with visualized personal content, and posts created by professionals received higher rewards.

STUDY 2: CHAPTER DISCUSSION AND CONCLUSION

For everyday content generators, the use of information about oneself can be an effective and convenient strategy for creating content. In the Steemit environment, creating content using personal information can effectively create value for individuals. Hence, examining the content of posts can directly provide an indication of value-creation strategies and priorities. Drawing on self-identity and aspects of daily life, many ordinary users create original content for community engagement and more rewards in the Steemit platform. This examination probed various practices of identity expressions and personal information management in user-generated posts.

Self-posting activities offer new opportunities and challenges for controlling self-identities in online spaces. On one hand, posting can include another piece of personal information subject to commodification by large companies' automated data tracking and sophisticated data mining techniques in ubiquitous technological environments. On the other hand, it can provide a new chance to control personal online presence for various purposes by experimenting with what and how much personal information is posted online in order to achieve a certain goal. Steemit offers a direct way of looking at rewards.

Even within the unique incentive structure of blockchain-powered social media, people with different goals take varied positions in the creative domain and negotiate their privacy boundaries accordingly. The findings focus on three key identity enactment strategies: identity separation, identity integration, and information balancing practices. In

regard to the value realized by practicing each of these identity management strategies, posts presenting an integrated identity tend to be rewarded higher than those separated. In addition, regardless of various multimedia presentations, posts with visualized personal content tend to be highly valued the platform.

First, most user posts chose to adopt an identity-separation strategy by keeping a distance between professional and personal life in their posts, presumably in order to safeguard privacy. The extent of details about oneself illustrate two different forms of self-display, namely reflective and self-display practices, which have a different communicative intention (inward v. outward).

Less frequently, but with more inferential personal details across various life domains, other user posts represented both professional/personal identities together under some specific circumstances, presumably in the desire for self-verification and self-promotion. Self-verification posts tend to contain very sensitive personal information about one topic in detail while self-promotion posts disclose various celebratory aspects of self.

Finally, visual identity expressions such as selfies are associated with limited use of other modes of communication. Posts with selfies are inclined to simply describe what is in the picture, without other personal details in texts, audios, videos, or hyperlinks.

At the intersection of privacy protection, content creation, and the gig economy, people have different motivations and attempt to take a different role in the Steemit ecosystem. As a result, framing and placing oneself somewhere between a private person, a content creator, and a gig worker influences one's strategy to express and manage one's identities using different layers of personal information. Furthermore, posting represents the different content strategies of average users who "sell" themselves on social media, not those of celebrities or influencers. This study finds that users deploy storytelling and

diverse modes of communication to use alongside descriptive or implied personal information to enact identities in a more compelling or balanced way. While the current findings could not provide evidence that shows whether the practices of Steemit users are different from what they do in other social media environments, it can be suggested that many people on Steemit are still privacy-conscious, and that opportunities and risks are carefully negotiated in their posting activities even in the platform's reward design.

Self-monetization in this era of UGC becomes one prominent form of online engagement in digital space. Despite its negative connotation, the identity practices exercised within the context of tangible incentives imply that the promise of rewards did not always result in indiscreet self-exposure. Indeed, the control of online personal presence requires more than individual practices at user levels. Within the business logic of large tech companies like YouTube and Facebook, user agency in information control will be in conflict with their excessive data collection, sophisticated data mining, and targeted profiling strategies. On the contrary, Steemit, as a social media platform as well as a self-marketing platform, provides an alternative environment where user agency can be less bounded or obfuscated by these business methods of centralized authorities.

All in all, the above findings contribute to extending the multiplicity of user agency in the position of content/data generator: at its broadest, content creation can be a combination of creative efforts and identity expressions with some (unstructured) personal data. The observed user practices of identity management can be seen as another entanglement of data privacy and user agency (Kennedy, Poell, & van Dijck, 2015) that highlights a process of data generation that enhances the agency of individual users. As well, the study can complicate the conception of personal branding in relation to the information regarding work and private life (Ngai, Tao, & Moon, 2015). The self-enacted

position between professionalism and amateurship can serve as the first step in deciding which side of self to disclose with what personal details for many content-generating users. As a growing number of people seek direct/indirect benefits from their creations on many social media platforms, an in-depth understanding of user's strategic identity management and personal information tactics will be increasingly important to investigate the user agency in power dynamic around personal data.

Chapter 6: Technological Condition for User Success

The goal of the survey is to investigate how one's digital behavior is embedded within the architecture of technology in our everyday lives. It uses the concepts of competence and dependence as both affect actual digital use behaviors and performance on platform. As already discussed in prior chapters, digital behavior, called user performance here, becomes an indicator of users' success on the platform. It is comprised of a set of tangible and intangible values such as the number of followers, reputation, productivity, and estimated account value. As the research setting is deeply ingrained in social media environments, the study explores the role played by privacy and data right awareness in the user performance system. This approach allows us to provide a glimpse of the promises and realities of the Steemit model as one alternative in which people may explore and realize the potential of a data economy on their own.

The study responds to the following questions and hypotheses.

3.1 How do technological competence and dependence affect value-making performance?

H 3.1.1 Technological dependence is associated with technological competence.

H 3.1.2 Technological competence positively affects the use of Steemit/Internet.

H 3.1.3 Technological dependence positively affects the use of Steemit/Internet.

H 3.1.4 Steemit/Internet use positively affects value-making performance.

H 3.1.5 Technological dependence negatively affects the relationship between technological competence and Steemit/Internet uses.

3.2 How do technological experiences and capabilities affect value-making performance in relation to awareness of data rights?

H 3.2.1 Technological competence positively affects awareness of privacy and data rights.

H 3.2.2 Technological dependence has negative or no impact on awareness of privacy and data rights.

H 3.2.3 Awareness of privacy and data rights negatively affect the use of Steemit/Internet.

H 3.2.4 The relationship between technological competence and Steemit/Internet use is negatively mediated by awareness of privacy and data rights.

STUDY 3 METHOD: ONLINE USER SURVEY

Data Collection

Steemit users were surveyed from January, 2021 to mid-May, 2021 in order to learn the extent to which digital technologies has penetrated every aspect of life, and how this embeddedness influences value-making performance on Steemit, a rewarding social media platform. The participants of this survey consist of people 18 years of age and older who use English language from different parts of the world.

I used the daily active wallet (DAW), which indicates the number of unique digital currency wallet addresses interacting with the Steemit platform, as the community

population. As of April in 2021, the number of DAW was 17.66K., A power analysis suggests the sample size of 150 responses is adequate.¹¹ A total of 153 valid responses were collected, based on the researcher's recruitment posts and comments that open to any Steemit users who are interested in participating the survey.

Survey Instruments

The goal of the online survey with participating Steemit users is to ascertain their Steemit experience and technology expertise. The survey questionnaire consists of five different sections:

- The first section assesses the users' Steemit usage, including use frequency and the number of social media accounts that are linked with their Steemit account (SU).
- The second section investigates users' general Internet usage (IU).
- The third section assesses factors regarding technological capabilities (CP) and perceived dependency of technology (PD).
- The fourth section examines users' concern for information privacy (CFIP).
- Finally, the survey collects the users' demographic information.

To assess Dependency (PD) components, the survey items were drawn from the Technology Access and Digital Inclusion project developed by the research team at the department of Radio-TV-Film at the University of Austin in cooperation with the City of Austin from 2010, with slight modifications (Strover, Straubhaar, Chen, Gustafson, Schrubbe, & Popiel, 2014). Specifically, the PD component includes task-based dependence (PD_T) and social-centered dependence (PD_S), following Park et al. (2019).

¹¹ With this sample size, the survey data have an 8% margin of error rate at the 95% confidence level. Although professional standards suggest that an acceptable margin of error used by most survey researchers typically falls between 4% and 8% at the 95% confidence level, these findings constitute an exploratory effort, and consequently a larger error term can be tolerated.

Technological capabilities address four different factors: the operational skills (CP_OS), information navigation skills (CP_IS), social skills (CP_SS), and creative skills (CP_CS). These are adapted from Van Deursen, et al. (2017) with further adjustments to be suitable for the scope of this study. Table 8 below shows the items associated with each construct.

Constructs	Components	Instruments
Perception of tech-dependence (PD)	Task-based (PD_T)	1. The Internet is very important in my life.
		2. I keep up with events by using the Internet.
	Social-centered (PD_S)	3. I use the Internet for work-related tasks.
		4. I use the Internet for personal utilities such as paying bills, shopping or making purchases online.
Tech-competence (CP)	Operational skills (CP_OS)	5. I stay in touch with family & friends using social media.
		6. I use social media to share things that are important to me.
		7. I make friends and interact with other people using social media.
		8. Social media is important to meet my socialization needs.
		I know how to...
		1. Upload files and content (e.g. texts, pictures, music, videos, and web pages)
		2. Save or store files and content
		3. Bookmark a website
	Information navigation skills (CP_IS)	4. Adjust privacy settings
		5. Connect to a WIFI network
		6. Turn on my computer, logon and do basic tasks
		7. Install/download apps on a mobile device
		8. I find it hard to decide what the best keywords are to use for online searches.
		9. I find it hard to find a website I visited before.
		10. Sometimes I find it hard to verify information I have retrieved.
		11. All the different website layouts make working with the internet difficult for me.
		12. Sometimes I end up on websites without knowing how I got there.
	Social skills (CP_SS)	13. I know which information I should and shouldn't share online.
		14. I feel comfortable deciding who to follow online (e.g. on services like Twitter or Tumblr).
		15. I know how to change who I share content with (e.g., friends, friends of friends, or public).
		16. I am able to find relevant communities, networks, and social

	media that correspond to my interests and needs.
Creative skills (CP_CS)	17. I would feel confident writing and commenting online. 18. I know how to create something new from existing online images, music, or video. 19. I know how to make basic changes to the content that others have produced. 20. I know how to design a website. 21. I would feel confident putting videos, photos, or music content I have created online. 22. I know how to write computer code in any language.

Table 8: Measures of technological embeddedness.

As referenced in Chapter 2, prior research shows that four dimensions can gauge *concern for information privacy* (CFIP); these items work equally well with a US and an international sample (Bellman et al., 2004; Stewart & Segars, 2002; Smith et al., 1996). The four are concerns around information collection (CFIP_C), improper access (CFIP_A), errors (CFIP_E), and unauthorized secondary use (CFIP_U), as measured by the 13 items in Table 9. We are interested in examining how concern for privacy affects using the Steemit platform through the influence of other components such as technological competence, dependence, and Internet usage. With further adjustments of the CFIP items to fit the social media context of the current study, all items of the questionnaire were measured using a 5-point Likert-type agreement scale (1=strongly disagree to 5=strongly agree).

Dimensions	Instruments
Collection (CFIP_C)	1. It usually bothers me when social media sites ask me for personal information. 2. It bothers me to give personal information to so many people on the social media sites I am registered with. 3. I am concerned that companies are collecting too much personal information about me through the social media sites I am registered on.
Improper access (CFIP_A)	4. Social media sites should take more steps to make sure that unauthorized people cannot access personal information on their computers. 5. Databases that contain personal information should be stored in a highly secured location.

	6. Social media sites should delete a user's account for illegally accessing other users' personal information.
	7. Computer databases that contain personal information should be protected from unauthorized access—no matter how much it costs.
Errors (CFIP_E)	8. Companies should take more steps to make sure that personal information in their files is accurate. 9. Companies should have better procedures to correct errors in personal information. 10. Companies should devote more time and effort to verifying the accuracy of the personal information in their databases.
Unauthorized secondary use (CFIP_U)	11. Social media sites and companies should not use personal information for any purpose unless it has been authorized by the individuals who provide the information. 12. When people give personal information to a social media site or company for some reason, the company should never use the information for any other purpose. 13. Social media sites or companies should never share personal information with other companies unless it has been authorized by the individual who provided the information.

Table 9: CFIP scale (modified from Bellman et al., 2004).

Measuring User Performance

The analysis uses data from users' Steemit platform account. In order to assess a user's performance on the platform rooted in behavioral phenomena, the study focuses on four elements of users' activity record that are closely related to user performance (UP): reputation (up_rp), estimated account value (up_eav), number of followers (up_f), and number of posts generated by users on a monthly basis (up_mp). Reputation is, according to Steemit, "one way Steemit measures the amount of value a user has brought to the community." Every user's reputation score starts at 25 (this can go down to a negative scale) and operates within a log10 system, meaning that a score of 50 is about ten times higher than a score of 40. Basically, one's reputation score goes up when other users vote for her content, so a high reputation score can imply that one's contribution of high-quality content over a period of time. The estimated account value is, approximately, based on an average value of Steem and Tron crypto coin in US dollars held by a user account.

Data Analysis

We used the Partial Least Square-Structural Equation Modelling (PLS-SEM) using SmartPLS 3, commercial software that provides the most up-to-date and comprehensive software implementation of the PLS-SEM methodology. PLS-SEM is considered an appropriate SEM method for the study considering the sample characteristics and research design of this study because it not only handles all kinds of model complexity with constructs and indicators with small size samples, but also makes no distributional assumptions (Hair, Hult, Ringle, & Sarstedt, 2017). It also offers flexibility for exploratory research with real-world data (Hair, Risher, Sarstedt, & Ringle, 2018).

Evaluating the specified PLS-SEM result requires a two-step process: measurement model assessments and structural model assessments (Hair et al., 2018). The former is to ensure the validity and reliability of the construct measures and the latter is to evaluate the predictive power of the path model. For the measurement model, we checked the indicator loadings, composite reliability (CR) or Cronbach's alpha, as well as the average variance extracted (AVE) to assess a construct's convergent validity. To evaluate discriminant validity that calculates the extent to which each construct is distinct from each other in the path model, the study checked heterotrait-monotrait (HTMT) ratio values. As a next step, structural models are evaluated based on the following: the coefficient of determination (R^2), the statistical significance and relevance of the path coefficients, as well as the predictive relevance (Q^2). Taken together, the study applied the above criteria in order to evaluate the component measurement and structural models.

Pairing the online survey responses with secondary data that indicate the extent to which each survey participant became successful on the Steemit platform, the independent variables are technological embeddedness (tech-competence and tech-dependence); actual

use of the Internet and Steemit is treated as a mediating variable; user performance on the platform is a dependent variable. Demographic variables were used as exogenous, control variables. Figure 8 delineates the conceptual path model for analysis. Since the relationships between the subconstructs of technological competence and dependence, and privacy awareness are not yet sufficiently researched, the conceptual path only maps the relationship between upper-level constructs.

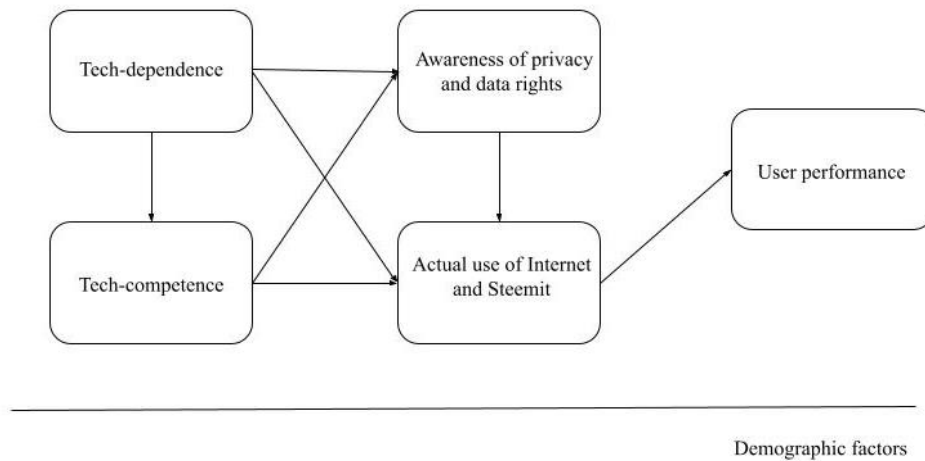


Figure 8: Conceptual path model.

The study first constructs the path model comprised of technological competence, dependence, actual use, and user performance. Drawing on this model as a base structure, we explore the position and association of the components regarding the awareness of privacy and data rights.

STUDY 3 RESULT: TECHNOLOGICAL CONDITION FOR USER PERFORMANCE

Descriptive Statistics

Valid responses were collected from 150 Steemit users across 29 different countries¹². Human development indicators in key dimensions such as health, education, and standard of living (UNDP, 2021), were examined for the sample using the [Human Development Index \(HDI\)](#) developed by the United Nations Development Programme (UNDP). We find that 63 subjects are from the countries below an average HDI of 0.71 for 2021, and 87 participants from countries above the average.

Table 3 displays demographic information of the participants. Men account for 75% of the responses collected while females constitute only 25%. Participating Steemit users were relatively young; the mean age of the participants was 32 and approximately 70% of the sample was aged less than 36 years. In terms of education, 58% of the users are highly educated, followed by about 40% of the participants who have earned less than a 4-year undergraduate degree. In regard to employment status, over 55% of the participating users were full-time, part-time, or self-employed workers; 33% were students and about 23% of the sample were unemployed. Most of the users perceived their social class as middle class (38%), followed by working class (28%), upper or upper-middle class (17%) and lower class (9%).

Item	Labels	Frequency	%
Countries (HDI_based)	Developing	63	58
	Developed	87	42
Gender	Male	111	75.5
	Female	36	24.5

¹² These countries include Venezuela (36 responses), Nigeria (19), Philippines (16), India (11), USA (11), Pakistan (10), Ghana (9), Bangladesh (5), Malaysia (3), Cameroon (3), Greece (3), Algeria (2), Canada (2), UK (2), Albania (1), Denmark (1), Ecuador (1), Germany (1), Indonesia (1), Italy (1), Kenya (1), Malta (1), Mexico (1), Namibia (1), Spain (1), Sri Lanka (1), Syria (1), Turkey (1), Vietnam (1), and unclassified (3).

Education	Less than high school	5	3.4
	High school graduate	30	20.1
	Tech certificate, 2-year college degree	28	18.8
	4-year undergraduate degree (BA or BS)	54	36.2
	Graduate or professional degree (MA, MD, JD, or Ph.D.)	32	21.5
Employment (multiple responses allowed)	Full-time employment	45	30
	Part-time employment	28	18.7
	Unemployed	27	18
	Student	49	32.7
	Self-employed	10	6.7
	Others (e.g., retired, pensioner or rentier; Housewife, -man, Disabled)	7	4.6
Perceived social class	Upper	7	4.7
	Upper-middle	18	12.1
	Middle	56	37.6
	Working	41	27.5
	Lower	14	9.4
	Prefer not to say	13	8.7
Age	Min=18, Max=77, <i>Mean</i> =32.2, <i>SD</i> =11.4		

Table 10: Demographic information.

In terms of general Internet and Steemit usage, the results showed that most participants are heavy Internet users who almost constantly access the Internet (69%), spending more than 40 hours (37%) or 21-40 hours (24%) per week on the Internet. To a lesser extent, 34% of the participants responded they constantly use Steemit and 39% several times a day: 28% spend 1-2 hours on Steemit a day, 23% more than four hours, 19% 2-3 hours, 19% less than an hour, and 10% 3-4 hours. These people use a variety of social media; on average, they were also on 4.6 other social media platforms. Of these, they only disclosed a subset of these accounts ($M = 1.7$) on Steemit, linked or associated with their Steemit posts. Overall, the mean scores of tech-competence and tech-dependence was

4.0 (SD=0.48) and 3.9 (SD=0.74), respectively, which is slightly lower than their level of concern for information privacy (M=4.34, SD=0.62).

In terms of user performance, an average reputation score was 54 with a SD of 11.34. For the sample population, the median number of followers was 74 (SD=1102.93), monthly-generated posts numbered 57 (SD=106.63), and the estimated account value in USD was \$87.46 (SD=1182.08). There was considerable variation among users' performance on the platform.

Measurement Model Assessment

Assessing the measurement model confirms the reliability and validity of the composite measures established for analyzing the path structure. The reliability of instruments can be first determined using the indicator loading value of each construct, which represents the sheer contribution of the indicator to define the latent construct. Hair et al. (2017) suggested that the acceptable cutoff value of equal or greater than 0.7 as reliable; researchers can also conditionally retain an item with a loading range between 0.40 and 0.70 when the item removal does not affect the value the composite reliability (≥ 0.7) or average variance extracted (AVE)¹³ (≥ 0.5), or the study is exploratory and the indicator contributes to content validity of the construct. The Cronbach's alpha and composite reliability equal to or greater than 0.7 are considered a desirable threshold for

¹³ AVE is the common measure to evaluate convergent validity on the construct level (Hair et al., 2017). Defined as the grand mean value of the squared indicator loadings of a construct, the threshold value of 0.5 or greater indicates that the construct explains more than half of the variance of its indicators.

the internal consistency reliability.¹⁴ As shown in Table 11, most scales are reliable and satisfy the criteria; some items were kept in the construct despite its loading score lower than 0.7 (bolded loadings). This is due to the nature of index that needs further testing and refinement for future study of the field and the potential of each item. Since each of these bolded loadings fall between 0.4 and 0.7, it is reasonable to assume that it still contributes to the content validity of each latent construct. Measures of the use of Steemit (SU) and Internet (IU) had lower Cronbach's alphas (below 0.7) but remained in the model following their acceptability in regard to the composite reliability and AVE.¹⁵

Constructs	Items	Loadings	Cronbach's alpha	Composite reliability	Average Variance Extracted
PD_T (Perceived tech-dependence: task-based)	pd_t_a	0.728	0.718	0.826	0.544
	pd_t_b	0.819			
	pd_t_c	0.736			
	pd_t_d	0.659			
PD_S (Perceived tech-dependence: social_centered)	pd_s_a	0.758	0.813	0.876	0.640
	pd_s_b	0.788			
	pd_s_c	0.821			
	pd_s_d	0.830			
CP_OS (Tech-competence: operational skills)	cp_os_a	0.840	0.882	0.908	0.587
	cp_os_b	0.820			
	cp_os_c	0.743			
	cp_os_d	0.752			
	cp_os_e	0.714			
	cp_os_f	0.760			
	cp_os_g	0.723			
CP_IS (Tech-competence: information)	cp_is_a	0.592	0.832	0.876	0.589
	cp_is_b	0.859			
	cp_is_c	0.851			

¹⁴ It is reasonable to report both criteria because Cronbach's alpha tends to underestimate the internal consistency reliability, whereas the composite reliability score results in the opposite (Hair et al., 2017).

¹⁵ According to Garson (2016), composite reliability can be a preferred alternative to Cronbach's alpha, which is too conservative, especially in PLS-based research. For all the constructs, the AVE values are greater than 0.5, thus establishing convergent validity of each construct.

navigation skills)	cp_is_d	0.804			
	cp_is_e	0.698			
CP_SS	cp_ss_a	0.556	0.702	0.813	0.526
(Tech-competence:	cp_ss_b	0.739			
social skills)	cp_ss_c	0.824			
	cp_ss_d	0.754			
CP_CS*	cp_cs_a	0.771	0.692	0.810	0.517
(Tech-competence:	cp_cs_b	0.726			
creative skills)	cp_cs_c	0.650			
	cp_cs_d	0.715			
SU	su_a	0.765	0.486	0.748	0.503
(Actual Steemit	su_b	0.769			
use)	su_sns	0.570			
IU	iu_a	0.755	0.519	0.757	0.512
(Actual Internet &	iu_b	0.749			
SNS use)	iu_sns	0.635			
UP	up_rp	0.728	0.771	0.848	0.583
(User performance)	up_eav	0.754			
	up_f	0.745			
	up_mp	0.826			
CFIP_C	cfip_c_a	0.713	0.635	0.792	0.565
(Collection)	cfip_c_b	0.898			
	cfip_c_c	0.617			
CFIP_A	cfip_a_a	0.796	0.829	0.887	0.662
(Improper access)	cfip_a_b	0.840			
	cfip_a_c	0.864			
	cfip_a_d	0.750			
CFIP_E	cfip_e_a	0.860	0.832	0.898	0.747
(Errors)	cfip_e_b	0.909			
	cfip_e_c	0.875			
CFIP_U	cfip_u_a	0.897	0.868	0.919	0.791
(Unauthorized	cfip_u_b	0.909			
secondary use)	cfip_u_c	0.861			

Table 11: Measurement model results.

* Each item can be interpreted corresponding to how the survey questions are arranged in the tables above. For the CP_CS construct, two indicators, cp_cs_e (web design skills) and cp_cs_f, (computer programming skills) are removed considering its low factor loading (0.407 and 0.098, respectively) and a very low AVE (0.372).

Note. With the study goal to further elaborate the role of privacy perception (CFIP_) in one's use of technology, the study included the four dimensions of privacy perceptions on by one in testing the base path model; this approach resulted in sequencing the four path models one by one, but it only had a marginal effect on other constructs (coefficient differences between 0.001 and 0.002).

In order to determine the discriminant validity of each construct, the study estimated HTMT values and the result ensured each construct is sufficiently distinct from other constructs (see Appendix 4 for detailed output).

Structural Model Assessment

The path model of this study predicts the outcome, e.g., the level of user performance on the platform, by examining one's perception toward digital technology in everyday lives (Park et al., 2019) and concerns for information privacy (Bellman et al., 2004) as constructs that are associated with digital skills and actual use.

As suggested by Hair et al. (2018), possible collinearity should be detected to make sure it does not bias the regression outputs. The variance inflation factor (VIF) should ideally be close to 3 or lower to avoid probable collinearity issues among the predictor constructs. The results showed that the VIF values of the predictor constructs of the model for this study met an appropriate criterion, ranging from 1.0 to 1.77. Collinearity is not a problem for the model.

To assess the structural model, key criteria include the level of R^2 values,¹⁶ the significance of path coefficients, and the predictive relevance Q^2 values (Hair, Sarstedt, & Ringle, 2019).¹⁷ As shown in Figure 1 below, the base path model, where privacy concern factors are not yet taken into account, has an R^2 value of 21% for social-centered dependence, 10.5% for competence of operational skills, 9.6% for competence of information-navigation, 43.5% for social competence, and 37.9% for creative competence,

¹⁶ R^2 values are a measure of the model's explanatory power that calculates the proportion of variance that is explained by each endogenous construct.

¹⁷ Unlike traditional covariance-based SEM, PLS-SEM does not have universal goodness-of-fit measures. With the focus of prediction, reproducing the relationship among the variables, it uses the listed coefficients to assess model fit ("model quality") (Garson, 2016).

6.4% for Steemit use, 14.3 for Internet use, and 15.3% for user performance. Table 12 reports the coefficient for each variable that can verify the significant of the path relationship.

Subject	Path	Path coefficient (β)	t statistics
Task-based tech-dependence → web 1.0 competence and actual use	Task-dependence → social-dependence	0.495	6.612**
	Task-dependence → operational skills	0.324	3.789**
	Task-dependence → Info-navigation skills	-0.196	2.579*
	Task-dependence → Internet use	0.266	2.426*
	Task-dependence → Steemit use	0.051	0.493
Social-centered tech-dependence → web 2.0 competence and actual use	Social-centered dependence → social competence	0.115	1.735
	Social-centered dependence → creative competence	0.195	2.591*
	Social-centered dependence → Internet use	-0.111	1.111
	Social-centered dependence → Steemit use	0.077	0.798
Web 1.0 competence → Web 2.0 competence	Operational skills → information-navigation skills	0.311	5.634**
	Operational skills → social competence	0.636	11.311**
	Operational skills → creative competence	0.373	4.523**
	Info-navigation skills → social competence	0.016	0.258
	social competence → creative competence	0.239	3.03**
Web 2.0 competence → actual use	social competence → Internet use	0.152	1.437
	social competence → Steemit use	-0.128	1.195
	creative competence → Internet use	0.135	1.345
	creative competence → Steemit use	0.197	2.273*
Internet use → Steemit use	Internet use → Steemit use	0.095	0.918
Actual use → user performance	Internet use → User performance	0.150	1.907
	Steemit use → User performance	0.342	3.626*

Table 12: Path results.

Note. Critical t-values. *1.96 ($p < 0.05$); **2.58 ($p < 0.01$).

Compared to the R^2 value that is often referred to as “in-sample” predictive power, the Q^2 values calculates the path model’s predictive accuracy based on a quasi-out-of-sample data: it reuses the sample by omitting every n th data point and re-estimates the model to predict the omitted “true” data with remaining data points (Hair et al., 2017). As

a technique combining out-of-sample prediction and in-sample explanatory power, a higher Q^2 indicates a higher predictive accuracy.

Research suggests that the Q^2 value larger than zero for a particular endogenous construct shows the path model's predictive relevance of the structural model for that construct. As a guideline, Q^2 values larger than 0, 0.25, and 0.50 consider the small, medium, and large predictive relevance of the path model. With the blindfolding procedure of SmartPLS 3 software, the analysis demonstrated that the explored path model has predictive powers for all the constructs in the model: the Q^2 values were 0.13 for PD_S, 0.06 for CP_OS, 0.05 for CP_IS, 0.21 for CP_SS, and 0.18 for CP_CS, 0.01 for SU, 0.05 for IU, and 0.07 for UP.

Furthermore, the path coefficients of Figure 9 and Table 12 exhibit all the paths that are significant, thereby fully or partially supporting H3.1.1, H3.1.2, H3.1.3 and H3.1.4. Bootstrapping analysis with 5,000 bootstrap samples is applied to assess the significance of path coefficients. Looking at the relationships between the constructs, the task-based technological dependence (PD_T) and social-centered dependence (PD_S) have positively affected some aspects of digital competence (CP) and actual use (IU and SU) except for a negative influence of task-based dependence on information navigation competence ($\beta = -0.196$; $p < 0.05$).

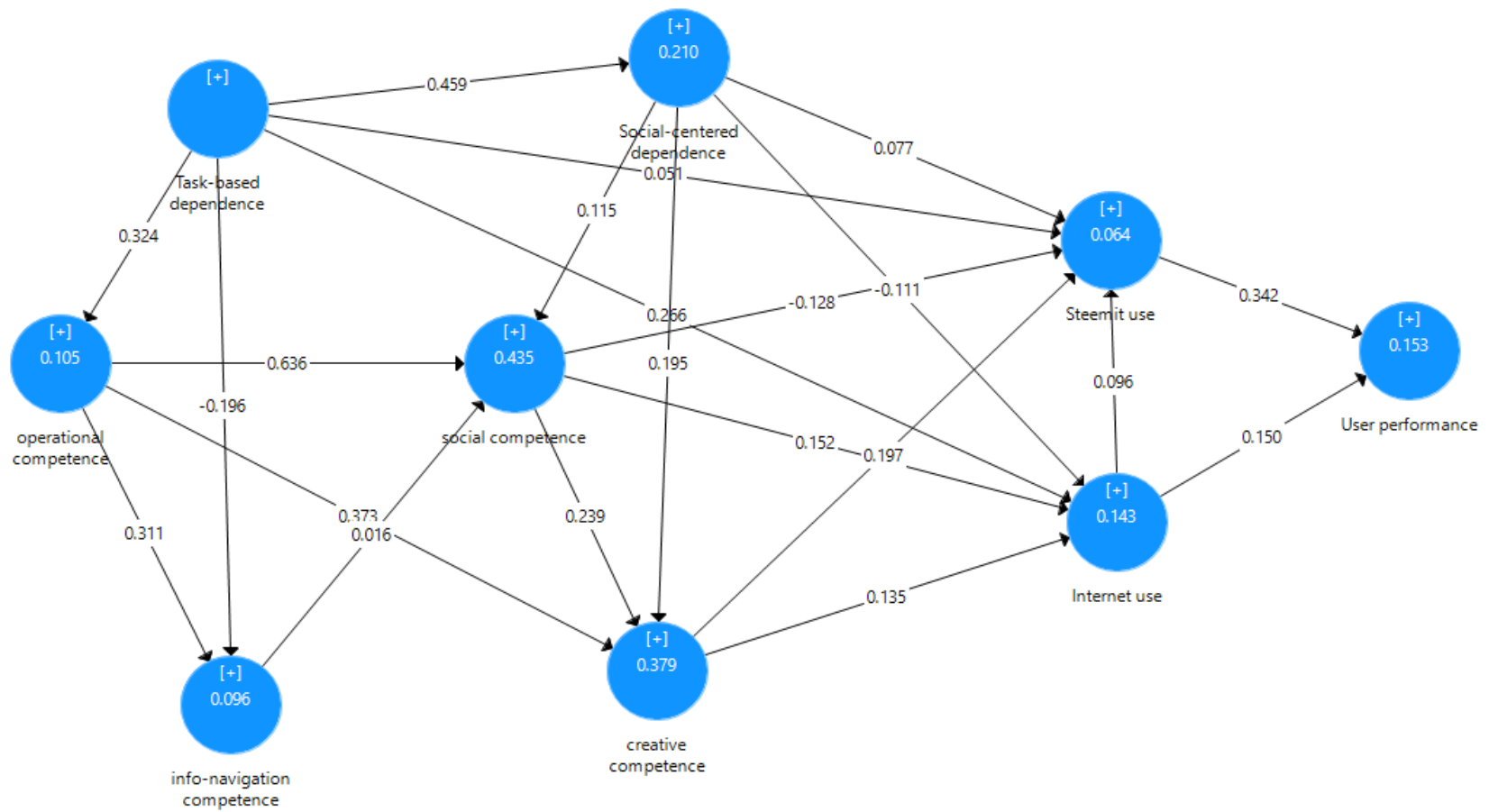


Figure 9: Base path model result.

Given the size of coefficient, the construct of tech dependence is more strongly associated with digital competence than actual use. In terms of whether the development of competence leads to the actual use of the Internet and Steemit, only the path between creative competence and Steemit use, i.e., time spent on Steemit and the number of other social network accounts that are linked with Steemit, was significant ($\beta = 0.197$; $p < 0.05$). Consequently, the use of Steemit positively and significantly predicted how a user performed on the Steemit platform ($\beta = 0.342$; $p < 0.05$).

Demographic variables including age, gender, employment, education, perceived social class, and country based on the HDI classification are used as control factors (Table 13). While there are many developing recommendations to handle control variables in PLS-SEM research, this study selected the single-item approach—examining the effect of control variables one by one on the dependent variable, considering the limits of sample size, difficulty of interpretation, and reliability and validity challenges of multi-item approach (De Battisti & Siletti, 2019).¹⁸

Subject	Path	Path coefficient (β)	<i>t</i> statistics	Effect
Control variables	Age → UP	0.322	4.639**	Significant
	Female → UP	-0.005	0.079	Not-sig
	Unemployed → UP	-0.288	5.213**	Significant
	Less-educated → UP	-0.072	1.094	Not-sig
	Lower class → UP	-0.020	0.287	Not-sig
	Lower HDI countries → UP	-0.206	3.012**	Significant

Table 13: Effect of control variables.

Note. For dichotomous variables, reference groups are those not stated in the path above, e.g., males as the reference group for females.

Note. Critical *t*-values. *1.96 ($p < 0.05$); **2.58 ($p < 0.01$).

¹⁸ De Battisti and Siletti further suggested that the separate or simultaneous application of single-item variables means the same differences as in regression models when considering a unique multiple regression with all controls together or multiple regressions differing only by the controls.

The study re-estimated the model by including the four dimensions of privacy concerns consisting of the concerns for information collection, information errors, improper access, and unauthorized secondary use. Each of these privacy variables was separately included in the base model and tested to empirically account for the multidimensionality of privacy concerns and risks in relation to the perception of tech dependence, tech competence, and actual use that are associated with the user performance on the platform. For the path model, it is assumed that privacy is influenced by both the tech-dependence and web 2.0 digital competence constructs to invoke subsequent online behaviors. Awareness and concerns of information privacy can be seen as being formed through prior conditions of cognitive understanding and emotional responses to the current digital environment, including the risks and challenges of data collection and sharing. (Li, Luo, Zhang, & Xu, 2017). As the current study's information privacy constructs are primarily rooted in social media contexts, these factors are more likely to be closely associated with web 2.0 based resources and capabilities. The same procedure for path analysis was performed on each of the four different models. No collinearity was detected in all the path models. Additionally, this study compared R^2 and Q^2 values of all indicators of the four privacy models, and it demonstrates that the inclusion of each privacy indicator does not significantly influence the prediction accuracy or the percentage of variation of other indicators in each model.

In terms of path analysis, Table 14 reports the significant path coefficients for the four path models involving respective privacy dimensions. Most of the privacy constructs are determined in large part by the tech-competence variables. Regarding Model 1, the path between the privacy construct of information collection and the creative skill of technological competence is significant ($\beta = 0.239$; $p < 0.05$), indicating that the creative

competence increases the concern for providing personal information to be collected on social media sites. Tech-competence focused on social skills positively influenced the concerns for unauthorized access to personal information ($\beta = 0.496$; $p < 0.01$) that also led to an increase of actual use of the Internet and social media sites ($\beta = 0.207$; $p < 0.05$). Here, the indirect effect of the privacy concern for unauthorized access to address the relationship between tech-competence on social skills and Internet use was marginally significant ($t = 1.958$; $p = 0.05$). A concern for errors in personal information was significantly predicted by technological dependence on social media ($\beta = 0.259$; $p < 0.01$) and competence of digital social skills ($\beta = 0.284$; $p < 0.01$). Model 4 demonstrates the path between the concern for unauthorized secondary use of personal information and the competence of social ($\beta = 0.344$; $p < 0.01$) and creative skills ($\beta = 0.223$; $p < 0.05$), representing that both web 2.0 competence constructs significantly affect this type of privacy concern.

The result demonstrates that the awareness of privacy and data rights is positively affected by technological competence (H3.2.1) while negatively predicted by tech-dependence (H3.2.2). However, the significant path between the effect of the concern for inappropriate access to personal information that predicts more use of the Internet contradict the hypothesis (H3.2.3). Also, this concern about inappropriate access positively mediated the effect of social competence on Internet use, giving results contrary to our hypothesis (H3.2.4).

Subject	Path	Path coefficient (β)	<i>t</i> statistics
Model 1:	Social-centered dependence → concern for data collection	-0.170	1.929
CFIP_C	Social competence → concern for data collection	0.131	1.229
	Creative competence → concern for data collection	0.239	2.412*

	Concern for data collection→ Internet use	-0.025	0.249
	Concern for data collection→ Steemit use	-0.051	0.549
Model 2:	Social-centered dependence → inappropriate access concern	-0.008	0.123
CFIP_A	Social competence → inappropriate access concern	0.496	5.008**
	Creative competence → inappropriate access concern	0.159	1.728
	Inappropriate access concern → Internet use	0.207	2.386*
	Inappropriate access concern → Steemit use	-0.152	1.327
Model 3:	Social-centered dependence → concern for errors in information	0.259	2.991**
CFIP_E	Social competence → concern for errors in information	0.284	2.883**
	Creative competence → concern for errors in information	0.111	1.009
	Concern for errors in information→ Internet use	-0.139	1.651
	Concern for errors in information→ Steemit use	0.143	1.331
Model 4:	Social-centered dependence → concern for unauthorized	-0.039	0.636
CFIP_U	secondary use		
	Social competence → concern for unauthorized secondary use	0.344	3.549**
	Creative competence → concern for unauthorized secondary use	0.223	2.263*
	Concern for unauthorized secondary use → Internet use	-0.017	1.428
	Concern for unauthorized secondary use → Steemit use	0.174	0.159

Table 14: Path results for privacy constructs

Note. Critical t-values. *1.96 ($p < 0.05$); **2.58 ($p < 0.01$).

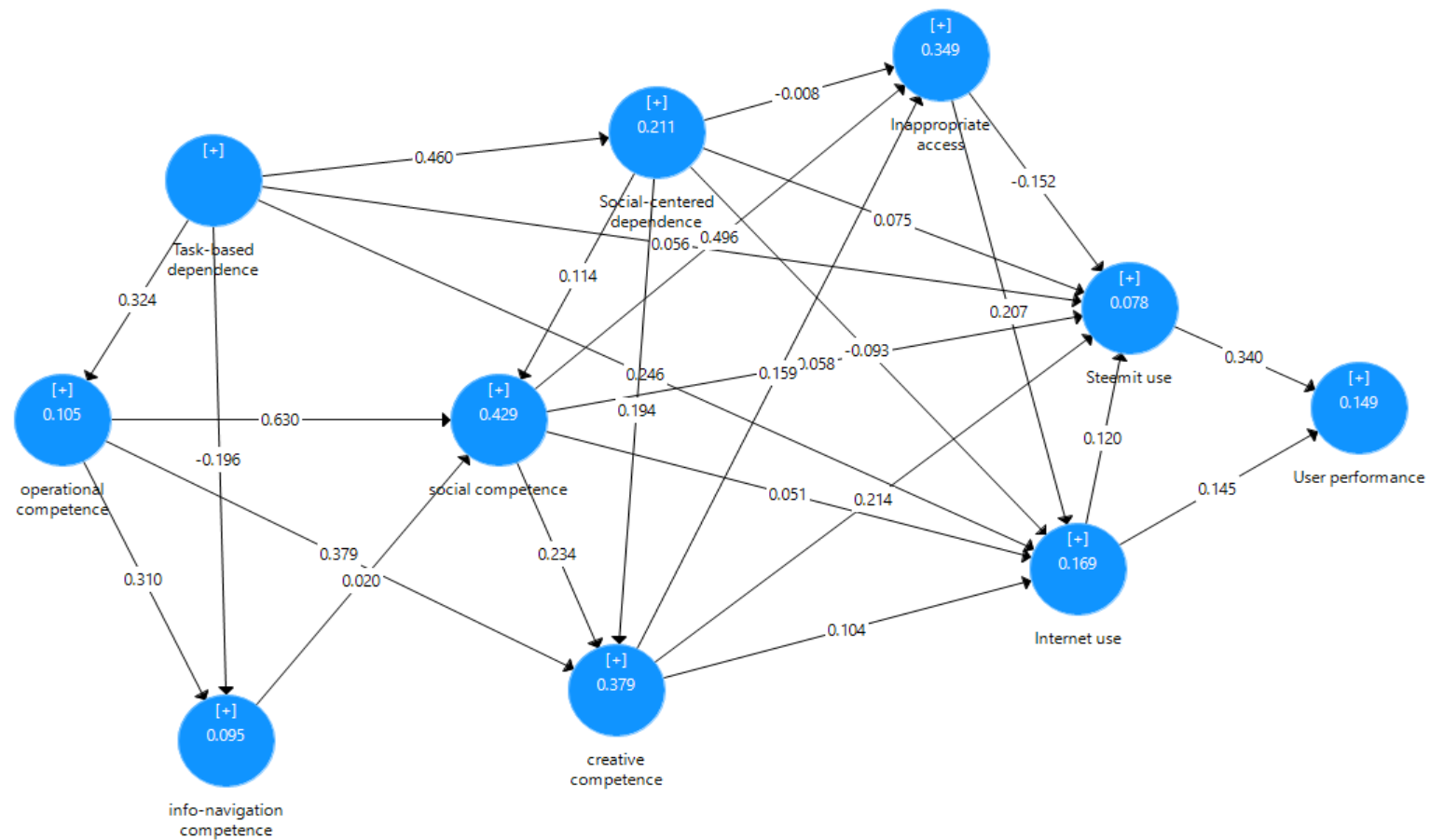


Figure 10: Path model expanded with privacy construct of improper or unauthorized access.

STUDY 3: CHAPTER DISCUSSION AND CONCLUSION

The findings of this study suggest that the user performance on the social media platform Steemit is positively and significantly predicted by creative tech-competence, in turn associated with other competence constructs including operational skills and social skills. Overall, the results based on international user data not only validates the potential of creative competence as a key construct for users' successful performance on the platform in regard to their contribution of UGC, but reinforces the broad structure of prior literature that underscores the multiple dimensions of digital competence figuring in the sequence of developmental progressions of skills, use, and outcomes (van Deursen et al., 2017). These results show how the self-reports of one's dependence on digital communication technologies for everyday tasks and social interactions covary with web 1.0- and web 2.0-based technological capabilities as well as with the frequency and intensity of the use of the Internet and social media.

The construct of privacy in this study implies a deliberate, informational self-determination or control over personal data that extends beyond general concern for privacy; by influencing actual platform and Internet use behavior, privacy behaviors do imply the possibility of indirectly impacting individuals' performance on social media platform. Particularly, the interplay between privacy and digital capabilities highlights that this notion of privacy in the current digital environment is likely to be related to one's digital skills and knowledge dimensions. Specifically, while concern for general data collection and data accuracy is primarily predicted by digital social skills, variables regarding concerns for unauthorized access/ secondary use of personal information are closely associated with creative abilities. One explanation is that the concern for unauthorized access and use can be greater for those who regularly generate and post

content online; they want to maintain greater control over their production, and the scope of these “productions” increasingly grows, including content containing the information about oneself as well as tracking data such as location, IP address, or browsing activities. This, in turn, suggests a context where mature users of digital technologies who have developed a sense of digital privacy can further influence the way they facilitate, adjust, and implement capabilities in digital creative practices.

In addition, the study found a significant effect of some demographic factors on the model of user performance. The results indicate that users from countries with higher reported HDI scores were more likely to be successful in the Steemit platform than users of lower HDI countries. Likewise, users who are unemployed were less likely to have a high ranking in the metrics of user performance than the employed. This illustrates the vulnerability of certain populations that can be amplified by country-specific socioeconomic conditions and work status.

Despite the importance of the above findings, considering the relatively small sample size of 150 participants, these results should be interpreted with caution. In addition, the current path model requires a more comprehensive measure of time. For instance, user longevity – when one started using the platform - may provide another possible explanation for the difference in rewards for UGC activity between novice and experienced users. User longevity is not captured in the data.

Chapter 7: Discussion and Conclusion

REFRAMING DATA OWNERSHIP AND CONTROL IN REVOLUTIONARY DIGITAL SPACE

Our rights to personal data have long been precarious within the system of modern digital businesses that collects, uses, and sells personal data. Using a blockchain-based reward social media platform, the present study questioned and examined the potential of this revolutionary system to empower individual data subjects to exercise ownership and control over their own data. To reiterate the core research questions, the study first challenged the contemporary discourse around personal data rights for individual data subjects to exercise ownership and control over their own data. It showed how the reward platform Steemit can be an alternative to complement the gaps in the existing policy and regulatory arenas. The next questions focus on users' actual data activities and practices on this new type of platform: what types of personal data were voluntarily shared, how were they shared, and how was sharing personal data actually rewarded? Users' self-generated content became the basis for empirical investigations. Finally, the study explored the sociotechnical conditions that may maximize user benefits on this platform.

The current regulatory discourses and industrial practices disregard or undermine the role of individuals in the battleground for data rights (Chapter 2). Responding to public concerns about their data rights, current efforts worldwide have focused on strengthening data protection legislation as well as overseeing commercial companies. Their efforts are rooted in two perspectives: (1) protecting data privacy as constitutional rights that are inalienable and untradeable for a natural person (Wilks & Christie, 2013), and (2) using the public law approach that views shared data as public goods essential for "human flourishing" (Benkler, 2006; Cohen, 2012) and social justice (Johnson, 2014, 2016; Heeks & Renken, 2016). The study argues that such viewpoints could treat end users as mere

“consumers” or “victims” or lead to individuals’ interests being dispersed and disorganized. Additionally, we show how the discussion of the privacy “tradeoff” common in mainstream explanations of why people give up personal information when they use popular applications and platforms itself has a fairness flaw, given information asymmetries of data markets and dilemmas in current contractual practices and technical designs.

Revisiting the notion of privacy as property complements the limits of current conditions. To motivate individuals to stand up and act for themselves for their own rights, personal data should be more than about protecting privacy; instead, having economic or an instrumental value, “it is valued as a means for achieving certain other ends that are valuable” (Solove, 2002). This study views the instrumental values of personal information as powerful drivers for one’s desire of ownership and control. While this personal property framework has been challenged for its feasibility in existing markets, our study’s conception of a new type of decentralized reward platform enables individual data subjects to actually use, sell, and value information about themselves within the system.¹⁹ In doing so, we also provide the rationale for valuing personal data that is contributed in the user-generated content context. Our findings affirm that there is intrinsic value in one’s self-representation and expression in digital space and it can indeed be linked with the legal property base of the database under the rationale of adding value to data based on substantial individual investment, creativity, and originality.

¹⁹ This vision is rooted in Web 3.0 environments, where the World Wide Web is built on blockchains. Compare to web 2.0, in which content hosting companies firmly hold control and revenue from UGCs, blockchains enable a mass decentralization of publications. This means that individual users can directly control and monetize what they publish on the web (Palmer, 2021; The Investopedia Team, 2020).

Value Space for the “Qualified Self”

By analyzing content in a primary digital space on Steemit (Chapter 4), I find that the information receiving considerable investment and value in the platform’s currency was comprised of how people describe their personal traits, values, and beliefs or how they narrate original stories of personal experiences and events. The latter are more closely related to the “qualified inner self” than the standardized consumer profiles and behavioral and demographic data used in data-driven marketing. Such information practices were prevalent particularly during the boom of the cryptocurrency markets in which people had huge aspirations toward the market and presumably were more motivated. During the contrasting market phases of boom and bust, the composition of information space appeared differently. Steemit users formed a more diverse personal information environment with multifaceted self-representation strategies, which achieved greater values during the time when the cryptocurrency price was at an all-time high, compared to the late period when the market fell sharply.

In addition, the findings demonstrate the ceaseless continuation and expansion of “qualified self” practices on the revolutionary digital space. They echo some of Humphries’ observations regarding the drive for self-presentation, in which she affirms that people have sought to describe themselves in multiple ways, including in pocket diaries and photo albums from past centuries (Humphreys, 2018). My current study not only underscores that this is still a widespread practice, but also that it has real and considerable value in the collective eye. This observation suggests the potential of new blockchain-powered reward social networks as an alternative mode of platform governance that values mundane user contributions to the platform.

Importantly, even though our conceptual approach towards analyzing the value of privacy and personal information is based on previous research framing privacy as property (arguing that personal information may be constituted by economic value or capital), the results suggest the multidimensional nature of personal information as it created value on Steemit. In broad scholarship addressing forms of capital, personal values, beliefs, philosophies, taste, demographic and cultural environments, expertise, and virtual and real-world relationships, and relationships to other users and community memberships have been identified as main sources of value-creating and exchanging activities; Bourdieu's 'Forms of Capital' (1986), for example, presents capital in three fundamental guises of economic capital, cultural capital, and social capital. Mobilizing all three is consistent with the view of value in for-profit digital networked media such as YouTube, in that it has greatly increased the potential of value realization for a variety of end-users "by giving them more freedom to determine their understanding of value itself" (Grünewald & Haupt, 2014, p.3), allowing them to create social networks and to realize cultural capital as well as economic capital. Decentralized networked media such as that offered by blockchain-based installations, also promises to *give users even more freedom to determine the value itself* by allowing users' votes, i.e., collective user decision, to play a vital role in deciding how much to pay for each individual's contribution. Compared to conventional commercial social media platforms, we can expect that this realization of value in every aspect of self is more likely to empower users to exercise stronger ownership and control over whatever they make or contribute on the platform, regardless of whether or not it is sensitive personal information. The varied market conditions examined here arguably provide a look at how diverse information activities and valuations the market can make both when it is low and high.

Agency in Action

When compensation is assured, can it lead to excessive or other undesirable conduct in collective sharing of personal information? Based on the findings here (Chapter 5), a majority of user posts reflect practices that implemented and practiced privacy-aware identity management strategies: more than half of user-generated posts did not contain any personally identifiable information. Among the posts including personal information, about half of those posts-maintained boundaries between personal and professional representation. Those posts containing visualized self-presentation in a form of selfie or v-log tended to reduce the use of other multimedia components to express identities. In terms of valuation, however, posts with more self-disclosure and creativity were likely to achieve greater value: posts earned higher rewards when they combine personal and professional identities, when a significant part of them are original to the author, or when artistic creativity or other competencies and skills are exhibited. In addition, visual personal information tends to be more rewarded.

These practices have implications for the possibility of deliberative public engagement in personal information sharing and identity expressions within the new digital spheres such as Steemit. In this study, users of the decentralized reward platform appear to be aware of the risks in investing personal information in their content, and simultaneously acknowledge the efforts that other users put in for further self-disclosure or displaying special talents or skills (by upvoting them, for example). Presumably, depending on different motivations and goals, user-generated posts might have expressed self-identities using different facets of personal information and negotiating privacy boundaries in varied ways. This context emphasizes and expands the presence and multiplicity of individuals' agency over their own data (Kennedy, Poell, & Van Dijck, 2015) by engaging user

practices of identity management in the context of content creation as natural persons, content creators, and data generators. The analysis shows individual investment, creativity, and originality in peoples' posts, and suggests that these types of forums may indeed be acceptable (or even preferable) for people who wish to express themselves and enjoy fruit of their labor (Solove, 2002, 2012). The analysis illustrates that granting or strengthening property rights (or applying the law of value) to individual units of data production can in fact enliven a social media site and allow the realization of value just as has been shown in the case of databases. Lastly, privacy-aware identity practices observed on Steemit could be the first step in proving the promise of this new system that ensures user benefits in balancing rewards and privacy.

Creative Competence as a Key Driver of Data Competence

The third analysis explores the role of creative competence in creating value on these platforms. We hypothesized that creative competence can be a key predictor of users' successful performance on the decentralized reward platform, and this inquiry is important in order to better understand the circumstances in which people might be able to realize value under their own direction and control. Our path model confirms the sequential progression of technological competence from basic operational skills to social competence and to creative competence (Van Deursen et al., 2017). The study demonstrates the positive impact of technological dependence on creative competence, suggesting the indirect effect of the dependence construct on user performance. While the results indicate that being 'embedded' in the technology positively influences generative digital use and related outcomes, the information literacy component works differently: it does not influence advanced Internet competence while being negatively affected by tech-

dependence. The findings reveal that these peoples' awareness of privacy and data rights was driven by competence rather than dependence, and although insignificant in the current study, this awareness generally appeared to be a negative factor in actual use. Under these conditions, people from developing countries or who may be unemployed are less likely to succeed than their counterparts on this platform.

At the outset of this study, we highlighted the ownership rights to data as the ability to sell or derive benefit from data in an effort to find the place of individual users within the data economy. The results show how creative technological competence can fulfill this specific demand of data competence in which people not only learn to realize real-world value through data but also to increase their critical awareness of companies collecting, using, and sharing user data. It is noteworthy that this critical awareness of data rights and privacy may operate independently of traditional media and information literacy competencies (information-navigation skills). While studies on competencies or literacies of personal data are still in their nascent stage, the current finding provides an empirical link to connect technological competence with some of the five conceptual domains of personal data literacies underpinned by Pangrazio and Selwyn (2019), including data identification, data understandings, data reflexivity, data uses, and data tactics. In this context, the present findings signal another area of global digital inequality between developed and developing countries in terms of the benefit that users can derive from digital spaces. Beyond the first access and use divides, these results show how a third level of digital divide – disparities in actual returns – (van Deursen & Helsper, 2015) in the new technology field is still widening, reproducing existing inequities on a global scale.

Limitations and Suggestions

While the limitations of each set of studies have already been discussed in each chapter, this section provides some points where the overall design of this present study on the social impact of the new technology architecture of social media platforms as a whole could be further improved in several ways.

First, considering the new technical platform architecture that situated this study in the unique context, there is not yet sufficient evidence to differentiate user behaviors and practices on Steemit from what they do in other social media environments. Cross- and multi-platform level analysis would be needed to fully understand and estimate the potential of Steemit and other decentralized reward platforms. Second, a more consistent set of research data for each study in terms of time frame, the type of corpora, and survey participants, or a panel or longitudinal design, might have enhanced depth of analysis and contextual diversity of the same phenomenon of users' personal information practices; this study's cross-sectional approach using different corpora or time frame for each study limits an in-depth observation of users and how their data practices might change over time, presumably in response to benefits they receive. Finally, it should be noted that Steemit governance has undergone some changes in its leadership, interface, and functional features: although the underlying incentive mechanism is the same, the current results need to be approached with some caution, considering the possible impact of these changes.

As well, there are various types of decentralized applications running on different types of blockchain networks (e.g., public versus private, centralized versus decentralized, O'Leary, 2017) that can result in different use cases and different forms of governance. As Steemit is not the only use case that can be envisioned from blockchain-based social media

platforms, the study of future blockchain networks and emerging/ developing applications deserves more scholarly attention in the field of media and communication research.

IMPLICATIONS OF THE STUDY

Social Media User Studies: Identity, Privacy, and Value

While the existing media studies have extensively covered the subject of identity and privacy within users' information behaviors and the societal and cultural value they bring, few attempts have been made so far to relate it to tangible economic value that individual end users can enjoy. The findings of the current study provide empirical evidence to substantiate the discussion over the value of identity and private information in social media space, supplementing previous research on user participation and their information behaviors on social media by adding the context of *voluntary* self-disclosure and *democratic* value realization. These practices are inspired by decentralized incentivization mechanisms of the blockchain-powered social media platform. Extending the theorization of Bechmann and Lomborg (2013) on value creation via user participation, the current finding not only exemplifies the collapse of a classic producer/audience binary, but also effectively incorporates dual perspectives on social media users' value creation, in line with ideas of sense-making (creative explorations of the self, from user-centric perspective) and of straightforward economic gain (from the industry perspective).

Moreover, what can be further highlighted from our findings is that economic value creation can be the result of creative self-explorations. The results capture the context in which creative self-display, storytelling, and other forms of qualitative engagements in self-presentation online, as mirrored in Humphreys' "qualified self" (2018), can have a significant role in realizing value. Contrary to the heated attention paid to the subject of the

“quantified self” under the increased datafication model, the perspective of the “qualified self” has been treated as that of old times, somehow taken for granted. Research on self-expression on social media has concerned emphasized its narcissism, and its often incongruent personas between digital self and offline self (Cheng, 2004; Grieve, March, & Watkinson, 2020). However, recent studies have revitalized the cultural meaning of identity practices in the age of social media (Humphreys, 2018; Belliger & Krieger, 2016). The present study joins those efforts in affirming the desirability of broader versions of self-representation and how they are valuable to users. The results here report many “qualified” representations of ourselves in media traces, and document their reward in this particular platform (Humphreys, 2018).

Finally, this study could further contribute to the burgeoning literature of (qualified) social media identities in its discussion of power dynamics in surveillance capitalism (Zuboff, 2019) through the valuation of self in our media traces in distributed and decentralized architecture of blockchain-based social media platforms. By introducing the subject of this new system of communications and networks into the field of communication and media studies, the study offers a novel perspective to evaluate identity work which could further address concerns about the datafication and quantification of individual users by commercial social media platforms in the spirit of market capitalism. The observed information behaviors of users on this system may also offer new insight into social, cultural, and economic ramifications of identity work.

Creative Competence and Digital Divide

The present study argues the role of creative competence in regard to successful value-making performances on blockchain-based social media platforms. While the

general definition of “creative competence,” which refers to the ability to create original or remixed content online, was mentioned in the earlier section, the in-depth meaning of creative competence of this study should be articulated further. That, in part, is the goal of the third study.

The creative competence factor specified in this study’s path model (Chapter 4) constitutes four different dimensions: writing/commenting online, making basic changes to content produced by others, making substantial changes to content produced by other, and uploading/creating something new such as videos, photos, music content, etc. These dimensions had been found important in other studies, and we hypothesized they may affect peoples’ abilities to engage the STEEMIT platform. However, when testing the model as a whole, this study had to remove the elements of web design and programming skills, which were considered creative competencies in previous studies (Choi et al., 2020), due to inconsistency, i.e., low reliability of the scale. Consequently, the resulting elements of creative competence in this study entail a broad variety of content creation or remix activities which do not require high sophistication or technical specialization.

Apart from the possibility of the need for more instruments to assess creative competence, one implication of our findings prompts greater scrutiny of the Steemit platform’s affordances. Steemit is similar to Reddit (detailed in Chapter 3) in that it allows users to easily generate, edit, rate, and/or link to content or to others. The level of technological competence to remain in good standing in the Steemit environment (or similar decentralized application (dApp) environments) simply is not that high. As well, the creative expressions shared by people in posts did not have to be strictly original. It could be original or remixed, a simple photograph or a simple statement; the value

recognized by other community members appeared similar, as long as the content exhibited some creativity.

Having this in mind, it can be argued that, at least in the present study, the possible gaps in creative competence for explaining peoples' privacy management behaviors may emerge more from the *willingness* to create or contribute *anything* in order to engage the platform than from the difference of sheer technical knowledge or skills. This point underscores the potential complexity of the creative competence construct as to what actually constitutes creative competence in different digital contexts, whether there are subtypes of creative competence, and if so, how they develop or apply in general or in specific to function well in digital settings.

While an ample literature studies the phenomenon of digital competency in the field of digital and media literacy and digital divide studies, the relationship of each dimension of digital competence, such as basic skills, information skills, social skills, and creative skills, or the sub-clusters or types of respective dimensions, have been under-researched to date. On top of a few studies demonstrating sequential progress (van Deursen, et al., 2017) or consecutive progress (Choi et al., 2020) of digital skills, the current specification of creative competence may imply broadening the breadth and depth of digital divide research. It may be overly reliant on developmental paths that are out of step with what platforms demand (van Deursen & Helsper, 2015).

Methodological Implications

It is only recently that computational content analysis approaches such as LDA topic modeling to analyze large amounts of text documents have gained interest within media and communication research communities. Although this approach is increasingly

being employed with its capacity and efficiency for performing large-scale data analysis, researchers have recommended careful consideration of when to apply it most appropriately, indicating the conceptual limitations and the lack of initial application of domain knowledge by human experts (Puschmann & Scheffler, 2016). Heeding this advice, this current study applied a mixed-method approach using both computational content analysis and traditional content analysis to examine self-generated content with macro-and micro-level perspectives, bringing more room for contextual interpretation while retaining the benefits of large-scale analysis.

The LDA method provided data-driven characterization of documents, an approach which is less susceptible to human bias. The proportional view of topic shares within documents helped to assess topic diversity that may not have been effectively assumed by human judgment. Moreover, with the ability to aggregate documents on metadata levels such as time, language, and payout values, the study was able to examine the association of topics with received rewards and times of market boom and bust. Altogether, the advantages of the LDA approach enabled productive comparative analysis of corpora, which were aggregated into the Mania and Blow-off phases.

To address the problem of granularity that occurs frequently in LDA-based (and other computational topic modeling techniques) research, a traditional content analysis approach combining qualitative and quantitative components was applied as well. By establishing a theory-driven conceptual framework, this study discovered more diversified identity management and privacy-aware information strategies that may have gone unnoticed in the computational topic modeling approaches. Traditional content analysis also allowed this study not only to answer the value question of personal information with

depth from the detailed distribution of value assigned to each user post by other community members, but also to include visual elements of self-generated posts as a unit of analysis.

While many communication and media studies use a single-method approach – either computational or traditional content analysis – to analyze topics from a set of documents, this research employed diverse methods from a multi-level perspective to provide a depth picture to understand users’ personal information practices reflected in the form of UGCs and realized economic value. In doing so, the analyses of the present study were based on actual self-created user posts and related metadata, which overcomes possible biases of self-reporting and inherent information asymmetry reflected in how commodity market services (conventional social media sites, for example) often “produce” user data. Combined with a user survey design implemented to investigate sociotechnical conditions under which users undertake value-creating activities, the utilization of the mixed method design helped this study to provide a comprehensive view to understand not only actual personal information that is invested, shared, and valued, but also external conditions that lead one to be successful on incentivized blockchain-based social media.

Practical Implications

The vision of a redistributive data economy conceived and implemented through decentralized blockchain-powered social media environments in this study challenges the monopolistic tendencies of the data economy dominated by giant tech platforms. With the help of technology, end-users have been able to receive economic rewards for their creative self-expressive engagements through a wide variety of user-generated content. The blockchain environment suggests a contrast to conventional platforms in terms of how data rights and privacy might be handled.

Let's briefly review several existing solutions and discourses about data rights and data wealth distribution under scrutiny currently. As detailed in Chapter 2, critics are posing ideas and frameworks for considering Information fiduciaries, data trust, data taxes, and dividends within the domain of regulation, policy, and law. While these are loosely linked to the mechanism of the market, they basically posit the need for setting rules for the market. From the standpoint of tech-based solutions, there are three representative examples. First are the recent ideas of "solid open-source protocol" from Sir. Tim Berners-Lee with a group of developers (2016). Their proposal aims to separate online applications from user data; users can store their data securely and separately in one place they have control over and then decide where to contribute their data. Another type of existing technical alternative is an alternative web browser called Brave that blocks ads and trackers in our web searching and retrieving; when users choose to spend their time viewing ads and content, it pays "Basic Attention Tokens (BATs)" to users. Last is a growing ecosystem of decentralized applications (dApps) running under a reward mechanism similar to Steemit; as mentioned, it aims to fairly reward each person's contribution to the platform. Considering that there is no one-size-fits-all solution to address the current personal data controversies, each of these proposed solutions could offer a meaningful complement to strengthen our data rights.

While the current study results will only be a starting point to examine the role of the dApps ecosystem as another technical solution for our data rights, the operating logic of incentivization, as well as its technological affordance, differentiate what it could induce end-users to exercise/strengthen their ownership rights to data. First, the affordances of Steemit have maximized the flexibility of self-explorations across a wide variety of self-generated content. It is primarily user-driven, user-centric creations that contribute

to/realize the value of the platform by giving people the freedom to choose how they want to explore, express and communicate about themselves; in this regard, the control Steemit offers to users not only includes protecting or securing data but also allows them to invest, sell, or profit from data at their own will. Second, the low identity affordance and customizability of the content can be combined to become a forum where one can freely exhibit her creativity or express ideas or opinions with fewer features and lower anxiety associated with being judged or criticized. Compared to other technical solutions centered on data-oriented managements and functions only, the design of Steemit advances an understanding of the user as an expressive and creative actor. Furthermore, the incentivization mechanism can be a driver of “gamification” of activities on Steemit. While the current study posited that the higher the reward, the higher the motivation to create/contribute, there is also the argument that lower rewards do not necessarily discourage motivation, but may be an extra boost to the activity that users are already enjoying (Thelwall, 2018). Like game playing, even without monetary reward, users may enjoy the challenge of increasing the number followers associated with enhancing their chances of getting more upvotes.

Situating personal data rights within the context of flexible self-expression, creativity, and gamification, Steemit initially envisioned an environment where users could exercise personal data rights as well as engage in activities conveying positive affect and value. Opportunities for enjoyment or fun were structured to bypass tedious or annoying laborious content practices. Despite recent challenges in Steemit governance, its fundamental vision could still serve as a model for how future platforms can ensure user control over data.

CONCLUSION

Our rights to personal data exercised on Steemit have been shown to change how we can control our data in the context of the Web 3.0 environment. The incentive mechanism of blockchain-based social media platform Steemit, coupled with its technological affordances, enabled users not only to autonomously balance their opportunities of value creation and risks of information disclosure but also to democratically value each person's contribution.

Against this backdrop, personal information practices in a variety of forms of UGCs on Steemit can be characterized not only by creative and qualitative expressions of self but also by active information investment and diversified self-expression strategies stimulated by an optimistic outlook for the cryptocurrency market boom. The value of personal information was found to be significant among the contextualized display of self around personal values, philosophies, and beliefs as well as personal stories, greater than or comparable to that of typical consumer profile data that big tech companies usually collect. These findings suggest that the way people value our data differs from conventional market logic and that the collective prospects and beliefs toward the currency or platform may provide a turning point for individual data subjects to leverage and share their personal information on the platform.

Despite the presence of rewards, users' self-expression does not always reveal more information for more incentives but exhibits privacy-aware identity practices. By remaining completely anonymous, maintaining boundaries between personal and professional identities, or balancing self-presentations using limited multimedia elements, users are aware of privacy risks on the platform. However, the more investment of personal information, the higher the reward tends to be. In other words, if one invests/shares more

information to express oneself in-depth, these efforts are likely to be highly recognized by other community members who vote to distribute rewards. Visualized self-representation such as selfies and vlogs or creative artistic expressions are also highly valued. Altogether, the incentive mechanism of Steemit had the potential to enhance the value of UGCs with a substantial investment of personal information, creativity, and originality. This underlying value rationale is concurrent with the basis on which legal property rights are granted to the database.

Our result shows that creative competence is an important condition for successful functioning in the blockchain-based social media network environment, but it does not necessarily have to be sophisticated or specialized skills such as computer programming or web design. In the context of the present study, the creative competence to enhance user's performance may not be the level of competency but the users' *willingness* to create something new on the platform. In this regard, our result implies how existing concerns about the level of education that differentiates one's digital skill acquisition (van Deursen et al., 2017; Elena-Bucea, Cruz-Jesus, Oliveira, & Coelho, 2020) can be alleviated from the perspective of the digital divide. The blockchain-based platform serves as a space of flexible, creative, and playful self-expression and interactions, where people freely set and negotiate their boundaries of empowerment in the process of selling, deriving benefit, or creating value by utilizing their own data beyond mere protection or passive control.

Appendices

Appendix 1.1: Representative posts for topic spaces during Mania phase (Chapter 4)

Topics (% of the M-corpus)	Keywords	Topic contribution (%) ²⁰	Sample representative posts
Theme 1 (32.1) Self- expression	make, get, time, people, know, see, start, want, go, think	0.640	"I Was Thinking A Lot Quite Long Time Who Was I And What I Wanted To Do In My Life What Abilities And Opportunities Life Gave Me You Know Some People Born Without It But It Is Not Problem Why Because You Can Always Work Hard Dream On And Try To Reach Your Goals You Need Motivation And You Will Reach Whatever You Want In Your Life And Still I Am Thinking What I Can Do I Am Not Quite Talented Girl I Have Talent In Eating Make Up And Sleep But Does It Will Help Me To Reach My Goals Of Course No I Need To Be More Active And Work Harder And Harder And You Can See Me Being Here Trying To Work Hard To Make My Dreams Come True Maybe Someday I Will Become A Role Model For Someone Who Knows I Have Many Goals And Dreams In My Life To Reach And Make Them Come True But I Am Realistic And Know Very Well That Everything Needs To Work Hard You Need To Deserve Something You Want And I Am Not Afraid I Will Challenge Myself And See Who I Will Be Next Day Next Month And Next Year I Need To See Progress In My Job I Will Build My Own Homeless Animals Shelter I Will Help Them Care And Feed Them I Will Make My Family Proud Of Me I Will Make My Haters Jealous Of My Success This Is What Keeps Me Alive If I Will Not Start Today Now Then When Do I Will Ever Have More Ability Motivation And Power To Become The Person I Want To Be Step By Step No Matter How Hard It Will Be I Will Make My Dreams Come True This Is Who Am I This Is What I Desire I Feel That I Have A Great Future I Am On The Right Way ██████████"
		0.608	"HELLO Welcome to my own Canvas Ego Osaji from Nigeria I would like to think my line of thoughts is often misunderstood Guess a few even think I am weird but hey Weird is good at least I like to think so when I am constantly being questioned I take family as important and that s because I have seen firsthand the importance of family Don t get me wrong I know some good people but family shows up even at your worst some people might disagree with this but mine does Growing up in Nigeria has been an experience At times I am tempted to think it s the hardest place to be I have always been independent and very vocal I am quick to say how I feel and people find it offending well Sorry if I won t let you talk me down I never worry about things especially things I can t change A place filled with negative energy is not a place u will likely find me I like my space and thoughts The time I spend with my earpiece music is the best way to escape your thoughts a notepad and a pen are my best Challenging myself is something I do on a steady It gives the opportunity to rave about something new and exciting And when I fail at it I try again and again until there is a breakthrough Little wonder I started a YouTube channel while working a hectic 95 job not sure I can say I have succeeded at that yet Writing has always been fun and the only way I express myself when am not speaking a lot is through words Words are everything Take a bold step explore move

²⁰ LDA topic modeling assumes that a document has multiple topics. It calculates how each topic is weighed in the document and presents the percentage of topic contribution to the document. In this way, the study selected the sample posts for which a particular topic showed the highest contribution rate.

		0.554	<p>from that step It will only take a bit of courage Nothing comes easy I know how long it took me to decide on this post I rather pray hard and work smart Don t let other peoples story blindfold you from your own reality Life is never going to give the same results no matter how similar the solutions are Make sure you are doing what makes you happy You owe yourself that much Am so glad I get to share this with you all and look forward to having a great time on Steemit Let s connect Let s share Let s take little steps together”</p> <p>“Hello dear friends My name is [REDACTED] and I have no idea how to construct this Introduce Yourselftext I usually just go with what comes to my mind but with this one I wanted to make it special So I looked up ways to introduce yourself online Let s see how that goes shall we Maybe you even learn something for your own Introduce Yourselftext Establish a connection with your reader I don t know about you but I am the queen of procrastination meaning that I always think about the things I could should do But I just end up not doing them For a long time Until the deadline is right at my doorstep I try to tackle said problem with planning but sadly planner peace is not achieved easily You have to test and test and test in order to find the system that works for you I have not arrived in my planner yet but I would like you to accompany me on this journey to a more organized and productive life Be Personal For a long time I have felt alone with my love for planning and organising I am by no means the most organised person on earth but I just get excited over simple things like storage boxes different types of paper and things I can DIY for my flat There are lots of projects floating around in my head that I would like to share with someone And that someone could be you State Facts that planning is good for your health Imagine that Being mentally rewarded by planning a topic which you might find headache inducing right now It can help you to organise your time accordingly Who knows Maybe you can finally make some time for your hobby which you might have neglected for way too long I do not want to pressure anyone into anything I simply want to try out certain techniques together with you Engage in discussions about specific topics and just enjoy exploring with you Making mistakes is part of this whole process The only real mistakes are the ones we don t learn from Henry Ford With that being said Hello This is me this is what I would like to do”</p>
Theme 2 (23.8) Self-disclosure	love, friend, name, thank, share, good, new, year, hope, learn	0.665	<p>“Hello steemit Let me introduce myself to my friends After I heard and learned about steemit from my friends I am very motivated to join this steemit program So my name is [REDACTED] I was born in PEUREULAK on [REDACTED] And I live in [REDACTED] And my profession is still in learning or learning about other sciences both religious knowledge and other general knowledge My hobby is playing football and badminton I have one brother and two sisters and I love them so much I m joining here I really hope to get to know each other familiar with friends and should be with friends can help me in studying the science of steemit more deeply with the support and help of my friends are very grateful And I really hope with information from my friends can learn a lot of science later as well as the information I am describing should be useful to me and my friends later I am [REDACTED]”</p>
		0.627	<p>“Hii steemit Let me introduce myself to my friends After I heard and learned about the steemit from my friend s am very motivated to join this steemit program So my name is [REDACTED] and I live in AcehIndonesia precisely in [REDACTED] and My profession is now still in the study or exploration of sciences both religious knowledge of other general science Join me here I really hope to get to know each other with friends and should be with friends can help me in studying the sciences of steemit more deeply with the support and help my friends are very grateful And really hope that with the information from my</p>

		0.611	<p>friends can learn a lot of science later so with the information that I describe should be useful for me and for friends later For the attention of my friends say many thanks Do not forget follow [REDACTED]</p> <p>"Hello steemian around the world today I want to introduce myself because I am a new user here therefore I need a lot of help from all friends in steemit Before that I want to thank the CEO of steemit ned because with steemit I can share various posts on my blog hopefully my post will be useful for everyone And I also really hope to get support good karma as app owner esteem and I also look forward to the support of my brother who has been in steemit for a long time surpassing google purepinay steemphcebu Because I m inspired from all of you MY IDENTITY My name is [REDACTED] I am the first child of two brothers who have one younger brother and one person I come from the philippines precisely in the [REDACTED] I work in entrepreneurship and I also like to make sketches and animated drawings and I am also very much looking forward to the support of the cebu steemit community thank you and here are some pictures of my sketch hope you like it Thank you for visiting my blog [REDACTED]"</p>
Theme 3 (13.7) Personal narratives	year, day, go, take, time, back, leave, come, home, move	0.553	"I would like to invite you to a tour of my home in Oslo Norway As the title suggests I want to share a tour with steemit of my Home Its not exactly the Norm that your used to I guess so check it out I am living here with a mate and and 3 dogs between us staying here since the beginning days of August 2017 I thought Id do first a blog post about where im living and how I got to live here There are 1000s of unoccupied bulidings ranging from factory s to houses train stations and old farms in literally every county People like myself have realised that these empty places should be used providing people with a place to live or to be used as community centers etc If your ever travelling and are doing the hostel way but want to travel for free instead then you can see this symbol on buildings or stickers on lampposts etc....."
		0.542	"My journey to steem has been a long time coming My life has been filled with many challenges that I ve had to overcome long the way everything from my father being in and out of prison to meth labs to foster care to adoption and even being hung in a tree But for now Ill start at the beginning My name is [REDACTED] and I m [REDACTED] years old I m originally from [REDACTED] To be more specific I began my life Portland This is a photo of the home I grew up in as a child Portland is a miserable area of Louisville were the median income is less than 20 thousand dollars However this isn t were my life gets interesting Back on [REDACTED] my father went to prison for the first time he was charged with transporting greater than 70 lbs of pot from Texas back to Kentucky but was apprehended in Tennessee The original police stop was for following to close to a semi It was during these years from 2000 to 2004 or the age of 1014 that the lives of both my brother and I had changed Soon after my father went to prison....."
		0.486	"And so let s start with the fact that I live in an old house although almost two generations of my relatives lived there but not about this now Let s move on to the story itself In general I went to bed I usually sleep with a nightlight but that night I did not find it and I will not hide that I was insanely lazy to look for it in the pile of rubbish under the bed I will not say it garbage it s bags with clothes and all that In short she did not look and fell on her side I m asleep And I have a very strange dream we have such a small field in which the goats graze there is a lot of grass and there is a path to a small grove on this path along the sides of the nettle and black caterpillars on it like I m walking along this path is the last all my friends have already passed but instead of the trampled path there is a burdock I go by means of this mug and suddenly I see this caterpillar jumping through the mug and a bunch of

			butterflies fly over me straight on my head....."
Theme 4 (11.0) Personal relationships	life, music, love, family, child, believe, man, always, person, dream	0.422	"Hello steemians my name is [REDACTED] but my friends call me [REDACTED] I m a medical student and in stay in [REDACTED] I m not your typical medical medical student I believe life should be enjoyed because you only live once A little about myself I come from a small family family of six actually I have no sister so my mum is the only female in our house My dad is an engineer while my mum is a nutritionist Food is an important part of our lives so yeah we re all foodies I decided to study medicine and specialise in obstetrics and gynecology because of some complications my mum experienced while she was pregnant with me I have this burning desire to establish effective health centers in rural areas of my country to help reduce the mortality associated with child birth That s the four of us with our mum You can call us her bodyguards Now more about me music is a very important part of my life I love music I live and breathe music I used to write when I was younger dont ask my why I stopped cos I dont know I m not an emotional person but good music makes me emotional Music is meant to inspire and uplift I believe music can heal the soul I believe in love I believe in taking chances I believe in dreams I believe in miracles and I believe there s a God
		0.349	""Hello steemit community I am [REDACTED] from New Delhi India I am [REDACTED] years old and i am working as a software developer in new Delhi I am very self motivational person and i like those person who are self motivate I think that no one can motivate you until you will not self motivate I like to meet and communicate a new person because every person has different think and different experience about the beautiful world I am very happy after become a part of this amazing community because with the help of this i will meet and communicate with new peoples and i can share my thinks and ideas with them it will help me to know that what think peoples have about the beautiful world I have more interest to know about the world About my family I am very lucky because i have my mother and father who love me very much i am only single son of my parents I love my family I belongs to very reputed family There are three members in my family me and my parents I would like thanks to god for giving me a so lovely family I think that a person who have hisher parents is a very rich person in the world If anyone believe in god than i want to say that person your god is your parents Everyone should respect hisher parents due to them you are in the world I hate that person who does not respect hisher parents Parents are the very big gift from the god Today s some peoples do not respect of hisher parents i want to tell them please meet those peoples who have lost hisher parents and talk about the parents you will understand the value of parents Some peoples leave hisher parents due to money and come out from hisher country i want to say those person sometime money is not everything you can get every things of the world by money but if once you lost your parent you will never get back so please respect hisher parents Those peoples who do not respect hisher parents they never get a real happiness I request to everyone please never heart hisparents I m so happy to be a part of this amazing community Thanks to all Regards By [REDACTED] Don t forget to follow and upvote"
		0.347	"Good day everyone I am [REDACTED] I just joined this wonderful community and I think I have to introduce myself to the community I am a God fearing man and am anointed as well God has blessed me with a wonderful ministry God has given to me a power to ask and be fulfilled Anything I touched has never failed because the power of my father is in me I travelled to many countries to spread gospel and many people have repented from their sins through my preaching All you need to

			<p>keep at the back of your mind is that the Lord loves you If you have that believe the Lord will soften your mind to listen and understand his message The devil has corrupted many souls lead many people to wrong path through the power given to him But because of the love God has for His people he gave His holy book to us Who ever believe in it shall never perish I m here to transform people s lives people who are ready to give their lives to Christ I also want to have a church located here on steemit to feed people with words of God and to lead them to the eternal life Feel free to consult me whenever you have need any assistance God has been doing his work and he will continue to do it May God bless you all I do need strength Because the bible says For wisdom is a defence and money is a defence but the excellency of knowledge is that wisdom giveth life to them that have it Ecclesiastes 712 And also A feast is made for laughter and wine maketh merry but money answereth all things Ecclesiastes 1019 Support the gospel mission on earth and God shall reward you in heaven"</p>
Theme 5 (8.9) Instrumental focus: Steemit platform	use, create, content, post, video, account, project, comment, follow, website	0.607	"vertical daily Steemit report service on Steemitcom searches collects and analyses Steemit posts daily based on quality in most popular Steemit category s Here are results for introduceyourself category sorted by our top earning algorithm Please support authors of this articles upvoting following TOP posts....."
		0.451	"Technology is one of the most effective communicating method of living one s life to the fullest in this century Without technology this generation will be on a spot INTRODUCTION TECHHUB is a blogging page which focus on brining ease life to blockchains this project page will anchored by a first class computer and software engineering teacher which will be blogging about software computer resources on how steemians can be their own computer guru in all areas of networking and blogging....."
		0.428	"Hey new and also old Steemians I want to tell you about these votingfollowing tools made for Steemit They are great tools for minnows to gain votes and at the same time stop bad content and plagiarism The first tool is called Steem Engine They make it simple by explaining it in smalls steps....."
Theme 6 (4.2) Personal history of education & work	school, business , company, be, work, teacher, education, study, computer, student	0.293	"I am [REDACTED] into writing multilevel business and student unionist a progressive Nigerian Gained admission to the Federal School of Surveying Oyo for National Diploma certification in Surveying and Geoinformatics and later to the Federal University of [REDACTED] for Bachelor of Technology in same course study As a student intrapreneurtrader [REDACTED] has been influenced by among others the King of network marketing in Nigeria [REDACTED] is presently a 400L student at the [REDACTED] where he represented the department of [REDACTED] at faculty and currently representing at university level of the legislative arm of student union governance likewise served over 20 thousand students as the Chief Press Secretary to the Speaker of the student union 1617 parliamentary session My love and serious philosophic zeal for better governance has made me involve in advocacy during tertiary education little wonder he was able to propose amendment of 2013 SVG Constitution which was adopted which a joy of student legislature and raised an eye over epileptic process of transition of power hereinto this brought new dawn wherein executive members of the department handover before leave for Industrial Training likewise full involvement of lower level students in process decision making and execution of policies"
		0.291	"Assalamualaikum Wr Wb Hello all the steemit community friends from Sabang Island to Maroke Island and all over the world I am a beginner steemit community I was told and taught by my best friend [REDACTED] thank you for he who has informed me with enough patience My name is [REDACTED] and often called [REDACTED] I

		0.271	<p>live in [REDACTED] and the Republic of Indonesia I am the fifth child of six siblings we are all four men and two women Father named [REDACTED] and Mother [REDACTED] My education history is Madrasah Ibtidaiyah Negeri MIN Peusangan Siblah Krueng and graduated in 2005 the best graduates Peusangan State Junior High School Siblah Krueng and graduated in 2008 Secondary School SMA Negeri 2 Dewantara and graduated in 2011 [REDACTED] Faculty of Teacher Training Education FKIP Department of Primary School Teacher Education Program PGSDS1 Bachelor level My daily activities are a classroom teacher at [REDACTED] I am now almost 3 years of service in guiding the children in the school I am also a teacher of teaching the religion of Islam Mengaji at home My Motto Failure is the beginning of success"</p> <p>"Hello I am [REDACTED] I am An Educator Tech Enthusiast Developer Media Pro Computer Engineer Entrepreneur Data Scientist Community Builder [REDACTED] has over 3years of experience in Community Growth Currently [REDACTED] is a Computer Teacher at [REDACTED] where I teach Computer Science and also training the student on Computer Pratical Hardware Software I am Committed to the growth of the young ones around me I have volunteered at different program for the community growth eg CYFI Project Skill Up TeenCode Andela Learning Community I held different post when I was in Federal College of Education Technical where i study Computer Science Education Integrated Science Edu such as Google Student Ambassador Mozilla Firefox Student Ambassador Nigerian Association of Computer Science Student NACOSS President NACOSS South West PRO I am here because it affords me the opportunity to work together with likeminded people under the guidance of experienced mentors to make and deepen my impact in my chosen sector of interest ie education and tech also enables me access to networks and developmentrelated resources which would prove useful and invaluable for my career as a passionate social worker and community developer and as well helps me to add value to the young generations Thanks to [REDACTED] for your support Looking forward to a great experience on steemit"</p>
Theme 7 (3.0) Creative work	travel, photo, art, picture, photogr aphy, draw, place, movie, artist, adventu re	0.357	<p>"hello friend i am draw a simple art for you please watch this 3d drawing drawing step by step3d modelseasy drawings Easy 3D 25 Best Ideas about 3d Drawings on Pinterest 3d writing FunArt Pencil Drawing How to Draw 3D Dew Drop on Ldraw 3d eaf 5 pencil drawing tutorial drawings out of words how to draw with man e word cat into a cat how to turn girl into a cartoon how to turn words into drawings how to draw 3d art on paper step by step 3d drawing drawing step by step3d modelseasy drawings how to draw 3d art with pencil how to draw 3d art easy how to draw 3d hole on paper step by step 3d drawing techniques how to draw 3d drawings step by step with pencil for beginners"</p>
		0.276	<p>"Hi Im [REDACTED] and I run [REDACTED] Studios Its a Toronto based Wedding Photography Fashion and Commercial imaging studio located in the heart of Toronto We provide curate some of the most indelible images for weddings fashion editorial and commercial photography due to the detailed work of owner [REDACTED] [REDACTED] alumnus in photography whom approaches her work as an artist first and foremost My practice in imagery excels in her editorial work for such iconic magazines such as Story of Fashion in which she paints a story narrative in each image The style is evident in my wedding photography in which is very distinct in the industry Not only I do capture and find the most endearing moments that are raw and spontaneous but my editorial fashion side adds a special artistry that makes each image more than a wedding image they look like a moment of magic and art My team and I are proud to be initiating this new step in their business"</p>

		0.253	<p>to provide a special connection from your wedding to your family from your commercial goods to your clients or your editorial to your readership demographic I focus on the connection and has a unique feel for each images effectiveness and lasting effect My work is about the character personality and a lasting image that is invaluable This up above is ME And yes at times I ll get in front of the camera"</p> <p>"Hello friend My name is [REDACTED] I m photographer traveler freediver underwater photographer sea hunter and blogger from Magadan city in the Russia I like to photograph wildlife and its inhabitants I like to photograph wildlife and its inhabitants Photography is my main profession I have been shooting landscapes and animals for 14 years Magadan is in the north and I take my pictures in different weather conditions A few years ago I was carried away by underwater hunting and freediving It s so interesting that I started diving at different times of the year Once I dived under the ice to the sunken ship The underwater world is so beautiful that I changed the gun to the camera Under the water I take my pictures only on a delay in breathing I was glad to introduce myself Have a nice day"</p>
Theme 8 (1.5) Instrumental focus: STEEM coin and other cryptocurrencies	crypto, invest, bitcoin, cryptocurrency, coin, exchange, trade, dollar, buy, money	0.300	<p>"Hello My Coin Brothers and sisters today i just wanted to share this with You My Cryptocurrency Picks For 2018 on 7 January 2018 i invest 200 on altcoins now am i invest 1000 and on my calculation its going to be 400 000 Depending on my own reasearchs so lets Start with Coin number 1 1 Bytecoin byte is a privacy crpto curuency last year it started the year with 10 Million Dollars Market Cap now the Market cap of this coin is over 18 Billion Dollars and i bought 26K bytecoins 2 Ripple Ripple its number 3 of market caps of all coins which is amazing and its the best coin for banks at these moment last time i bought 57 ripples now i bought 100 ripples which bring this to over 157 ripples 3 Verge Verge its a privacy coin also i already had 35 coins now i bought 1400 XVG which make it on total 1435 Verges 4 TRON am not sure of this coin features but its getting a lot of partnerships these days and i bought 2k Trons for long term 35 years 5 Stellar i bought 189 stellars i still have 400 Please suggest some altcoins to me and i will do my reaserchs and buy them anyway am not good adviser and i didnt tell anyway to buy these altcoins do your own researchs and if you like it you buy it hope you enjoyied reading see you in the next blog D"</p>
		0.255	<p>"Hello guys today we are going to talk about crypto gonna change your like so enjoy reading about why you should invest in bytecoin What is Bytecoin Bytecoin BCN is the first cryptocurrency based on the CryptoNote technology and launched in July 2012 with an open source code designed for anonymous cash settlement BCN protects the user s privacy with impassive and anonymous transactions Bytecoin Features 1Unlimited instant international payments the Bytecoin network works as fast as the Internet Payments take a little time to process as confirmation is required for payment Bytecoin claims that transactions are processed in about 2 minutes 2Bytecoin protects the user s money with secure and modern crypto algorithms which can not be hackedThe creators of Bytecoin argue that breaking a crypto currency will require a lot of expensive electricity and the computing power of supercomputers Why you Should invest in Bytecoin last year when i bought 2k bytecoins it was as i remember 000043 and now the price of 1 bytecoin is 001 thats mean what cent also its the father of the famous crypto curuency Monero its one of the best coins on the market now listen to this if you buy 100 000 bytecoins and you keep if for one year you gonna get 100k if the price of this coin hits 1 and im sure it gonna happened soon anyway i hope you enjoyed reading see you"</p>
		0.217	<p>"My name is [REDACTED] a graduate of Computer Science from Nigeria I work as a full time Media Specialist in TV program transmission from a reputable Media Firm in Lagos I have my keen interest in Bitcoin and</p>

			Cryptocurrency trading and training people on how to trade My vision in life is to eradicate poverty and build a community of young millionaires across Africa continent though putting them through on how to trade and and how to make money in the crypto world A friend introduce me steemitcom and I find the platform as one of the best platform to actualize my dream Kindly follow me solad to see my tutorials on what bitcoin is all about and how to trade bitcoin with other crptocurrency using technical analysis"
Theme 9 (1.2) Lifestyle & leisure	food, eat, cook, cat, recipe, colleagu e, fully, water, garden, meal	0.374 0.268	food recipes "Create Food Lovers who like the road show certainly less afdol if you ga nyobain typical food in the place you visit probably from the many traveler who often jalan2 there are some who do not know know or try a typical food from the place that has been visited for that I will share a little knowledge about the typical foods that you can later try if you visit places of tourist in Indonesia and the following 21 Food Cuisine Typical from various regions in Indonesia"

Note. Excerpt posts are unstructured as these are directly scraped through Steem API and went through preprocessing processes.

Note. To protect the privacy and identities of post authors, personally identifiable information such as user IDs, names, addresses, and other sensitive information has been blacked.

Appendix 1.2: Representative posts for topic spaces during blow-off phase (Chapter 4)

Topics (% of the B-corpus)	Keywor ds	Topic contribution (%)	Sample representative posts
Theme 1 (49.1) Self- disclosure	time, make, love, good, work, know, want, people, get, new	0.887	"My Steemit Presentation Hello everyone My name is [REDACTED] and I m new to Steem I don t know how to make a proper introduction but I guess I just did didn t I so I m just going to talk a little bit about myself I am [REDACTED] years old [REDACTED] in less than a month and I m from Venezuela My native language is Spanish and I m currently studying French and English I like writing in English it helps me practice and I think I could reach more people this way That being said I m still thinking about writing some stuff in Spanish and maybe even in French but not just yet I m still a newbie I guess we ll see I love languages and I want to learn German and Italian in the near future and certainly others Hopefully I will become an interpreter soon I will be talking about languages from time to time maybe recounting my experience with languages or some tips to learning or maybe just stuff that I like and I think you guys could like as well Another passion of mine is film I love movies I love watching movies I love researching movies I love reviewing movies and hopefully one day I will be making them Also I ve always loved acting and I have been in a couple of plays so there s that too D But in the meantime I will be writing movie reviews I love writing I might also write about my journeys about things that I love maybe funny and interesting things that have happened to me and maybe just maybe I could write some fictional stories Other topics of interest of mine are philosophy and politics but I don t know if I ll be writing about those any time soon again I guess we ll see I m very excited about writing in this website as I ve said I ve always loved writing and I think the tools that I need to express

		0.864	<p>myself are laid out here So yeah I think that s it Thanks a lot for your time ██████</p> <p>"Hello everyone My name is ██████ and the time has finally come to do my first ever post on Steemit I would like to introduce myself shortly but properly and let you know what i hope to learn from this community I m ██████ years old finished Bachelor of Science in Computer Engineering from ██████ This platform was introduced to me by ██████ so credits to him by bringing me here The platform got me so excited because of the knowledge that ██████ and ██████ shared to me so I have my high hopes that I can learn more or even contribute to the community I don t really have an hobby where I can brag about I just have fun on things that I find very interesting especially if I have fun on it with my friends For the time being I think I m having fun playing volleyball playing mobile and computer games and watching vlogs and animes Thank you to ██████ who help me a lot on how this kind of platform works I d also like to tag my fellow colleagues namely ██████ ██████ and ██████ Thanks Steem ██████ Follow me"</p>
		0.858	<p>"Aloha Steemit community I want to introduce myself so you will know who I am and what I do My name is ██████ aka ██████ I am a free spirit from French living in Portugal and traveling around the world With me I have always my Canon 77D Camera because I am also a photographer My passion is skating surfing and motorcycles I love wild things that s why I choose this name Also you can find me on Instagram with the same username I heard about Steemit from ██████ I met her on one of my journeys and she said it would be cool if I registered so here I am I will show you my photography s and tell about my traveling city s and other stuff I am really exited how this platform works and how the community is So we will see what the future brings Much love Your ██████"</p>
Theme 2 (18.2) Instrumental focus: Steemit platform	use, content, blockchain, project, create, account, user, support, platform, need	0.622	<p>"Hello Steemians Meet entrust Website Link entrust is a bot focused on rewarding 100 to delegators aiming to help Steemit users promote their posts We have made this to give value to the community by sharing all the bid rewards to delegators and intending our voting power to run the system Given the name itself we want to build a service to our bidders and delegators that they can entrust entrust is focused on developing into a sustainable service that benefits the community....."</p>
		0.621	<p>"STEEMITALIA steempostitalia is a dynamic project created almost two years ago by a small group of Steemit Italian users Since the beginning our goal was to make the experience on Steem Blockchain and the use of the blogging platform Steemit more accessible and usable for Italian users So we created a community based mainly on that supports the Italians and the Italian speakers helps new users with no knowledge of the blockchain ecosystem and tries to spread steem as much as possible in our country For this purpose we offer specific guidelines and personalized tutoring processe....."</p>
		0.616	<p>"Project Emprendedores 200 is a community that is in the process of expanding for the improvement and performance of the blockchain and the growth of the platform called Steemit dedicated to attracting and attracting new users content creators artists in general among others in order to train them train them and support them so that they can make good use of their skills within the conception of the Newsteem the Smart Media Tokens SMT of steem and other adjacent platforms....."</p>
Theme 3 (12.5) Personal narratives	go, say, time, take, end,	0.565	<p>"All my life I never believed it could happen till it happened Death is the nickname of a man who was basically our nextdoor neighbor he is 7 feet tall and has a bald head I was barely 16 when I almost got killed by Death he haunted me for months and the whole story ended up this way I woke up one morning just to find out that I was in his house and</p>

	look, never, back, come, move	0.509	<p>was tied to a pillar in his house The question in my mind was how did I get here but before I could ask myself he spoke you re quite a kid you know I like you but there s a problem about you which I m taking personal I could barely breath at this moment because my heart was pumping”</p> <p>“My friends definitely know me as eccentric even extreme I have huge ambitions and dreams of changing the world and I want the world to see this story as it plays out LIVE TriggerAlert What is This I think I have one of the most interesting lives of anyone you d ever meet The experiences I ve been through and choose to be a part of are like no other The way I handle myself and take on the world will shock some and confuse others I ve been through quite a lot these last few years and I have been video recording my life consistently for 18 months and have rapidly increased the frequency to where I might as well be livestreaming the whole thing if that were logistically realistic This will be the greatest story ever told</p>
		0.436	<p>“I d like to tell others what night life is like for me I get 2 maybe 3 hours of sleep a night Between my insomnia and anxiety sleep is more of a myth for me Last August my dog died of parvovirus even though he was UTD on shots watching him slowly and painfully die killed me making the decision to end his suffering ruined me and mentally messed me up even more All I could think about was death and the fact I will die one day and I cannot stop it but it s not just death itself I fear it s will I suffer like Bowser did what about my kids is there something waiting for me at the end of it all I started to get sharp unbearable pains behind my left eye and temple area the beginning of the year it scared me because I thought to myself this is it it s a tumor and now it s my turn to die slowly went to my doctor and he kept saying I didn t show any red flags.....”</p>
Theme 4 (6.8) Personal relationshi ps	life, child, family, feel, mother, man, parent, girl, young, teacher	0.327	<p>“Blessings to all Christians of this platform I am new here a Christian friend recommended me steemit There are ways that seem true to man but in the end it leads to death In life there are two kingdoms that of God and that of men of image The kingdom of men is more compatible with our flesh in it are the riches the power and all the interests that are proper to men The kingdom of God is the will of Him not ours it is his life not ours it is all to Him for Him It is more convenient to rely on the strength of our arms to trust in something more than thinking to look for to call and ask”</p>
		0.326	<p>“Raising children can be very challenging Some children are very naughty and disobedient all the time while other children are only naughty on occasion Keep in mind when dealing with a naughty child that you should recognize that it is the behavior that is bothering you rather than the child The child may have an unmet need and their behavior may be an attempt to get that need met You can help the child by providing a safe space for them to tell you what they need Take steps to create boundaries deal with tantrums deal with bad behavior and reinforce good behavior with the child and you will be raising wellbehaved children in no time If you are taking care of children that are not your own you can take steps to teach them to behave without undermining the authority of their parents”</p>
		0.305	<p>“Hi there My name is [REDACTED] I am a Soul Teacher and Tao Calligraphy Practitioner My main purpose to share is about the power of self healing the power of the soul and the heart My second purpose is to help people to deeper understand Bitcoin and therefore start accumulating Satoshis 1 BTC 100000000 Satoshis The soul is a light being Everyone and everything has a soul every plant animal planet organ cell and human The soul has incredible power for selfhealing and</p>

Appendix 2: A sample of Steemit user recruitment document (Chapter 6)

Hello Fellow Steemians,

This post is to invite interested members of Steemit to complete a short survey about your Steemit activities. The survey is related to the completion of a doctoral dissertation research of Soyoung Park who is pursuing a doctoral degree in Media Studies from the Moody College of Communication at University of Texas at Austin. Upon completion of the survey, you will receive 5-STEEM coin upon for your provided user ID.

This current research is interested in understanding how one's digital experiences and capabilities influence his/her value-making activities and personal data management practices in the context of a blockchain-based data trading platform. This survey will ask you a series of multiple choice and other types of questions about your Steemit activities and your perceptions of, knowledge of, and experience with information and communication technology practices.

By surveying Steemit members, the study hopes to gain insight for some broader understanding of information privacy online and personal data management practices, lessons for individuals to secure rights and benefits for their contributions to our digital communities, and guidelines for designing better digital literacy education curricula.

Your participation is voluntary and you may skip any questions or stop your participation at any time. All information collected for the study will be kept confidentially and used for the scholarly research purposes only.

The study should take no more than 15 minutes.

Thank you in advance for your participation.

Appendix 3: The table of Heterotrait-Monotrait ratio (HTMT) Results (Chapter 6)

The heterotrait-monotrait ratio (HTMT), proposed by Henseler, Ringle, & Sarstedt (2015), uses correlations to empirically evaluate the extent to which a construct is distinct from other constructs. Put simply, HTMT is an estimate of the true correlation between two latent constructs, assuming these two constructs were perfectly reliable. An HTMT value close to 1 implies that the estimated constructs are very similar, representing a lack of discriminant validity; the common cutoff criterion should be 0.9 or lower, or 0.85 for more conservative interpretations. The result table below confirms that the HTMT values are less than 0.85, ensuring each construct to be substantially distinguished from the others.

	CP_CS	CP_IS	CP_OS	CP_SS	IU	PD_S	PD_T	SU	UP
CP_CS									
CP_IS	0.117								
CP_OS	0.673	0.291							
CP_SS	0.689	0.177	0.789						
IU	0.422	0.434	0.662	0.427					
PD_S	0.337	0.309	0.175	0.209	0.178				
PD_T	0.354	0.173	0.4	0.325	0.462	0.598			
SU	0.317	0.158	0.186	0.174	0.51	0.109	0.218	0.283	
UP	0.092	0.122	0.097	0.144	0.315	0.096	0.171	0.062	0.538
CFIP_C	0.375	0.1	0.34	0.312	0.181	0.137	0.19	0.109	0.096
CFIP_A	0.516	0.265	0.66	0.75	0.482	0.149	0.295	0.186	0.079
CFIP_E	0.39	0.16	0.367	0.482	0.177	0.401	0.39	0.251	0.101
CFIP_U	0.483	0.233	0.586	0.582	0.416	0.178	0.462	0.51	0.315

References

- Acquisti, A. (2004). Privacy in electronic commerce and the economics of immediate gratification. In *Proceedings of the ACM Conference on Electronic Commerce (EC 2004)*, pp. 21–29.
- Acquisti, A., & Grossklags, J. (2005). Privacy and rationality in individual decision making. *IEEE security & privacy*, 3(1), 26-33.
- Acquisti, A., Taylor, C., & Wagman, L. (2016). The economics of privacy. *Journal of Economic Literature*, 54(2), 442-92.
- Aoki, K. (1996). (Intellectual) property and sovereignty: Notes toward a cultural geography of authorship. *Stanford Law Review*, 48(5), 1293-1355.
- Awad, N. F., & Krishnan, M. S. (2006). The personalization privacy paradox: an empirical evaluation of information transparency and the willingness to be profiled online for personalization. *MIS quarterly*, 13-28.
- Balkin, J. M. (2016). Information fiduciaries and the first amendment. *UCDL Rev.*, 49, 1183.
- Balkin, J. M., & Zittrain, J. (2016). A grand bargain to make tech companies trustworthy. *The Atlantic*, 3.
- Bandura, A. (1998). Health promotion from the perspective of social cognitive theory. *Psychology & Health*, Vol. 13 No. 4, 623-649.

- Bandura, A. (2002). Growing primacy of human agency in adaptation and change in the electronic era. *European Psychologist*, 7(1), 2-16.
- Bandura, A. (2004). The growing primacy of perceived efficacy in human self-development, adaptation and change in the electronic era. *European Psychologist*, 7(1), 2.
- Barth, S., & De Jong, M. D. (2017). The privacy paradox—Investigating discrepancies between expressed privacy concerns and actual online behavior—A systematic literature review. *Telematics and Informatics*, 34(7), 1038-1058.
- Batenburg, A., & Bartels, J. (2017). Keeping up online appearances: How self-disclosure on Facebook affects perceived respect and likability in the professional context. *Computers in Human Behavior*, 74, 265-276.
- Bechmann, A., & Lomborg, S. (2013). Mapping actor roles in social media: Different perspectives on value creation in theories of user participation. *New media & society*, 15(5), 765-781.
- Belliger, A., & Krieger, D. J. (2016). From quantified to qualified self. *Digital Culture & Society*, 2(1), 25-40.
- Bellman, S., Johnson, E. J., Kobrin, S. J., & Lohse, G. L. (2004). International differences in information privacy concerns: A global survey of consumers. *The Information Society*, 20(5), 313-324.
- Benkler, Y. (1999). Free as the air to common use: First Amendment constraints on enclosure of the public domain. *NyuL Rev.*, 74, 354.

- Benkler, Y. (2006). *The wealth of networks: How social production transforms markets and freedom*. New Haven: Yale University Press.
- Berthold, S., & Böhme, R. (2009). Valuating privacy with option pricing theory. In T. Moore, D. Pym, & C. Ioannidis (Eds.), *Economics of information security and privacy* (pp. 187–209). New York: Springer.
- Bird, S., Klein, E., & Loper, E. (2009). *Natural language processing with Python: analyzing text with the natural language toolkit*. Sebastopol: O'Reilly Media, Inc.
- Blei, D.M., Ng, A.Y. and Jordan, M.I. (2003) Latent Dirichlet Allocation. *The Journal of Machine Learning Research*, 3, 993-1022.
- Bourdieu, P. (1986). The Forms of Capital. In J. Richardson (Ed.), *Handbook of Theory and Research for the Sociology of Education* (pp. 241 – 258). New York: Greenwood.
- boyd, d. m., & Hargittai, E. (2010). Facebook privacy settings: Who cares? *First Monday*, 5, p. 8.
- Boyle, J. (1996). *Shamans, software, and spleens: Law and the construction of the information society*. Cambridge, Mass: Harvard University Press
- Boyle, J. (2008). *The public domain: Enclosing the commons of the mind*. New Haven, Conn; London: Yale University Press.
- Brandeis, L., & Warren, S. (1890). The right to privacy. *Harvard law review*, 4(5), 193-220.

- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101.
<https://doi.org/10.1191/1478088706qp063oa>
- Byrne, J., Kardefelt-Winther, D., Livingstone, S., & Stoilova, M. (2016). *Global Kids Online: research synthesis 2015-2016*. Florence: UNICEF Innocenti.
- California Consumer Privacy Act of 2018 (2018). Accessible at:
<https://oag.ca.gov/privacy/ccpa>
- California Privacy Rights Act of 2020 (2020). Accessible at:
https://oag.ca.gov/system/files/initiatives/pdfs/19-0021A1%20%28Consumer%20Privacy%20-%20Version%203%29_1.pdf
- Caraway, B. (2011). Audience labor in the new media environment: A Marxian revisiting of the audience commodity. *Media, Culture & Society*, 33(5), 693-708.
- Castells, M. (1998). *End of Millennium: The Information Age: Economy. Society and Culture*, New Jersey: Blackwell Publishers, Inc.
- César, J., Debussche, J., & Asbroeck, V. (2017). *White Paper - Data ownership in the context of the European data economy: proposal for a new right*. TOREADOR research project funded by the European Commission. Retrieved from:
<https://sites-twobirds.vuture.net/1/773/landing-pages/white-paper-form.asp>
- Cheng, V. J. (2004). *Inauthentic: The anxiety over culture and identity*. Rutgers University Press.

- Choi, J. R., Straubhaar, J., Skouras, M., Park, S., Santillana, M., & Strover, S. (2020). Techno-capital: Theorizing media and information literacy through information technology capabilities. *New media & society*, 23(7), 1989-2011.
<https://doi.org/10.1177/1461444820925800>
- Choi, S. (2018). When digital trace data meet traditional communication theory: Theoretical/ Methodological Directions. *Social Science Computer Review*, 0894439318788618.
- Chowdhury, G. G. (2003). Natural language processing. *Annual review of information science and technology*, 37(1), 51-89.
- Christl, W. (2017, June). *Corporate Surveillance in Everyday Life. How Companies Collect, Combine, Analyze, Trade, and Use Personal Data on Billions*. A Report by Cracked Labs, Vienna.
- Cohen, J. E. (1998). Lochner in cyberspace: The new economic orthodoxy of "rights management". *Michigan Law Review*, 97(2), 462-563.
- Cohen, J.E., (2012). *Configuring the Networked Self: Law, Code, and the Play of Everyday Practice*. New Haven: Yale University Press.
- Couldry, N., & Mejias, U. A. (2019). Data colonialism: Rethinking big data's relation to the contemporary subject. *Television & New Media*, 20(4), 336-349.
- Culnan, M. J., & Armstrong, P. K. (1999). Information privacy concerns, procedural fairness, and impersonal trust: An empirical investigation. *Organization science*, 10(1), 104-115.

Daniels, J. (2019, Feb 12). *California governor proposes 'new data dividend' that could call on Facebook and Google to pay users*. CNBC. Accessible at:

<https://www.cnbc.com/2019/02/12/california-gov-newsom-calls-for-new-data-dividend-for-consumers.html>

Davis, J. L. (2014). Triangulating the self: Identity processes in a connected era. *Symbolic Interaction*, 37(4), 500-523.

De Battisti, F., & Siletti, E. (2019). On the Use of Control Variables in PLS-SEM= Sull'Uso delle Variabili di Controllo nei PLS-SEM. In *Smart Statistics for Smart Applications*. pp. 793-798. Pearson.

Donnenwerth, G. V., & Foa, U. G. (1974). Effect of resource class on retaliation to injustice in interpersonal exchange. *Journal of Personality and Social Psychology*, 29(6), 785.

Duch-Brown, N., Martens, B., & Mueller-Langer, F. (2017). The economics of ownership, access and trade in digital data. The Joint Research Centre (JRC) of European Commission.

Egelman, S., Felt, A. P., & Wagner, D. (2013). Choice architecture and smartphone privacy: There's a price for that. In *The economics of information security and privacy* (pp. 211-236). Springer, Berlin, Heidelberg.

Elena-Bucea, A., Cruz-Jesus, F., Oliveira, T., & Coelho, P. S. (2020). Assessing the role of age, education, gender and income on the digital divide: evidence for the European Union. *Information Systems Frontiers*, 1-15.

- Epstein, R. A. (2018). A Not Quite Contemporary View of Privacy. *Harv. JL & Pub. Pol'y*, 41, 95.
- Fairfield, J. A. (2017). *Owned: Property, privacy, and the new digital serfdom*. Cambridge University Press.
- Fairfield, J. A., & Engel, C. (2015). Privacy as a public good. *Duke LJ*, 65, 385.
- Feiner, L. (2018, November 28). *Former FTC official says Facebook representative made false statements to international leaders about how the company handles personal data*. Consumer News and Business Channel. Retrieved from: <https://www.cnbc.com/2018/11/27/former-ftc-official-says-facebook-vp-misled-international-leaders.html>
- Feist Pubs., Inc. v. Rural Tel. Svc. Co., Inc., 499 U.S. 340 (1991). Accessible at: <https://supreme.justia.com/cases/federal/us/499/340/>
- Flath, C. M., Friesike, S., Wirth, M., & Thiesse, F. (2017). Copy, transform, combine: exploring the remix as a form of innovation. *Journal of Information Technology*, 32(4), 306-325.
- Galanxhi, H., & Nah, F. F. H. (2006). Privacy issues in the era of ubiquitous commerce. *Electronic markets*, 16(3), 222-232.
- Garson, G.D. (2016), *Partial Least Squares Regression and Structural Equation Models*, Statistical Associates, Asheboro.

- Glancy, D. (2010, August). Personal information as intellectual property. *Paper presented at the 10th Annual Intellectual Property Scholars Conference*.
- Retrieved from: https://www.law.berkeley.edu/files/bclt_IPSC2010_Glancy2.pdf
- Guidi, B., & Michienzi, A. (2021, June). The decentralization of Social Media through the blockchain technology. In *13th ACM Web Science Conference 2021* (pp. 138-139).
- Granovetter, M. (1985). Economic action and social structure: The problem of embeddedness. *American journal of sociology*, 91(3), 481-510.
- Grieve, R., March, E., & Watkinson, J. (2020). Inauthentic self-presentation on facebook as a function of vulnerable narcissism and lower self-esteem. *Computers in Human Behavior*, 102, 144-150.
- Grossklags, J., & Acquisti, A. (2007). When 25 cents is too much: An experiment on willingness-to-sell and willingness-to-protect personal information. In: *WEIS*.
- Grünewald, L. & Haupt, J. (2014). Value Creation on Youtube: How Musicians, YouTubers and Commercial Networks Create Social, Cultural and Economic Capital. Paper presented on *Young Scholars Workshop at the Vienna Music Business Research Days 2014*.
- Hair, J. F., Hult, G. T. M., Ringle, C. M., Sarstedt, M., & Thiele, K. O. (2017). *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)* (second edition). Thousand Oaks: SAGE Publications.

- Hair, J.F., Risher, J.J., Sarstedt, M. & Ringle, C.M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, 31(1), pp. 2-24.
<https://doi.org/10.1108/EBR-11-2018-0203>
- Hair, J.F., Sarstedt, M. and Ringle, C.M. (2019). Rethinking some of the rethinking of partial least squares. *European Journal of Marketing*, 53(4), pp. 566-584.
<https://doi.org/10.1108/EJM-10-2018-0665>
- Halbert, D. J. (2000). *Intellectual property in the information age: The politics of expanding ownership rights*. Westport: Greenwood Publishing Group Inc.
- Haraway, D. J. (1985). *A manifesto for cyborgs: Science, technology, and socialist feminism in the 1980s* (pp. 173-204). San Francisco, CA: Center for Social Research and Education.
- Hargittai, E., & Marwick, A. (2016). “What can I really do?” Explaining the privacy paradox with online apathy. *International journal of communication*, 10, 21.
- Heeks, R. & Renken, J., (2016). *Data Justice For Development: What Would It Mean?* Manchester. Retrieved from:
<https://www.gdi.manchester.ac.uk/research/publications/di/>
- Hirsch, D. D. (2010). The law and policy of online privacy: Regulation, self-regulation, or co-regulation. *Seattle UL Rev.*, 34, 439.
- Hoffman, M. D., Blei, D. M. & Bach, F. (2010). Online Learning for Latent Dirichlet Allocation. *Advances in Neural Information Processing Systems*, 23.

- Hogan, B. (2010). The Presentation of Self in the Age of Social Media: Distinguishing Performances and Exhibitions Online. *Bulletin of Science, Technology & Society*, 30, 377–86.
- Hood, C., (1994). *Explaining economic policy reversals*. Buckingham; Philadelphia: Open University Press.
- Hosein, I., Tsiavos, P., & Whitley, E. A. (2003). Regulating architecture and architectures of regulation: Contributions from information systems. *International Review of Law, Computers & Technology*, 17(1), 85-97.
- Howison, J., Wiggins, A., & Crowston, K. (2011). Validity issues in the use of social network analysis with digital trace data. *Journal of the Association for Information Systems*, 12(12), 2.
- Hughes, T. P., & Hughes, T. P. (2004). *Human-built world: How to think about technology and culture*. University of Chicago Press.
- Humphreys, L. (2018). The qualified self: Social media and the accounting of everyday life. MIT press.
- Investopedia Team (2020, Mar 21). *How Blockchain is Revolutionizing Content Distribution*. Investopedia. Accessible at:
<https://www.investopedia.com/news/how-blockchain-revolutionizing-content-distribution/>
- Janecek, V. (2018). Ownership of personal data in the Internet of Things. *Computer law & security review*. 34(5), 1039-1052.

- Jasanoff, S. (1996). Beyond epistemology: relativism and engagement in the politics of science. *Social studies of science*, 26(2), 393-418.
- Jeong, Y., & Coyle, E. (2014). What are you worrying about on facebook and twitter? an empirical investigation of young social network site users' privacy perceptions and behaviors. *Journal of Interactive Advertising*, 14(2), 51-59.
- Jeong, Y., & Kim, Y. (2017). Privacy concerns on social networking sites: Interplay among posting types, content, and audiences. *Computers in Human Behavior*, 69, 302-310.
- Jin, S. A. A. (2013). Peeling back the multiple layers of Twitter's private disclosure onion: The roles of virtual identity discrepancy and personality traits in communication privacy management on Twitter. *New Media & Society*, 15(6), 813-833.
- Johnson, J. (2014). From open data to information justice. *Ethics and Information Technology*, 16(4), 263-274.
- Johnson, J., (2016). The Question of Information Justice. *Communications of the ACM*, 59(3), 27-29.
- Jungherr, A., Schoen, H., Posegga, O., & Jürgens, P. (2017). Digital trace data in the study of public opinion: An indicator of attention toward politics rather than political support. *Social Science Computer Review*, 35(3), 336-356.
- Kennedy, H., Poell, T., & van Dijck, J. (2015). Data and agency. *Big Data & Society*, 3, 1-2. doi:10.1177/2053951715621569

- Kroll, J. A., Barocas, S., Felten, E. W., Reidenberg, J. R., Robinson, D. G., and Yu, H. (2017). Accountable algorithms. *U. Pa. L. Rev.* 165(3). 633-705.
- Lesk, M. (2012). The price of privacy. *IEEE Security and Privacy Magazine*, 10(5), 79–81.
- Lessig, L. (1998). The new Chicago school. *The Journal of Legal Studies*, 27(S2), 661-691.
- Lessig, L. (1999). The architecture of privacy. *Vand. J. Ent. L. & Prac.*, 1, 56.
- Lessig, L. (2002). Privacy as property. *Social Research: An International Quarterly*, 69(1), 247-269.
- Lessig, L. (2006). *Code 2.0: Code and Other Laws of Cyberspace*. Retrieved from: <http://codev2.cc/download+remix/Lessig-Codev2.pdf>
- Li, H., Luo, X. R., Zhang, J., & Xu, H. (2017). Resolving the privacy paradox: Toward a cognitive appraisal and emotion approach to online privacy behaviors. *Information & management*, 54(8), 1012-1022.
- Li, Y. (2012). Theories in online information privacy research: A critical review and an integrated framework. *Decision Support Systems*, 54(1), 471-481.
- Lipton, J. D. (2010). Mapping Online Privacy. *Nw. UL Rev.*, 104, 477.
- Livingstone, S. (2004). Media literacy and the challenge of new information and communication technologies. *The Communication Review*, 7(1), 3-14.

- Livingstone, S. (2008). Taking risky opportunities in youthful content creation: Teenagers' use of social networking sites for intimacy, privacy and self-expression. *New Media & Society*, 10(3), 393-411.
- Litman, J. (2000). Information privacy/information property. *Stanford Law Review*, 1283-1313.
- Lorentz, A. (2013, April). With Big Data, Context is a Big Issue. *Wired*. Retrieved from: <https://www.wired.com/insights/2013/04/with-big-data-context-is-a-big-issue/>
- Loshin, D. (2002). Knowledge integrity: Data ownership. *As quoted in Data Ownership, US Department of Health and Human Services*. Rockville, MD: Office of Research Integrity.
- Lu, X. (2011). A corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers' language development. *Tesol Quarterly*, 45(1), 36-62.
- Malik, M. M. (2018). *Bias and beyond in digital trace data*. Doctoral dissertation, PhD thesis. Pittsburgh, PA: Carnegie Mellon University.
- McHugh, M. L. (2012). interrater reliability: the kappa statistic. *Biochemica Medica*, 22(3), 276–282.
- McLuhan, H. M. (1964). *Understanding Media*. London: Routledge & Kegan Paul.
- Moreno, M. A., & D'Angelo, J. (2019). Social media intervention design: applying an affordances framework. *Journal of Medical Internet Research*, 21(3), e11014.
- Murphy, R. S. (1996). Property Rights in Personal Information: An Economic Defense of Privacy. 84 *GEO. L.J.*, 2381- 2385.

- Naylor, B. (2017, March 28). Congress Overturns Internet Privacy Regulation. *National Public Radio*. Retrieved from:
<https://www.npr.org/2017/03/28/521831393/congress-overturms-internet-privacy-regulation>
- Newell, S., & Marabelli, M. (2015). Strategic opportunities (and challenges) of algorithmic decision-making: A call for action on the long-term societal effects of ‘datification’. *The Journal of Strategic Information Systems*, 24(1), 3-14.
- Ngai, E. W., Tao, S. S., & Moon, K. K. (2015). Social media research: Theories, constructs, and conceptual frameworks. *International journal of information management*, 35(1), 33-44.
- Nissenbaum, H. (2004). Privacy as contextual integrity. *Wash. L. Rev.*, 79, 119.
- Oracle (2015). *Oracle buys Datalogix: creates the world’s most valuable data cloud to maximize the power of digital marketing*. Retrieved from:
<http://www.oracle.com/us/corporate/acquisitions/datalogix/general-presentation-2395307.pdf>
- O’Leary, D. E. (2017). Configuring blockchain architectures for transaction information in blockchain consortiums: The case of accounting and supply chain systems. *Intelligent Systems in Accounting, Finance and Management*, 24(4), 138-147.
- Nolte, C. G. (2015, June). Personal data as payment method in SNS and users’ concerning price sensitivity-a survey. In *International Conference on Business Information Systems* (pp. 273-282). Berlin: Springer.

- Palmer, S. (2021, July 11). *Web 3.0 Things You Should Know*. Shelly Palmer. Accessible at: <https://www.shellypalmer.com/2021/07/web-3-0-things-you-should-know/>
- Panahi, S., Watson, J., & Partridge, H. (2012). Social media and tacit knowledge sharing: Developing a conceptual model. *World academy of science, engineering and technology*, (64), 1095-1102.
- Pangrazio, L., & Sefton-Green, J. (2020). The social utility of ‘data literacy’. *Learning, Media and Technology*, 45(2), 208-220
- Pangrazio, L., & Selwyn, N. (2019). ‘Personal data literacies’: A critical literacies approach to enhancing understandings of personal digital data. *New Media & Society*, 21(2), 419-437.
- Park, S., Straubhaar, J. D., & Strover, S. L. (2019). Exploring Ambivalence in Technological Embeddedness: The Role of Technological Competence and Dependence in the Information Gap. *Journal of Broadcasting & Electronic Media*, 63(3), 433-454.
- Prosser, W. L. (1960). Privacy, 48 Calif. L. Rev, 383(10.2307), 3478805383.
- Puschmann, C., & Scheffler, T. (2016) Topic modeling for media and communication research: A short primer. *HIIG Discussion Paper Series* (No. 2016-05): Alexander von Humboldt Institut für Internet und Gesellschaft.
- Quan-Haase, A., & Wellman, B. (2005, February). How Computer-Mediated Hyperconnectivity and Local Virtuality Foster Social Networks of Information

- and Coordination in a Community of Practice. In *International Sunbelt Social Network Conference*.
- Raymond, N.A., (2016). Beyond “Do No Harm” and Individual Consent: Reckoning with the Emerging Ethical Challenges of Civil Society’s Use of Data. In L. Taylor, L. Floridi, & B. van der Sloot, (Eds.). *Group Privacy: new challenges of data technologies*. Berlin: Springer.
- Rayna, T., Darlington, J., & Striukova, L. (2015). Pricing music using personal data: mutually advantageous first-degree price discrimination. *Electronic Markets*, 25(2), 139-154.
- Regulation (EU) 2016/679 (General Data Protection Regulation) (2016). Accessed October 2020: <https://gdpr-info.eu/>
- Reichman, J., & Samuelson, P. (1997). Intellectual property rights in data? *Vanderbilt Law Review*, 50(1), 49-49.
- Reimsbach-Kounatze, C. (2015), "The Proliferation of “Big Data” and Implications for Official Statistics and Statistical Agencies: A Preliminary Analysis", OECD Digital Economy Papers, No. 245, OECD Publishing, Paris, <https://doi.org/10.1787/5js7t9wqzv8-en>.
- Ruhaak, A. (2021, Feb 24) *How data trusts can protect privacy*. MIT Technology Review. Accessible at: <https://www.technologyreview.com/2021/02/24/1017801/data-trust-cybersecurity-big-tech-privacy/>

- Sambra, A. V., Mansour, E., Hawke, S., Zereba, M., Greco, N., Ghanem, A., ... & Berners-Lee, T. (2016). Solid: a platform for decentralized social applications based on linked data. MIT CSAIL & Qatar Computing Research Institute, Tech. Rep.
- Samuelson, P. (2000). Privacy as intellectual property. *Stan. L. Rev.*, 52, 5, 1125-1174.
- Sanfilippo, M., Frischmann, B., & Standburg, K. (2018). Privacy as Commons: Case Evaluation through the Governing Knowledge Commons Framework. *Journal of Information Policy*, 8, 116-166.
- Sarigol, E., Garcia, D., & Schweitzer, F. (2014, October). Online privacy as a collective phenomenon. In *Proceedings of the second ACM conference on Online social networks* (pp. 95-106). ACM.
- Sawyer, S., & Jarrahi, M.H. (2014). Sociotechnical Approaches to the Study of Information Systems. *CRC Handbook of Computing*.
- Scharf, N. (2012). Life through a lens: A "Lessigan" model for understanding digital copyright infringement? *Journal of Internet Law*, 16(1), 18.
- Scheerder, A., Van Deursen, A., & Van Dijk, J. (2017). Determinants of Internet skills, uses and outcomes. A systematic review of the second-and third-level digital divide. *Telematics and informatics*, 34(8), 1607-1624.
- Schneier, B., (2015). Data and goliath: The hidden battles to capture your data and control your world. New York: W.W. Norton & Company.
- Scholz, L. H. (2015). Privacy as Quasi-Property. *Iowa L. Rev.*, 101, 1113.

- Schwartz, P. M. (2004). Property, privacy, and personal data. *Harv. L. Rev.*, 117, 2056.
- Schwartz, P. M. (2012). Information privacy in the cloud. *U. Pa. L. Rev.*, 161, 1623.
- Shen, Y., & Pearson, S. (2011). Privacy enhancing technologies: A review. *HP Laboratories*, 2739, 1-30.
- Smith, H. J., Milberg, S. J., & Burke, S. J. (1996). Information privacy: measuring individuals' concerns about organizational practices. *MIS quarterly*, 167-196.
- Solove, D. J. (2002). Conceptualizing privacy. *Calif. L. Rev.*, 90, 1087.
- Solove, D. J. (2012). Introduction: Privacy self-management and the consent dilemma. *Harv. L. Rev.*, 126, 1880.
- Sotto, L. & Simpson, A. (2015). *Data Protection & Privacy 2015: In 31 jurisdictions worldwide*. New York, United States: Law Business Research Ltd. Retrieved from: https://www.huntonprivacyblog.com/wp-content/uploads/sites/18/2011/04/DDP2015_United_States.pdf
- Spiekermann, S., Acquisti, A., Böhme, R., & Hui, K. L. (2015). The challenges of personal data markets and privacy. *Electronic Markets*, 25(2), 161-167.
- Steem Blueprint. (2017). <https://steem.io/steem-bluepaper.pdf>. Accessed: 2021-01-15.
- Steem Whitepaper. (June 2018). <https://steem.io/steem-whitepaper.pdf>. Accessed: 2020-11-30.
- Stevens, G. M. (2012, April 10). Data security breach notification laws. *Congressional Research Service Report*. Retrieved from: <http://dev.journalistsresource.org/wp-content/uploads/2012/04/R42475.pdf>

- Stewart, K. A., & Segars, A. H. (2002). An empirical examination of the concern for information privacy instrument. *Information systems research*, 13(1), 36-49.
- Strover, S., Straubhaar, J., Chen, W., Gustafson, K., Schrubbe, A., & Popiel, P. (2014). *Digital inclusion in Austin: preliminary results, 2014 survey*. Austin, TX: The University of Texas at Austin.
- Taylor, L. (2016). The ethics of big data as a public good: which public? Whose good? *Philosophical Transactions of the Royal Society, A*, 374(20160126), 1–13.
- Terranova, T. (2000). Free labor: Producing culture for the digital economy. *Social text*, 18(2), 33-58.
- Thelwall, M. (2018). Can social news websites pay for content and curation? The SteemIt cryptocurrency model. *Journal of Information Science*, 44(6), 736-751.
- Thimmesch, A. B. (2016). Transacting in Data: Tax, Privacy, and the New Economy. *Denv. L. Rev.*, 94, 145.
- Turow, J., Hennessy, M. & Draper, N. (2015). *The Tradeoff Fallacy: How Marketers Are Misrepresenting American Consumers and Opening Them Up To Exploitation*. Retrieved from:
https://www.asc.upenn.edu/sites/default/files/TradeoffFallacy_1.pdf
- UNDP, (2021). *United Nations Development Programme: Human Development Reports*.
<http://hdr.undp.org/en/humandev>

UNESCO (2013). *Media and Information Literacy: Policy and Strategy Guidelines*.

Paris: UNESCO. Retrieved from:

<http://unesdoc.unesco.org/images/0022/002256/225606e.pdf>

Vaidhyathan, S. (2012). *The Googlization of Everything:(and why we should worry)*.

Berkeley: University of California Press.

van Deursen, A. J., & Helsper, E. J. (2015). *The third-level digital divide: Who benefits most from being online?* In Communication and information technologies annual (10), pp. 29-52. Emerald Group Publishing Limited.

<https://doi.org/10.1108/S2050-206020150000010002>

van Deursen, A. J., Helsper, E. J., & Eynon, R. (2016). Development and validation of the Internet Skills Scale (ISS). *Information, Communication & Society*, 19(6), 804-823.

van Deursen, A. J., Helsper, E., Eynon, R., & van Dijk, J. A. (2017). The compoundness and sequentiality of digital inequality. *International Journal of Communication*, 11, 452-473.

van Dijck, J. (2009). Users like you? Theorizing agency in user-generated content. *Media, culture & society*, 31(1), 41-58.

Whitt, R. S. (2019). Old School Goes Online: Exploring Fiduciary Obligations of Loyalty and Care in the Digital Platforms Era. Santa Clara High Tech. LJ, 36, 73.

Wiebe, A. (2016). Protection of industrial data—a new property right for the digital economy? *Journal of Intellectual Property Law & Practice*, 12(1), 62-71.

Wilka, R. (2018). Privacy Commitments. *Wash. L. Rev. Online*, 93, 63.

Wilks, J. & Christie, A. (2013). *Rights in data handbook: Protecting and exploiting IP in data and databases internationally*. London, United Kingdom: DLA Piper.

Retrieved from:

https://www.dlapiper.com/~media/Files/Insights/Publications/2013/01/Rights%20in%20Data%20Handbook%202013/Files/Rights_in_Data_Handbook/FileAttachment/Rights_in_Data_Handbook.pdf

Wright, R. N., Camp, L. J., Goldberg, I., Rivest, R. L., & Wood, G. (2002, March).

Privacy tradeoffs: myth or reality? In *International Conference on Financial Cryptography* (pp. 147-151). Berlin, Heidelberg: Springer.

Yang, A., (2021). *Data Dividend Project*. Accessible at:

<https://www.datadividendproject.com/>

Zuboff, S. (2015). Big other: surveillance capitalism and the prospects of an information civilization. *Journal of Information Technology*, 30, 75–89.

Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. New York: PublicAffairs.

Zuckerberg, M. (2009). An Open Letter from Facebook Founder Mark

Zuckerberg. *Zuckerberg Transcripts*. 23. Retrieved from:

http://dc.uwm.edu/zuckerberg_files_transcripts/23