A Theoretical Investigation into *Instagram* Hashtag Practices

by

Moe N. Malik

A thesis submitted to the Faculty of Graduate and Postdoctoral Affairs in partial fulfillment of the requirements for the degree of

Master of Arts

in

Communication

Carleton University
Ottawa, Ontario

ABSTRACT

Hashtagging photos on *Instagram* represents a recent iteration of user-generated content organization. This thesis examines whether established theories of classification can explain these hashtagging practices by asking the following two questions "Are early conceptions of user generated classifications useful descriptors of hashtagging practices on *Instagram*?" And "Do older classification theories, developed prior to hashtagging practices still apply in a user generated context?" The first question examines nonformal classification systems such as folksonomies while the second examines Hacking's (1986, 1996) dynamic nominalism, Bowker and Stars' (1999) case study informed definitions, and Gruber (2007) and Iliadis' (2016) perspectives on the role of ontologies in classification. A hybrid walkthrough methodology was applied in *Instagram* to empirically examine the classificatory processes of content producers for thirty-five sets of hashtagged pools of photos. Overall most of the characteristics of formal classification systems and theories apply to *Instagram* user generated content and unsurprisingly, hashtagging practices on *Instagram* can best be characterized as being folksonomic.

ACKNOWLEDGEMENTS

I want to acknowledge Tracey P. Lauriault for her patient supervision throughout this research. Thank you to Merlyna Lim for being a thoughtful second reader.

Further thanks to Sandra Robinson for encouraging me to pursue a master's education in Communications and to Chris Russill for helping to make it possible.

I am grateful for Kasia who supported me patiently from the beginning and for Willow, our cat, for keeping me company on long days of research and writing.

TABLE OF CONTENTS

List of Tables

List of Illustrations

Appendices

CHAPTER 1: INTRODUCTION

1.1 Context

Every day, there are 90 million photos and videos shared on Instagram. The second largest social media platform after Facebook, Instagram is a photo- and video-sharing social networking service, where users apply dynamic, user-generated tagging called hashtag to make it possible for others to easily find photos and videos with a specific theme. Hashtagging practices allow users to classify contents easily and informally without the need of any formal taxonomy. How do we make sense of this informal, collective, bottom-up practice of classification? How do we theorize it? Can formal classification theories, which were developed before the advent of social media, be utilized to explain this emerging practice?

We interact with classifications on daily basis through everyday practices of sorting things out which allow us to make our way through the world. Classification as the process can be overlooked as it is so commonplace. In other cases, forms often restrict the classifying of people into gender and race categories. The data collected by these forms are

determined in a top down fashion by organizations according to particular standards and enforced by those organizations as they set the parameters and the options to choose. Experts and institutions are responsible for developing the classifications and applying them, while most of the population conform, to some extent, and choose the category they most fit into.

There are formal types of classification systems and informal types. The subject of this thesis is the hashtag practice in social media, notably in Instagram, which is a type of informal classification system called a folksonomy. A folksonomy is a classification that is user-generated to categorize social media content. This is classification from the bottom rather than being dictated from the top. Rather than being developed by a select group of professionals, folksonomies are created by everyday people. Rather than dictating to others how to apply a classification is applied, users are left to their own subjective judgement. Centrally determined standards do not necessarily occur in a social media context. There are different types of formal classifications such as taxonomies (see 2.3.4) and informal classifications can be called pools as I did when it came to label a largely diversified collection of hashtagged *Instagram* photo organized chronologically by the platform.

Several theorists have defined what formal and informal classifications are. For example, formal classifications were discussed by scholars such as Ian Hacking (1986, 1996) a philosopher of science and Geoffrey Bowker and Susan Leigh Star (1999) who, as science and technology studies (STS) scholars examined how things are classified in administrative infrastructures. Folksonomies is a concept coined and developed by Thomas Vander Wal (2007) an information architect who observed the emergence of user-generated classification practices in a photo sharing social media application called *Flickr*.

He believed that practices such as tagging, and later hashtagging, represented a new type of classification practice compared to the traditional or more formal of classifications. Most formal classification theories were developed before the advent of social media. One of the motivations behind the research conducted for this thesis was to understand whether formal classification theories are valid in social media applications.

1.2 Research Questions

In this thesis I chose to undertake theory testing, which is a way to assess the relevance of past research to present realities and did so by answering the following two research questions:

1. Are early conceptions of user generated classifications useful descriptors of hashtagging practices on Instagram?
2. Do formal classification theories, developed prior to hashtagging practices apply in a user generated context?

1.3 Theoretical framework

Methodologically, I chose to study hashtagging practices not from a user-oriented perspective by instead by examining how the practice manifests on the platform. This approach aligns well with how Bowker and Star (1999) conducted their classification case studies, namely they examined how classifications are used, where they occur, and how standards inform the process of classification in a given information infrastructure. They focused on the technical details related to classification. My study of classification processes on *Instagram* interface also focused on the technical details informing hashtag-based classification. I was therefore able to compare their findings directly to *Instagram* to see if they were applicable descriptors of hashtagging. I use Bowker and Star (1999) as an

example, since theirs was a common perspective among the scholars I examined. Tables 4 through 9 in chapter 4 show the direct comparison of scholarly definitions and my observation of hashtagging processes in *Instagram*.

Here, classification is associated with infrastructure or alternatively a platform as an infrastructure because I provide the technical environment that allow for the classification practices of photos to occur. My understanding of platforms is informed by Amrit Tiwana's (2013) framework of social media ecosystems, which explains the relationships between end users, applications, interfaces, and social media platforms. This ecosystem of interoperable technologies makes social media usage, including hashtagging, possible. I also accounted for social media platforms as spaces for economic activity because social media is a business venture for companies such as *Facebook* (Iansiti and Levien, 2004). Combining these technical and economic perspectives provided a fuller understanding of platforms.

Part of the outcomes of this research was the development of new thinking about classifications and thinking differently about platforms. For example, I question whether the characteristic of mutual exclusivity is essential to classifications. Additionally, in a social media context people classify not only to organize but also to promote their content. I also later I return to the platform ecosystem model to show how classification using hashtags is a powerful asset to social media platforms because of the individual liberty it accords to users compared to formal classification systems. Furthermore, I discuss cross-platform sharing using hashtags to show that classification for promotion extends the user's content beyond their immediate social circle and into larger audiences. Which benefits those who generate content and those who provide the platforms and their integrative

functions, since this forms part revenue strategies.

1.4 Research Objectives

The objectives of this research are twofold:

1. To assess how classification theories may be used to study hastagging practices on Instagram, and
2. To generate new thinking about classifications that can capture the observed reality of classification practices on social media, particularly on Instagram.

1.5 Research Contributions

This thesis contributes to communication and media studies and classification studies in general, because it examines a new form of media interaction such as user-generated content and classification and re-examines the traditional role of classification as and organizing tool. For instance, classification can now be thought of as a business activity when considering the way users employ hashtags to promote content related to their business and the business of sorting this content to aid the platform in capitalizing on that role. An avenue for further research is the notion of agency when it comes to classification, especially as the frequency of the use of a hashtag informs a recommendation system. Another under researched area is how user generated classification might be a revenue stream for platform companies.

1.6 Thesis outline

This thesis is structured as follows. Chapter 2 first explains Tiwana's (2013) framework and the business model of social media platforms. The chapter considers the centrality of algorithms to social media use because of their ability to intermediate between humans and computers and their ability to cluster content according to hashtags. The chapter also

includes a literature review of classification theories related to the central research questions. Chapter 3 discusses the methodological approaches adopted in this thesis such as a hybridized walkthrough method based on Light, Burgess, and Duguay (2016). Chapter 4 discusses how the walkthroughs were conducted and presents the findings. Chapter 5 presents the results of the descriptive coding analysis and unexpected outcomes Chapter 6 concludes with the results and a discussion of the limitations of the project as well as suggested directions for future research.

CHAPTER 2: DISCUSSION AND LITERATURE REVIEW

2.1. Introduction

In this chapter, I introduce social media platform ecosystems and to a lesser extent discuss their capitalistic nature as well as the role of algorithms in theses platforms. In addition, I present a literature review of key classification theories which form the theoretical framework for this thesis.

2.2. The Social Media Ecosystem and the Platform Economy

Amrit Tiwana provides a useful approach to understand what a platform is in instrumental terms in *Platform Ecosystems: Aligning Architecture, Governance, and Strategy* (2013), while Zysman and Kenney's (2016) in *The Rise of the Platform Economy* situate the discussion of platforms in a business model. Amrit Tiwana's (2013) framework (figure 1 below) describes platforms generally.

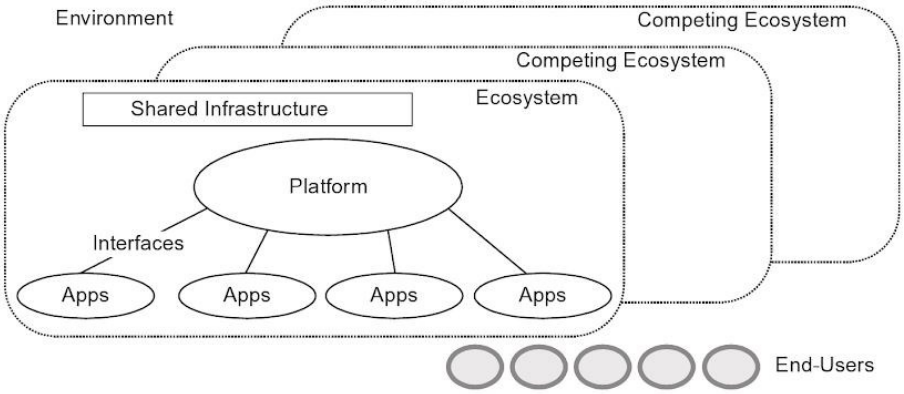


Figure 1: Tiwana's platform ecosystem framework (2013)

Tiwana characterizes platforms as a “software-based system that provides core functionality shared by the apps that interoperate with it” (2013:5). Applications are complimentary to the platform (6) and software needs to be compatible and interoperable with the app to enable access to interaction with the social media environment on the Internet.

Applications are characterized as extensions of the platform; ergo, they are software subsystems that mesh with the platform to allow interactivity (Tiwana 2013: 5-6). Although increasingly popular with laptops, applications are often synonymous with smartphones and tablets. The combination of applications and wireless internet connectivity and mobile devices afford users mobile interaction and participation in social media activities on social media platforms, whereby users mainly interact with the application. Users and producers of content in social media platforms are untethered from the sedentary activity of using a desk and laptop and tethered to their smartphone while their cell or internet service provides them a data plan or Wi-Fi. Applications also provide developers with revenue as social media platforms provide APIs and developer toolkits that enable them to develop third party applications as described in the "*Instagram* Developer

Documentation"(*Instagram*, 2018).

Tiwana (2013) notes that “apps are complimentary goods for platforms” and some are Proprietary included with a device and other apps are available for download for free or for a small fee. On an Android phone, for example, preinstalled applications include the Google suite, such as Gmail, and the Samsung suite, such as a text messaging app. Applications I have installed on my Android phone include VSCO (VSCO: play.google.com, 2018) and Storysave (Cottle, 2018), which are photo editing apps and free to use in their basic form. Buying extra filters for editing photos with VSCO costs upwards of a dollar (VSCO: play.google.com, 2018). Accessing the advanced version of Storysave means paying $2.99 (Cottle, 2018). In both cases, developers profit from the users who purchase the apps, the platform benefits by having services they did not develop added to their platform which brings them more users, prestige, and better and more content. Users benefit by having more options to manage and contribute their content. Innovation occurs among designers as the platforms change their code and APIs.

Proprietary applications that come with or are compatible with the phone’s operating system are also avenues of revenue for social media companies and their business partners. Proprietary apps allow users to access social media platforms in the way Tiwana (2013) describes in his framework of platform ecosystems. For example, the *Facebook* app affords access to the *Facebook* platform. The *Facebook* platform is free to use, but it features advertisements from other companies. Another stream of revenue is social media advertising, which exposes users to tailored ads to encourage the purchase of products from retailers such as *Amazon*. *Facebook* is an example of a proprietary app that affords access to a platform with advertisements from *Amazon*. This stream of revenue has been

challenged by the development of ad blockers developed independently to stop ads from appearing while the user.

Advertising generates revenue for the company marketing a product and for the company hosting advertisements on their platform. This model is similar to traditional print media advertising in which a newspaper generates money by including advertisements. These "host" companies are considered by Iansiti and Levien (2004) to be keystone firms. They are called thus because they are central to the health of the ecosystem "such that, in many cases, its removal will lead to the catastrophic collapse of the entire system" (Iansiti and Levien, 2004).

From an economic perspective, *Facebook* would collapse if *Facebook* the company was not central and active in revenue generation. Furthermore, *Instagram* would not generate revenue if *Facebook* as a company did not take an active role in generating revenue by dealing in advertising. Iansiti and Levien (2004) suggest that keystone firms "exercise a systemwide role despite being only a small part of their ecosystems' mass." If we think of *Instagram* as a branch of *Facebook*, it would die off economically without the active strive *Facebook* takes toward *Instagram*'s revenue generation. *Facebook* is a small part of its ecosystem but occupies a systemwide role that is crucial to ecosystem survival.

Instagram was started by Kevin Systrom and Mike Krieger to reframe their existing project *Burbn*. They chose to direct their focus to creating a mobile photography platform. Their choice of name was informal but apt, according to Systrom "an instant telegram of sorts. It also sounded camera-y" (Systrom, 2011).

In April 2012, Forbes reported that *Facebook* would buy *Instagram* for one billion dollars in cash and *Facebook* shares (Upbin, 2012). Mark Zuckerberg acquired *Instagram*

as he saw it as a natural addition to *Facebook* because of the prevalence of photo sharing on both platforms. According to Zuckerberg, buying *Instagram* was *Facebook*'s first purchase with the rationale that "providing the best photo sharing experience is one reason why so many people love *Facebook* and we knew it would be worth bringing these two companies together" (Upbin, 2012). Similarly, the purchase of *Instagram* by *Facebook* was the only time the platform changed hands (See section 1.4, Why *Instagram*?). *Facebook* as a keystone firm demonstrated the effective management strategies outlined by Iansiti and Levien (2004) by creating "value within the ecosystem" and striving "to share the value with other participants in the ecosystem." Some of the changes brought on by the acquisition include the addition of location tagging, increased photo editing capabilities, and the ability to share *Instagram* posts to *Facebook* profiles. The value, in this case, lay in affording users the most options when modifying and sharing their photos with people to whom they are connected. Creating value ensures that users are attracted to the platform and remain attracted to the platform; sharing the value ensures the longevity of a platform by balancing the lifespan of the platform and the richness of content therein (Iansiti and Levien 2004). Conveniently, this increases the market shares for *Facebook* because platform convergence creates a social media monopoly which increases the potential for capitalization.

For example, the terms and conditions specify that *Instagram* photos can be used by the company due to *Instagram*'s "non-exclusive, fully paid, and royalty-free, transferable, sub-licensable, worldwide license to use the Content that you post on or through the Service" (Terms and Conditions, 2018). Additionally, the terms of service make clear that *Instagram* collects, uses, and shares user information "across the *Facebook* products"

(Terms and Conditions, 2018). *Facebook*, *Instagram*'s keystone firm, could use visual media from *Instagram* for profit.

Competing ecosystems are the final consideration of Tiwana's (2013) framework. *Instagram* is one ecosystem among many. The same framework applies to other social media platforms such as *Facebook*, Twitter, *Snapchat*, and *Flickr*. Each has users who interact with the platform using apps and interfaces. Each ecosystem undergoes updates to improve existing affordances and to add new functionality. For example, during the data collection for this thesis *Instagram* was updated six times. The frequent updates seem to support what Tiwana (2013) and Iansiti and Levien (2004) suggest about platform competition whereby rival ecosystems lead to innovation that promotes the survival and thriving of an ecosystem (Tiwana, 2013).

Conversely, social media participation often leads to users overlooking their production and consumption practices. Participating in sharing photos, liking photos, commenting on photos, and carrying on direct messaging conversations are repetitive and predominant *Instagram* practices. I found that these practices involved me deeply as a researcher during the walkthrough as both creative production and digital labor. Zysman and Kenney (2016) critically discuss *Amazon*, *Uber*, and other forms of social media economies such as: creative economies, sharing economies, and gig economies in terms of prosumption:

> "sharing" has a more than passing resemblance to the putting-out economy that existed before factories, when companies would ship materials to people to assemble items such as shoes, clothing, or firearms in their homes. In the current manifestation of putting out, the platform operator has unprecedented control over

the compensation for and organization of work, while still claiming to be only an intermediary (Zysman and Kenney, 2016).

The authors refer to this type of piece work as part and parcel of the "'platform economy,' or 'digital platform economy,' as a more neutral term that encompasses a growing number of digitally enabled activities in business, politics, and social interaction" (Zysman and Kenney, 2016).

A platform economy perspective remains relevant in the studies of social media. *Amazon Web Services* has tailored ads to *Instagram* users since 2016. The growth and integration of the marketplace of platforms especially those that focus on user generated content platforms warrants deeper understanding, for "if the industrial revolution was organized around the factory, today's changes are organized around these digital platforms, loosely defined" (Zysman and Kenney, 2016). Studies tracing the integration of platforms and their structural changes to the economy are essential to understanding current business practices, the nature of "work" and the flow of capital. Similarly, Zysman and Kenney (2016) are correct when they suggest that the integration forms a new ecosystem.

Tiwana (2013) and Zysman and Kenney (2016) provided strong insight into the back-end workings of social media as a platform ecosystem and social media as an economic system.

2.3. Algorithms: Human Computer Enmeshing

Algorithms function as dynamic intermediaries between humans and computers and platforms rely on algorithms to sort content. Johnathan Roberge and Robert Sefert (2016) in *What are Algorithmic Cultures?* note Robert Kowalski's (1979) discussion about the components and goals of an algorithm. N. Katherine Hayles (2000) discusses human-

computer enmeshing, while Donald MacKenzie (2006) considers algorithms to be engines central to the process of user generated classifications.

Although the focus of this thesis examines surface activities of platforms such as hashtagging, it is nonetheless necessary to acknowledge the back-end processes that inform the Instagram interface. For platforms to do their work of sorting, managing accounts and assessing preferences, and to store these vast troves of data algorithms are required. In their work on algorithmic cultures, sociologists like Roberge and Seyfert (2016) and Kowalski (1979 in Roberge and Seyfert, 2016) define algorithms as "algorithm= logic + control" to signify that they are a combination of mathematical formulas and descriptive language coupled with the intent to control. Computing logic is used to instruct a computer and the latter is the application of that language to a subject. Hashtagging can be understood as a human-readable label which is a word or some text that humans can read, understand and associate with something; a hashtag symbol is part of the label, and the label with the hashtag is recognized by a software program on a computer as a classification that collects similarly hashtagged objects. The software includes an algorithm that reads the label and clusters like tagged images together. Algorithms are the intermediary that enables human-computer interaction as well as relatively new machine learning processes such as Automatic Image Annotation (AII) techniques as discussed by Giannoulakis and Tsapatsoulis, (2016).

Some argue that algorithms are the site of power in online activities. Alexander Galloway (2012) notes that "the point of power today resides in networks, computers, algorithms, information, and data" (in Roberge and Sefert, 2016). Each component is significant to the user experience of an online ecosystem, but algorithms are arguably

central to the functionality of social media ecosystems. A computer cannot do anything unless it is instructed in language that interoperates between its software and the human user. The two are enmeshed because of that dynamic (Hayles, 2000). Similarly, networks, linkages of computers that can communicate with each other, would not exist without language that interoperates between users and computers as well as between computers. Algorithm's driving power behind user generated classification supports Donald MacKenzie's (2006) claim that algorithms are engines (in Roberge and Seyfert 2016). It is important to acknowledge that algorithms ensure interoperability across networks; however, the focus of this project requires the more specific acknowledgment that algorithms drive user-directed data classification practices. That said, it is important too to remember that human and non-human contribution is key to algorithm functionality (Roberge and Seyfert 2016). Now that I have outlined algorithms and human-computer enmeshing, I have explained platforms and their business models I will proceed to the literature review of classifications.

2.4. Classifications

2.4.1. Overview

I approach formal classifications from the infrastructure case study-based theorizations of Geoffrey Bowker and Susan Leigh Star (1999), and Martha Lampland and Susan Leigh Star (2009). I then discuss Ian Hacking's (1996, 2006), dynamic nominalist approach to formal classifications. To discuss user generated classification systems, I refer to Thomas Vander Wal (2007) initial theorization of folksonomies as well as related scholars such as Scott Golder and Bernardo Huberman (2006). Finally, I consider whether ontologies

underlie folksonomies because of their centrality to formal classification systems, and to the web in general.

2.4.2. Where are Classifications Used?

Bowker and Star (1999) argue that classification systems are found anywhere there are objects, people, or data to sort into logical arrangements. Furthermore, Bowker and Star (1999) suggest that there are shared characteristics associated with infrastructures in which classifications systems are embedded. The characteristics are broad on purpose because there are many different classification systems in different locations, but the authors suggest that the locations in which they are embedded share common essential characteristics. As will be seen later, these were used to inform part of the data analysis especially when I suggest that hashtagging is a classification practice embedded in *Instagram*.

Infrastructure according to Bowker and Star (1999) is "a historical process of the development of many tools, arranged for a wide variety of users, and made to work in concert" which also aptly describes *Instagram* as a platform with user affordances (Bowker and Star, 1999). *Instagram* (2010) as a platform is a relatively recent iteration of social media compared to older platforms such as *Facebook* (2004) and *Flickr* (2004). Some of the tools developed in *Flickr*, such as a photo uploading interface and photo tagging were taken up by *Instagram*. Tagging, for example, became hashtagging thereby representing a historical development of user tools. When uploading photos in through the Flickr interface, users chose or created their tags. Likewise, users choose or create their hashtags when uploading content at the interface of *Instagram*.

Bowker and Star (1999) observed several characteristics associated with the process of categorization, such as multifunctionality, standardization and mutual exclusivity. For example, the characteristic of “multifunctionality, supporting a wide set of agendas” in an infrastructure applies to *Instagram* as a platform (Bowker and Star, 1999). Some users of *Instagram* are business people who use the platform to promote their products and drive the sales of their products on other market platforms such as *Etsy* (a platform in which Do it yourself (DIY) crafters sell their products). Other users are travelers who use the platform to document their journeys. Yet others are musicians who use the platform to interact with fans and document their tours.

Infrastructure usage is another characteristic, where usage of an infrastructure is learned as part of membership and "associated with communities of practice" which well describes *Instagram* users as well as hashtagging (Bowker and Star, 1999). A user is a member because they sign up for the service and accept the *Terms and Conditions* (Terms, 2018) of the platform. Members are referred to as the user in the document and by signing up consent to follow a contract. The terms and conditions set the limits and rules of what users can do within the platform. Members of the platform learn how to use it by experience and by observation. Experience teaches members how to choose images from their phone gallery and proceed through the uploading process. The more a user practices uploading, the better they are at making selections, captioning images, and hashtagging their images. General platform usage is learned by members, including the practice of hashtagging, over time.

The characteristic of "ready to hand" describes the *Instagram* interface as well as hashtagging (Bowker and Star, 1999). From a personal perspective, the *Instagram*

application is literally at hand because it is located on my smartphone which is usually near me. To make the application closer to hand, I placed it on the home screen of my phone. I can access it immediately after I unlock the phone. Furthermore, hashtags are ready to hand during the upload process because the platform suggests hashtags to users during the upload process.

The characteristic of "embedded in other structures" is related to the interoperability of platforms as discussed above with platform economies and below in the discussion of ontologies and folksonomies. Many of the same hashtags on *Instagram* are embedded in other social media applications or structures such as *Twitter* and *Tumblr* allowing users to share their hashtagged content across the three platforms. Whether *Instagram* is embedded within another structure is open to debate because the authors do not define "structure." From a business perspective, this is true because *Facebook* owns *Instagram* (see 2.1) and one can argue that *Instagram* is embedded in the business structure of *Facebook*. From a technical perspective, *Instagram* is not situated within the technical structure of *Facebook*. Instead the two are made interoperable through the affordance of sharing between the two platforms. Rather than *Instagram* being nested in *Facebook*, the platforms are bridged as shown in Tiwana's (2013) framework.

2.4.3. What are Classifications?

This section sets out the broad definition of a classification system for the purposes of this research. I refer to classification characteristics provided in *Sorting Things Out* (Bowker and Star, 1999) and *Standards and their Stories* (Lampland and Star, 2009). I will examine literature about classifications to develop an increasingly specific and detailed characterization of the hashtagging as a classification (e.g. Bowker and Star (1999), a

classification that is a constructive process (e.g. Hacking, 1986; 1996), and finally a folksonomy (e.g. Vander Wal, 2007). I will conclude with an examination of ontologies to see if folksonomies are underpinned by ontological structures.

Crucially, Bowker and Star (1999) note that classification systems are everyday sorting methods that organize items into a common-sense framework such a sorting things that are to each other (59). For example, people using their kitchens tend to organize the items in their cabinets according to likeness, that is, plates with plates, bowls with bowls, glasses with glasses, and mugs with mugs. Although the items themselves may vary as individual entities, they are similar enough essentially to be categorized together. Furthermore, the example highlights another characteristic of classification system: they are often overlooked because they are so mundane that we use them without second thought. We practice the routine acts of classifying unaware of what we are doing until it is called to our attention (Bowker and Star, 1999).

Lampland and Star (2009) consider standards to be "phenomena worthy of study in their own right" (4). The standardization of data is a phenomenon rather than a static application because standards are applied but are always open to redesign by their developers (Lampland and Star, 2009, 182). Although they are never fully static, it is possible to outline a few general characteristics of standards as hallmarks that signal their presence when looking at classification practices. According to Lampland and Star (2009:5), standards:

1. "Are nested inside one another."
2. "Are distributed unevenly across the sociocultural landscape."
3. "Are relative to communities of practice."
4. "Codify, embody, or prescribe ethics and values" that influence individuals.
5. Are "linked to and integrated with one another" across technical systems.

Nesting refers to the standards fitting into each other, smaller into larger. The authors begin with the standard of having a telephone, a small standard, and follow the nesting process through which making an appointment to have taxes prepared leads to the large standard of paying federal taxes (Lampland and Star, 2009). Normative cultural expectations are prevailing standards that people are expected to meet—what most people take as "a given" or part of one's normalcy. Of course, not all people can meet normative cultural standards, Lampland and Star (2009) address in their next point.

Uneven distribution can be related to the question of "who can afford to meet the standard?" Owning a smartphone is a standard but meeting that standard requires the capital to buy a phone and a plan that includes calling, texting, and data usage. The latter are three small standards nested into the standard of owning a phone. It is faulty to assume all people have the money to meet the standards either partially or fully.

Lampland and Star (2009) note that "what is benignly standard for one person at one time may be a barrier, or even a life-threatening occurrence for another." (7) If a person does not have the money to own a phone and pay for a phone plan, then their chances of employment drop because the person cannot meet other job-related standards such as being able to be reached at any time (i.e. being constantly 'on.').

Relativity to communities of practice relate to the codification of ethics and values because standards cannot encompass everyone. Those who are encompassed within the standard are supposedly different from those who are not encompassed because the latter are resigned to the "other" category, which is a catch-all for people who do not belong in the more specific traditional standard categories such as "male/female" on forms (Lampland and Star, 2009). Standards are attainable by some and unattainable by others, enable access to some and forbid access to others, confer upon some inclusion and on others discrimination.

The integration of standards into systems enables aspects of one system to speak to another. Lampland and Star (2009) use the example of reading an email which requires access to the internet through an ISP, protocols, email software to read the message, machine code to present information on the computer's interface, etc. All the standards are entities related to their specific concern, such as providing internet access, but they speak to each other to afford the practice of reading and replying to an email.

I have outlined classification systems and standards generally, so now I will articulate classifications further by examining Ian Hacking's (1986, 1996) dynamic nominalism. A scholar of the history and philosophy of science, Hacking (1986, 1996) provides a strong framework to understand the production, implementation, and results of classifications.

2.4.4. Classification Process and Effects: Dynamic Nominalism

Dynamic nominalism as developed by Ian Hacking (1986, 1996) interrogates static nominalism, outlines the systematic production of classifications, and examines the effects classifications have on classified subjects. Hacking (1986) understands static nominalism as the belief that categories are "given by human beings rather than by nature and that these

categories are essentially fixed throughout the several eras of humankind" (165). That humans themselves think of and apply categories onto their reality to make sense of it appears to make sense. Early in the development of science, the cabinet of curiosities was a precursor to the public museums of today. I conducted research to illustrate the historical shift of classifications suggested by Hacking (1986) in his framework. Figure 2, below, shows a typical cabinet of curiosities, "Musei Wormiani Historia," owned by the aristocrat-naturalist Ole Worm (1655) as depicted in the frontispiece of Museum Wormianum. Looking closely at figure 2, boxes have labels such as "animalium partes" and "conchiliata." With the former housing miscellaneous animal parts and the latter large conch shells. Hacking (1986) wants to reconsider classifications, for example, Ole Worm (1655) did not apply the labels by his own "independent" thinking, but instead nature informed his choice of labels. Conches determined the label "conchiliata" not the other way around, thus setting the stage for a particular kind of classification that was taken up in subsequent sciences. Secondly, Hacking (1986) reconsiders the fixity of classifications throughout human history.

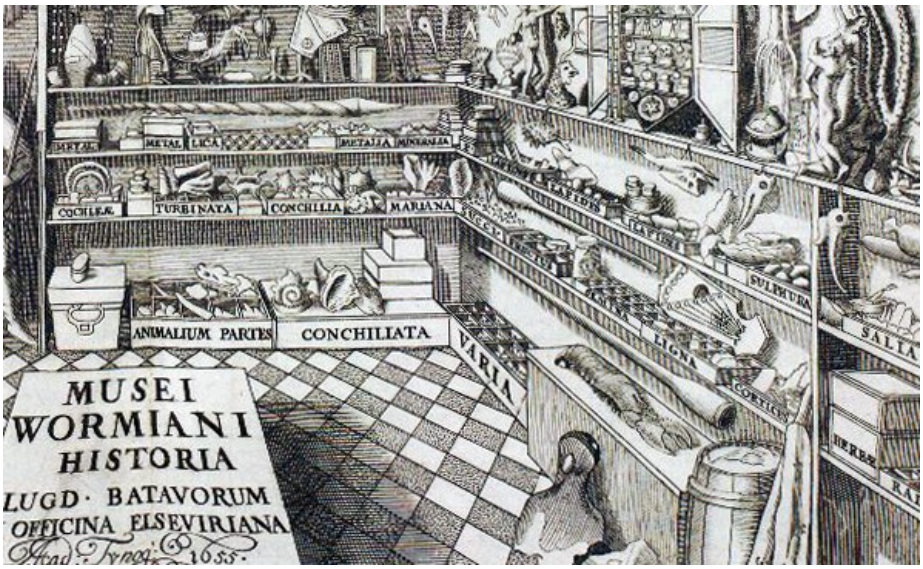
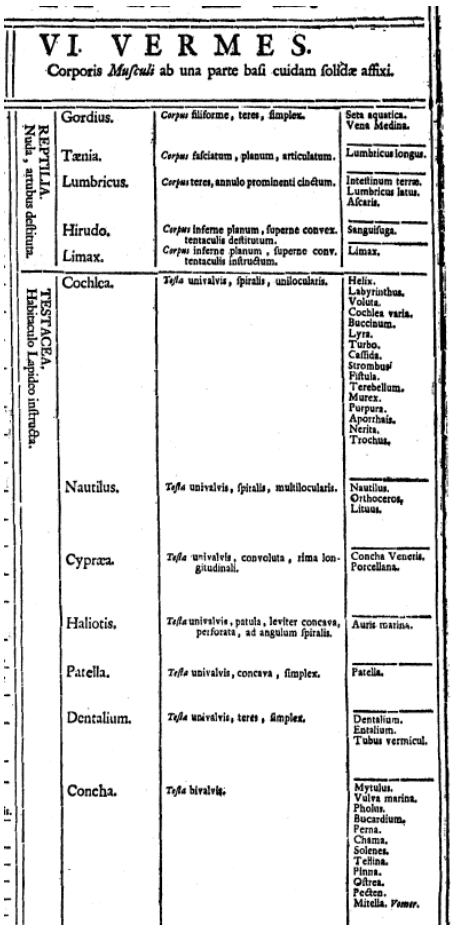


Figure 2: Detail of *Musei Wormiani Historia* showing "animalium partes" and "conchiliata"

That is, as knowledge about previously unknown subjects is accumulated from the study of nature existing categories evolve to become more accurate and new categories are added if needed. This is similar to the paradigm shift as theorized by Thomas Kuhn (1970). In the 1600s, the category "animalium partes" was necessary to gather the assortment of animal parts about which little was known except that they were from animals. As knowledge about those animals was collected, the unknown parts were categorized under increasingly specific labels. We can observe the historical change from "conchiliata" in Worm's 1600s to Carl Linnaeus' taxonomy in the 1700s.



VI. VERMES.

Corporis *Musculi* ab una parte basi cuidam solidæ affixi.

REPTILIA. Nudi, utrinque desituri.	Gordius.	*Corpus* filiforme, teres, simplex.	Seta aquatica. Vena Medina.
	Tænia.	*Corpus* fasciatum, planum, articulatum.	Lumbricus longus.
	Lumbricus.	*Corpus* teres, annulo prominenti cinctum.	Intestinum terræ. Lumbricus latus. Ascaris.
	Hirudo.	*Corpus* inferne planum, superne convex. tentaculis destitutum.	Sanguisuga.
	Limax.	*Corpus* inferne planum, superne conv. tentaculis instructum.	Limax.
TESTACEA. Habitaculo Lapideo instructa.	Cochlea.	*Testa* univalvis, spiralis, unilocularis.	Helix. Labyrinthus. Voluta. Cochlea varia. Buccinum. Lyra. Turbo. Caffida. Strombus. Fistula. Terebellum. Murex. Purpura. Aporrhais. Nerita. Trochus.
	Nautilus.	*Testa* univalvis, spiralis, multilocularis.	Nautilus. Orthoceros. Lituus.
	Cypræa.	*Testa* univalvis, convoluta, rima longitudinali.	Concha Venetis. Porcellana.
	Haliotis.	*Testa* univalvis, patula, leviter concava, perforata, ad angulum spiralis.	Auris marina.
	Patella.	*Testa* univalvis, concava, simplex.	Patella.
	Dentalium.	*Testa* univalvis, teres, simplex.	Dentalium. Entalium. Tubus vermicul.
	Concha.	*Testa* bivalvis.	Mytulus. Vulva marina. Pholas. Bucardium. Perna. Chama. Solenes. Telina. Pinna. Ostrea. Pecten. Mitella. *Vomer.*

Figure 3: Linnaeus' taxonomic classification of 1735

The acquisition of knowledge from nature shows how a box of conches called generally "conchiliata" becomes increasingly specific until the general classification of conches becomes subdivided and assembles specific types into families, subfamilies, and individuals (see figure 3). Linnaeus (1735) introduced the general animal kingdoms on which subsequent scientific classifications would be based, with some revision: mammals, birds, amphibians, fish, insects, and invertebrates. Looking closely under "vermes" (i.e. invertebrates), there is a family "testacea" and a subfamily "concha," and an assembly of individuals (Linnaeus, 1735). Presently, the classification is further complicated in its subdivisions; but the comparison of categorization between Worm (1655) and Linnaeus (1735), shows that Hacking (1986, 1996) is correct to assert that categories shift across

human history and to assert that nature is the source of categories rather than the human mind because natural entities precede their human-made classification (165). Instead of nominalism, Hacking (1986, 1996) proposes dynamic nominalism whereby kinds of subjects come into being as that kind is developed by the classification and "the category and the people in it emerged hand in hand' (165). This discussion of dynamic nominalism's key propositions sets the stage to understand the framework of dynamic nominalism.

Dynamic nominalism is an interaction between five components, driving engines, and resultant engines. The initial five components are: classification, object studied, institutions, knowledge, and experts (Hacking, 1996). Hacking (1996) originally proposed the following engines: correlation, medicalization, biologization, geneticizing, normalization, and bureaucratization, and resistance. Hacking (1986, 1996) examined the classification of people in medical science, so correlation, medicalization, biologization, geneticizing, normalization described the process. In studies of the classification of non-human entities, these engines do not apply. Medicalization for Hacking refers to molding deviant people into normal people, biologizing and geneticizing relieve deviant people themselves of the fault of being deviants by associating their deviation to a biological cause (Lauriault, 2012). When examining the classification of things or spaces, these engines needed modification. Lauriault (2012) looked at cartography and the construction of spaces and modified these engines. Instead of medicalization, she developed taking action (Lauriault, 2012). Instead of biologization and geneticizing she developed scientification (Lauriault, 2012). Taking action represents the isolation of a problem, development of a solution, and the application of the solution to the perceived problem. Scientification is the

body of knowledge mobilized to achieve the course of taking action. In this sense, Lauriault (2012) notes that scientification determines "a new course of action."

Classifications are more than labels and categories, they are "also the people classified." (Hacking, 2006, 3). The result is "a new kind of person, conceived of and experienced as a way to be a person" (Hacking, 2006, 2). A typology of people makes up those people in a particular way, so that the classification informs the way the classified people are perceived. In this research, I approach hashtags as a classification, but I do not disagree with Hacking (1996). I understand a hashtag as a category and a typology that brings a kind of entity into being. For example, the hashtag #selfie is a label applied to images, it is also a category in the form of a pool of images sharing that characteristic, which brings into being a particular type of image: portraits of the user with positive connotations (see figure 4 below).

The object is a material entity to be studied according to the original framework of dynamic nominalism. "We think of many kinds of people as objects of scientific inquiry," writes Hacking (1996), but it is troublesome to use the term "object of study" if studying classifications of people because people are not objects. Perhaps it is more humanizing to refer to people as material entities, that is, physical beings existing in the world. Material entities also accommodates actual objects such as trees, bones, and photographed subjects. In this research, the material entity is a photograph because it is the object classified. Although not studied in the strict sense, the user does consider their image before they choose to apply hashtag(s) to it.

Institutions in Hacking's (1996) framework include academic societies such as the International Society for the Study of Multiple Personality and Dissociation (4-5). These

are societies in which classifications are applied to entities and made formal, that is, official. Furthermore, institutions are the spaces in which classifications are refined and from which they are disseminated into the greater society (Hacking, 1996). In this thesis, I consider the *Instagram* platform to be the institution because it is the site in which users apply classifications to objects, hashtags to photos, publicly. Furthermore, the platform is the site of hashtag formalization. Formalization of hashtags is a matter of volume because thousands of users contribute thousands of images to popular hashtags. The same hashtags are suggested to users by the platform when they are about to post an image. The cycle of contribution and suggestion solidifies the hashtag as a formal classification.

Hacking (1996) suggests two kinds of knowledge for his purposes: expert knowledge held by professionals and popular knowledge "that is shared by a significant part of the interested population" (5). An interested population is a group of people who share common knowledge about a shared interest. For instance, amateur astronomers are a population interested in astronomy. Some of the knowledge they share includes measurements of altitude and longitude when using telescopes. Expert knowledge is specialized so that it is unlikely the amateurs know it or need it, but almost certain that the professionals know it because they need it. An example of expert knowledge is the spectrum of stars. Although useful to the experts, it is not relevant to amateurs because amateurs are interested in stargazing rather than determining the elemental composition of stars.

Most sciences can be divided into professional and amateur pursuits: astronomy, biology, paleontology, etc. Nevertheless, the two can interact in the arrangements of citizen science. For example, *Agent Exoplanet* is an initiative with professionals engaged in

exoplanet research informed by extra scientific data contributions from amateur astronomers. Professionals provide a research framework and instructions to the amateurs in return for the knowledge accumulated by the amateur's extra pairs of eyes and equipment (Haklay, 2013). The symbiotic exchange of structure for amateurs and data for scientists accomplishes more than professional scientists could by themselves. *Instagram* presents an interesting challenge to Hacking's (1996) conception of knowledge. Social media usage is not a scientific research process but a form of social engagement, but knowledge and classifications are nonetheless created.

Scientific professionals are certainly not the only type of expert. They are perhaps traditional experts because they are accredited (i.e. formalized) by an institution relevant to the field in which they work. Accreditation is signified by prefixes and suffixes such as Dr. to signify a doctorate in a field. These signifiers reflect a formal accreditation given to an individual by their institution for knowledge contributed to its field. Experts in the sciences are the focus of Hacking's (1996) framework. For this research, Hacking's (1996) idea of expert and popular knowledge is a division that does not conflate with observed hashtagging practices on *Instagram*. Instead, hashtag usage is popular knowledge mastered and applied by a large population of users. Users of hashtags constitute a type of professional in the non-traditional sense.

Experts, in Hacking's (1996) framework, are those who actualize the knowledge to classify from within institutions thereby studying, helping, or advising on the control of the classified people (5). Experts possess professional knowledge of a subject and apply it to the creation, application, and formalization of classifications. Furthermore, they also disseminate the classification thereby bringing into being new kinds of people and

reinforcing those new kinds. I suggest that all the users of hashtags on *Instagram* are experts. All hashtag users possess the knowledge of hashtagging practices (i.e. typing # and a word based on the subject represented in the image). Users of hashtags actualize the knowledge of the practice to the creation of hashtags in the first place, applying it to images, and formalizing the hashtag due to the cycle of user contribution and platform suggestion mentioned above. In this sense, the experts in this case are not a small body of people such as the International Society for the Study of Multiple Personality and Dissociation, but a large percentage of the nearly one billion users of *Instagram* ("*Instagram* monthly active users 2018, Statistic," 2018).

In summary, dynamic nominalism is the cycle by which classifications are applied to subjects from within institutions by experts acting upon their knowledge of the subjects. Consequently, a new kind of person or thing is brought into being by the classification (Hacking, 2006, 2). Although relevant to the study of sciences, the framework with few modifications applies to the understanding of hashtagging on *Instagram*. Hashtags are applied to photos from within the *Instagram* platform by users acting upon their knowledge of the subject represented in the photo and their knowledge of how to use hashtags. Hashtags are classifications, photographs are the entity to classify, *Instagram* is the institution, users are experts, and the knowledge is common to most users thereby undermining a professional-amateur division. Instead of a new kind of person being brought into being, a new kind of entity is brought into being. People took photos of themselves in the past using analogue cameras, but they were called auto-portraits or self-

portraits not selfies. To illustrate this, I have one of Vivian Maier's self-portrait from 1953 (Maloof, 2018) and a selfie taken by Tina Sosna (2017) (see figure 4).

Figure 4: Vivian Maier's self-portrait (left) and Tina Sosna's selfie (right)

The self-portrait and the selfie are not the same type of image. To be brief, the Maier (Maloof, 2018) image centers on documenting herself in her surroundings while Sosna's (2017) image reaches out to the viewer with a smile and added text. They are related but ultimately reflect different values and approaches to self-representation in photography. The selfie is an example of a comparatively new image brought into being through *Instagram* by a process similar to dynamic nominalism.

This concludes the discussion of the components of the dynamic nominalism framework. I explained how Hacking (1986, 1996) defines the components and I related them to hashtag usage on *Instagram*, and I will now proceed to explain the engines of discovery and the derived engines.

Hacking (1996) addresses engines of discovery and three derived engines. Engines of discovery drive the process of classification while derived engines are the results of classifications when they are applied to people. Counting is the first engine of discovery whereby people are subject to censuses. Hacking (1996) gives the example of autistic children in London, England and the result of the initial census being that 4.5 per 10,000 children in that city were autistic (11). In this research, counts are the body of entities in a hashtag (see figure 5). These are represented numerically when a user enters the final stage of uploading and chooses hashtags to use. For example, #cozy contains 7,257, 475 individual images. Quantifying is the next engine and addresses what the numbers mean.

Figure 5: Entity qualifications appear beside suggested hashtags during an *Instagram* upload

Quantifying is not the same as counting. Hacking (1996) refers to this as quantity and uses the example of the body mass index (BMI) to show a measure derived from the numerical representation of body fat in relation to height (10). A BMI number of 25, for example, is a measure of being overweight (Hacking, 1996, 10). In *Instagram*, the quantity of images in a hashtag influences its suggestion to users. A hashtag with 7,257, 475

individual images is more popular than one with 509,665 individual images. The numerical value is a measure of popularity in this case. The most popular hashtags appear in descending order when a user is uploading their image. In walking through the hashtagging process, I as the researcher readily chose the most popular hashtag in each case. High user contribution means a larger number of images and viewers thereby ensuring the platform suggests the most populated hashtags first and consistently.

Creation of norms follows quantification, such as "'the normal range' for the Body Mass Index" (Hacking, 2006, 11). BMI determines if a person is underweight, overweight, or within the normal range. A person is expected be within the normal range rather than a deviant. A norm is a standardized expectation analogous to standards as discussed above (i.e. Lampland and Star, 2009). Perhaps choosing the most popular hashtag is a normative behavior because it is normative expectation that that users on *Instagram* want their images seen by the greatest number of people.

Hacking (1996) approaches correlation as the relationship of one measure to another such as a BMI number of 30 (obese) with life expectancy. In this research the correlation might be of the number of individual images to popularity, so that a hashtag with 7,257,475 individual images is more popular and one with 509,665 individual images is less popular. They are both popular, but one is more popular than the other. Ergo, one is more desirable than the other if the normative expectation is the pursuit of visibility. With *Instagram*, correlation seems to be implied by the order in which the platform suggests hashtags to the user based on the number of individual images they contain.

The next three engines are medicalizing, biologizing, geneticizing. These three engines reflect what happens after counts are made, measures determined, norms

determined, and correlation between measures established. In Hacking's (1996) examination of medical science, these refer to understanding the causes of deviance from the norm and submitting deviant people to medicalization to bring them in line with the normal range. Biologizing and geneticizing refer to finding the source of the deviancy in the person's body itself or through the person's genetic history (Hacking, 2006, 11). Regarding this research, there does not seem to be an equivalent to the correction of deviation Hacking (1996) details. However, it is possible to be a deviant hashtag user. One example is using a hashtag to classify images improperly whereby a user can label an entity with a hashtag that has no relationship to it. Medicalizing, biologizing, geneticizing are analogous to acting against the deviant person to scientifically affirm that something deviant is in fact medically or scientifically normal. In the case of *Instagram* there does not appear to be any action taken against people who mis-classify using hashtags.

Normalization is one of the derived engines and is the act of trying to make deviants normal (Hacking, 2006, 11). As mentioned above, I did not observe corrective behavior between *Instagram* users in this research. Another derived engine is bureaucratization, which is a "system [that] sees itself as an objective way to determine who needs help" (Hacking, 2006, 11). Analogous to this system within *Instagram* could be the machinery that counts images and suggests hashtags because this helps users choose their hashtags. Hacking (2006) goes on to say that there is a feedback effect whereby the criteria used by the system define the meaning of categories (12). Bureaucratization accurately describes the cycle in which users contribute to the hashtags with the largest number of images because they are popular hashtags, further ensuring that the hashtag is suggested first and remains a popular, ergo desirable, choice for the user. In this case, the computational

elements of the platform might be part of a bureaucratization process.

The last derived engine is resistance, which Hacking (2016) characterizes as "kinds of people who are medicalized, normalized, administered, increasingly try[ing] to take back control from the experts and the institutions, sometimes by creating new experts, new institutions" (12). This engine does not apply to this research in the way Hacking (2016) theorizes because hashtagging images is the application of a classification to images rather than people. The medicalization, normalization, and administration of people can be damaging; so those groups of people will attempt to take back control from the experts and institutions responsible.

Although, behaviors resembling medicalizing, biologizing, geneticizing, and normalization as actions against deviancy were not found in this research, it is possible to resist as a user of *Instagram*. Two possible ways to resist are: deleting one's profile or turning one's profile setting from public to private. Deleting a profile removes the user from the institutional space altogether meaning that they are not bound by the terms and conditions of the company, the technical affordances of the platform, and the community's practices of hashtagging. The other way to resist is to remove oneself from the community of practice by making the user profile private because is not visible to the public and images from the user are not shared in the hashtag pools.

In summary, the engines of discovery and the derived engines are a framework for understanding how classification occurs from numerical representation, to the creation of measures, the administration of people, and their resistance to administration. Overall, Hacking's (1986, 2006) dynamic nominalism explains the production of classificatory knowledge.

In the next section, I will change focus and examine folksonomies to discuss user-generated classifications.

2.4.5. Folksonomies

Folksonomies were theorized by Thomas Vander Wal (2007) when he observed users of *Flickr* tagging their photos and wanted to conceptualize the practice. He accomplished this by writing a short blog post that is often cited as the coinage and definition of folksonomy in subsequent social media literature. Vander Wal (2007) did not think that *Flickr* tagging was subdivided enough to be a formal classification system because images were grouped into a pool rather than a compartmentalized framework such as the scientific taxonomy developed by Carl Linnaeus (1735) (see figure 3). He coined the term folksonomy to reflect a user generated taxonomy (Vander Wal, 2007). Although Vander Wal (2007) did not elaborate on why he considered tagging a form of taxonomy, folksonomy has been adopted by scholars in their discussion of tagging and hashtagging practices on social media platforms.

Defined by Vander Wal (2007), a folksonomy is the tagging of "information and objects (anything with a URL) done for one's own retrieval" in a social environment ("shared and open to others") by the "consumer of the information." Two modifications are necessary to relate this part of the definition to *Instagram*. First is that hashtagging does not enable retrieval for the user. Instead the hashtagged photo is also collected into a pool alongside others with the same hashtag, so other users will see photo. The uploader does not need to waste time looking through the photo pool because they are able to see all the photographs they have uploaded on their public profile. Secondly, the user is not a consumer of information only. Users produce information by taking and uploading photos,

but users also engage with other user's photos ergo consuming information. Rather than consumers only, users of *Instagram* are prosumers. Prosumer was coined by Alvin Toffler (1980) to reflect the blurring of the consumer and producer roles in a society saturated with products so highly customized that the consumers would become part of the production process. The diversity of images uploaded onto *Instagram* can be described as a highly customized product, one that reflects the user uniquely, that is consumed by other users each of whom is engaged in the production of their own unique product (i.e. a personal brand is sometimes used to refer to the overriding style of a user's images).

Vander Wal (2007) defined the tag itself as language derived from the user's personal understanding of their information or object. Certainly, users may think alike and tag a photo of a tree "tree" as I did below in figure 6.

20 May, 2018. Summer is finally here.
😌
#summer #tree #home #cozy

Figure 6: Associating "tree" to the image representing a tree

The common word to choose is tree because it is the main subject of the image in figure 6 and represents in visual form what most people agree is a tree (they recognize it as tree-like). The word "tree" represents a common sense understanding to which most people agree the object conforms to. Choosing "home" in figure 6 is probably not a common sense understanding on which most people agree because no obvious sense of home or feeling of homeliness is reflected in the image. However, I chose the word "home" to reflect the fact that I took the image from inside my home and that I felt at home sitting next to the window sill looking at the tree (the latter informing my choice of the word "cozy"). Words

such as "home" and "cozy" are used according to the subjectively informed choice of the individual. User language associations are the largely shared knowledge actualized by the community of practice in the institutional space of the *Instagram* platform. From Vander Wal's (2007) label generation perspective, an *Instagram* hashtag is as much a folksonomy as a *Flickr* tag.

Vander Wal (2007) identifies four data points associated with personal tagging: a folksonomy tool, a person tagging, the entity tagged, and the tag itself. The tag itself is a word such as "tree" or "home." The entity in this case is a user's photo as seen with Bowker and Star (1999) and Hacking (1986, 196). The user is the person associating the word to their image. The folksonomy tool differs between *Instagram* and *Flickr* because the latter relies on words alone to tag, but hashtagging requires a word and the hashmark (i.e. #). A user must type the hashmark before associating a word to their image. The tag itself is the word chosen by the user.

Vander Wal (2007) notes a social component to folksonomies that builds a community of practice: "if you know the object and the tag you can find other individuals who use the same tag on that object, which may lead (if a little more investigation) to somebody who has the same interest and vocabulary as you do." Again, the author does not define his terms such as "object," so this leaves them open to application to *Instagram* photos. The hashtag can be chosen by the user. Or, the hashtag can be suggested to the user when they are about to hashtag their image. Suggestions show the number of images in the hashtags, so it can make the user curious as to what else is in the hashtag besides their contribution. The investigation Vander Wal (2007) mentions is scrolling through the contents of a hashtag pool, viewing images that appeal, and following those images to the

profile of the poster. This has been the experience of this researcher since I joined the platform and wanted to find other users with similar interests. It is a time-consuming investigation but one that can build lasting relationships with other users. This speaks to Vander Wal's (2007) suggestion that the vocabulary of the hashtag and the subject of the hashtagged photograph can be a point of connection between users. Consequently, a community of practice emerges around certain hashtags and interests.

With modifications, Vander Wal's (2007) theory of folksonomy describes *Instagram* hashtagging practices despite being developed by the author to explain *Flickr* tagging practices. The generally open-ended framework allowed for the relation to this concept to *Instagram*. Subsequent researchers have refined the characteristics of folksonomies, and I will discuss their contributions to Vander Wal's (2007) initial folksonomy theory in the next section. For these purposes, I refer to Sheung-On Choy and Andrew Lui (2006), Tsutomu Ohkura, Yoji Kiyota, and Hiroshi Nakagawa (2006), and Scott Golder and Bernardo Huberman (2006).

Vander Wal (2007) considered the instrumental use of folksonomies but not the outcome of folksonomies. Choy and Lui (2006) suggest that folksonomies lead "to an emergent categorization of web resources in terms of tags and creates a different kind of web directory." The authors make a few useful contributions. First is that folksonomy usage develops a categorization that emerges over time. Over time content is hashtagged and collected into pools of images which continue to grow. An established platform such as *Instagram* has well established hashtags because they emerged as categories early in the history of the platform and have since gathered a dense volume of images; for example, #loveyourself contains "24.9 million posts" according to the *Instagram* explore tab. A

quantification of 24.9 million individual entities suggests the hashtag pool is no longer emergent but established. The emergence of a hashtag could be traced by waiting for social movements or social initiatives to emerge. They often create their own hashtag to publicize their ideological platform and social views while gathering more followers. Recent examples include #metoo and #notmypresident. Web resources in this case are the images generated by users, analogous to entities as discussed with Bowker and Start (1999) and Hacking (1986, 1996). Secondly, one of the crucial additions to Vander Wal's (2007) theory is that folksonomies create "a different kind of web directory" (Choy & Lui 2006).

I have used a few different words to try to describe the web directory, notably, "clustering" and "pools." Clustering refers to the moment after the image is hashtagged and uploaded, when an algorithm reads the hashtag and adds the image to the others with the same hashtag. Pool was chosen because it signifies openness, broadness, and a minimal amount of structure. Images in hashtag pools on Instagram resemble a shoebox of photographs rather than a meticulous arrangement of sub-categories as seen with Linnaean taxonomy (see figure 3 above). Images are organized only from recent to oldest in the pool. A user cannot search within the pool to narrow their results but must scroll through the pool and browse. Another important characteristic contributed by Ohkura, Kiyota and Nakagawa (2006) is folksonomies and the volume of content.

It is possible that Vander Wal (2007) did not note the number of entities handled by folksonomies because of his focus on user practices and the nature of *Flickr* functionality at the time. It is also possible that at the time of writing, *Flickr* tagging was not quantifiable. Nevertheless, hashtagging is well established on *Instagram* and requires a consideration of the assertion by Ohkura, Kiyota and Nakagawa (2006) that folksonomies "seem able to

deal with a large amount of content" (2). A user can see how many images are in a hashtag pool by searching the pool in the search bar of the explore tab. On 4 July 2018, I chose a set of hashtags to examine their volume to assess the assertion of Ohkura et. al. (2006) (see table 1).

Table 1: Entity volume of a set of hashtags

Hashtag Pool	Number of uses of that hashtag, date
#careletonuniversity	19.5 thousand
#ottawa	4.1 million
#samsung6edge	131, 601
#seanconnery	91, 631
#portugal	22.3 million

In the selection I attempted to represent a random sample of hashtags depicting people, places, and objects. Volume varies between hashtags, but the ones above seem to confirm that hashtag pools can be quite large. The final contribution to Vander Wal's (2007) considers whether a central authority is active in hashtagging.

Golder and Huberman (2006) suggest that folksonomy usage has "nobody in the librarian role or there is simply too much content for a single authority to classify; both of these are true of the web, where collaborative tagging has grown popular." In dynamic nominalism, Hacking (1986, 2006) suggested that traditional experts in scientific classification encompassed a few people formally accredited in their field of study and mobilize knowledge to engage in classification. As discussed in the section about

knowledge in dynamic nominalism, *Instagram* users are not divided into those with professional knowledge and public knowledge of hashtag usage. The knowledge involved in creating a hashtag and applying it to content is common knowledge to that public.

Hashtag users are not a small group such as traditional experts in the sciences but a majority group of content contributors or *Instagram* content producing 'experts' who try to use hashtags effectively to make their photos visible. Although there are many types of users for the purpose of this analysis all users are grouped into one since the focus here is the interface. Although, in future, distinguishing user types especially along the lines of niche communities discussed in the analysis could reveal different communities of practice. Overall, there does not seem to be any classification boundary keepers or enforcers, in the same way that librarians apply and manage collections using the Dewey Decimal System (DCC). In *Instagram*, the user decides which hashtag to use based on their subjective thinking, their intended purpose, recommendations from the platform during the upload, and by seeing how other users use the same hashtags. Any of the latter can inform an individual's final choice. Users learn how to use the platform and from others to tag their photos which is indicative of a loose form of collaboration common in virtual communities in a social media platform.

Ontologies are another form of formal classification although Thomas Gruber in (2007) *Ontology of Folksonomy* combines formal and informal classification. In addition, I will discuss Andrew Iliadis' (2016) *Black Art* and assess the potential relationship between ontologies and folksonomies.

2.4.6. Are Ontologies Involved with Folksonomies?

If a classification represents the hours on the face of a clock, the ontology represents the gears and balance wheels that turn the hands to classify the day by hours as seen in figure 7.

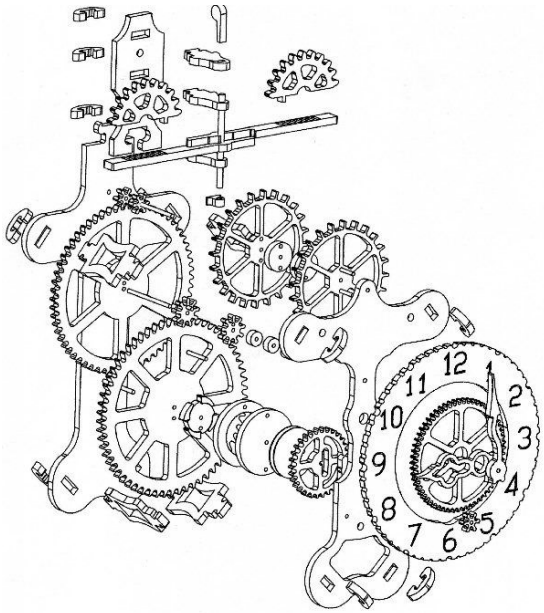


Figure 7: The relationship between underlying and overlaid technology

Gruber (2007) describes ontologies as an "enabling technology—a layer of enabling infrastructure—for information sharing and manipulation" through which "parties who have software/data/services to offer identify some common conceptualization of the data; they specify that conceptualization as clearly as they can; they build systems that interoperate those specifications" (2). If there is an ontology in *Instagram*, it is probably oriented toward information sharing rather than information manipulation because photos are shared by the user in a prosumer environment. Individuals are the parties involved, since each one has a unique product to share with the others.

Gruber (2007) defines ontologies as "specifications of conceptualizations at a semantic level" (Gruber, 2007, 2). A conceptualization in this case is the user's

understanding of the subject of their specific image. An image represents datum at the semantic level. For example, an image with a subject that is a branched organism with leaves is conceptualized as a tree. This is because a North American user learned at a young age that that particular organism is called a tree in the English language. A specification would therefore be "tree" or more specifically the sub-class "elm" if desired. Table 2, below, shows that there are many potential specifications for a single conceptualization.

Table 2: An example conceptualization and its related specifications

Conceptualization	Specification
"Tree"	Specific: "tree," "elm," "oak," "mahogany," etc.

According to Gruber (2007, page), the parties involved "specify a conceptualization as clearly as they can and then build systems that interoperate those specifications." The semantic specification of a conceptualization is hashtagging, that is, applying the hashmark and word "tree" to a photograph of a tree. For some users, the word "tree" is clear enough for their uses. For others, they may choose a more specific word such as "elm." Users develop semantic specifications of conceptualizations, but users do not build the system that facilitates interoperability. Further consideration of interoperability through *Black Art: Ontology, Data, and the Tower of Babel Problem* by Andrew Iliadis (2016) is necessary before trying to identify the system that facilitates user practices described by Gruber (2007).

Interoperability is central to Iliadis' (2016) focus on the Tower of Babel problem, which "states that each time a new database is constructed new terms are developed that represent an everchanging language thus complicating applied ontology-building, the goal

of which is to produce ontologies that can last over time" (23-24). The Tower of Babel problem represents the problem of making systems speak to each other, for example, a person who speaks French only and a person who speaks English only. Separating the two people apart are two different systems, but an accurate translation allows them to speak to each other.

Metaphorically, an ontology facilitates the translation of one set of specifications and another in a way that makes them compatible. Between two or three systems interoperability might be simple, but the difficulty of the Tower of Babel problem emerges when thinking about how many databases exist and how each one has different standards by which data are organized. There may be hundreds or thousands of standards that need to be brought into line. Further complicating the problem is the probability of a new database emerging often, so an ontology needs to account for existent databases that need to interoperate as well as accounting for databases that emerge. Accounting for emerging databases that need to interoperate is a challenge to the longevity of ontologies because the standards of existing databases are brought into line but a new database with new standards reflects a need to change the ontology to bring the new database into line with older databases.

Interoperability occurs when "ontologies can constrain wildly divergent data according to logically coherent principles of organization, thus producing semantic interoperability between heterogeneous datasets (Iliadis, 2016, 20). Note that the type of interoperability is specifically semantic, that is, related to the packaging and sharing of data across systems while maintaining the quality of the data. Ontologies intend to balance data

communicability while losing as little data quality in the translation. Semantic interoperability is the goal of an ontology (Iliadis, 2016; Pulsifer, 2008).

The types of data considered in this research were images in photos; with the images in the photos being datum at the semantic level. To refer to Gruber's (2007) terminology the images are the conceptualizations that are specified at a semantic level (2). The specifications at the semantic level are the hashtags because they refer to the subject of the image and are derived from the image by the user (Vander Wal, 2007). The hashtag individuates the image by labeling it, but it is also read by an algorithm that collects it with the other images with the same hashtag thus forming a pool. Hashtags can vary in *Instagram* as seen in figure 8 below: Arabic, Japanese, Portuguese, Ukrainian in Cyrillic letters, and English. Notice too that in some cases the native language hashtags exist beside purely English hashtags. The sole commonality between all the hashtags is the hashmark (i.e. #).

#photography #photographylover #photo #edit #editing #poetry #quote #quotes #instaquote #poem #arabic # شعر
#قصيدة #زهورة #وردة #طفولة

#gingerkitten #bobcut #森ガール #猫 #可愛い #大人可愛い #ヴィンテージ

#fernandopessoa #literatura #linguaportuguesa #historia #poema #portugal #lisboa #viagem #viajar #campinas
#curitiba #escritor #advogada #direito #letras #santacatarina #minasgerais #brasilia #blumenau #portugal_em_fotos
#biblioteca #opensador #amoler #vinhedo #gramado #viajarfazbem #paraty #livros #flip #livroseleituras

#ilovekiev #ukraine #киев #kiev #украина

#books #bookstagram #booklover #vintagebooks @vintageanchorbooks
amus #paperbacks #alvinlustig #leolionni #coverdesign #coverlove #igreads

Figure 7: Hashtags in other languages. Note the mixed languages of hashtags

The hashmark enables image clustering by algorithms, but it also allows sharing on other platforms where images are not the only form of data. *Instagram* images can be

shared on *Tumblr* and *Twitter*, which are platforms that also use hashtags but are not photo-centric. Nevertheless, an image hashtagged on *Instagram* and then shared on *Tumblr* is then clustered into the collection of content on *Instagram* and the collection of hashtagged content on *Tumblr*. The hashmark enables interoperability across platforms as it has become a standard across them all. and is a "logically coherent principles of organization," (Iliadis, 2016, 20).

Hashtagging does not suggest the presence of an ontology but instead suggests that in a platform ecosystem it is a connecter of content. Furthermore, hashtags do not represent a solution to the Tower of Babel problem because ontologies are designed to prevent "largely diversified repositories of knowledge" and make "repositories open to one another to produce new knowledge" (Iliadis, 2016, 26).

However, hashtag pools on *Instagram*, *Tumblr*, and *Twitter* are independent of each other; images can be shared across the platforms, but the pools of images exist only on each of those platforms. Furthermore, each pool is a largely diversified repository to the point that they can contain images unrelated to the hashtag suggesting that there is no platform-to-platform ontology. It would seem that ontologies are not the same as folksonomies.

2.5. Summary

The preceding chapter established *Instagram* as a platform within an ecosystem using the framework developed by Tiwana (2013) and introduced nuances to the *Instagram* platform by addressing the platform as a business model and the platform as the site of digital labor. Discussing algorithms showed how they facilitate the particular type of labor of

hashtagging images. Generally, algorithms are the site of power in online activities because they link computational systems and enmesh the user within them.

Formal classifications were addressed through Bowker and Star (1999), Lampland and Star (2009), and Hacking (1986, 1996) to show different perspectives on formal classifications and these were compared with informal classification, namely folksonomies by Vander Wal (2007) while the work of Golder and Huberman (2006) introduced contributions to folksonomy theorization. Although formal classifications are often supported by formal ontologies, it was unclear whether user generated classifications are supported by a similar sort of ontology.

In the following chapter, I will explain the methodological approach used to assess if the theories discussed apply to hashtagging practices in *Instagram*.

CHAPTER 3: METHODOLOGICAL APPROACHES

3.1. Introduction

In this chapter I introduce and discuss the methodological approach applied to answer the two research questions of this thesis, namely:

> Are early conceptions of user generated classifications relevant descriptors of hashtagging practices on *Instagram*?
>
> Do older classification theories, developed prior to hashtagging practices still apply in a user generated context?

First, I will discuss the rationale for this methodological approach, followed by a description of each method and how it was applied. The chapter will conclude with a discussion why the approach is suitable to this research.

3.2. Rational for the chosen methodology

There are a number of qualitative and quantitative approaches to the study of social media (Saldana, 2013). These might include semiotic analysis (Barthes, 1983), ethnomethodology (Thomsen, Straubhaar, Bolyard, 1998), discourse analysis (Saldana, 2013), and rhetorical

analysis. Semiotic analysis examines cultural artifacts as signs to understand what they mean and why. Ethnomethodology is a form of analysis that tries to explain the methods people use to produce and understand the social order in which they live. Discourse analysis examines artifacts, not just words and sentences, as language but also to understand discourse formation, context, and meaning. Rhetorical analysis takes a text apart to examine how its constituent parts work together to create an effect such as persuasion. Most of these examples contain methodological subtypes, for example discourse analysis includes critical discourse analysis, Foucauldian discourse analysis, etc. Furthermore, these methods are not mutually exclusive and can be combined with quantitative research methods to form a mixed methods approach. A qualitative walkthrough approach was selected for this research to assess whether established classification theories apply to hashtagging on *Instagram*. A quantitative approach could be applied, but I instead chose a fact-based theory approach.

Since Instagram is not heavily written about in the literature on classification, walkthroughs demonstrated how hashtagging and the making of classes work. Although *Instagram* continuously evolves by adding new features and updates, the walkthrough approach provided a time stamped approach to examine hashtagging processes while also enabling the collection of timely data detailing current hashtagging practices. To analyze collected datasets, descriptive coding related to classification definitions was applied because it is specifically designed for theory testing projects (Hsieh and Shannon, 2005; Saldana, 2013).

Other forms of coding in qualitative research include In Vivo coding, process coding, and values coding. In Vivo coding examines participant phrases to examine

language use in interview situations (Bryman, Bell, and Teevan, 2012). Process coding uses gerunds (-ing words) to examine the actions of actors to understand the emotions that underly their behaviors (Saldana, 2013). Values coding examines participant behavior and attempts to infer the beliefs that motivate their behaviors (Bryman, Bell, and Teevan, 2012). These coding processes reflect the focus on participant's characteristic of most coding practices. This research did not focus on the users of *Instagram* themselves. The analytic focus of this research was the relevance of the theoretical literature on classification to *Instagram* hashtagging practices in 2018. Therefore, I chose descriptive coding, whereby citations were derived from the literature discussed in chapter two to compare with the datasets collected during the walkthroughs. The following are the methodological steps applied to collect data for the analysis:

1. Read the scholarly literature.
2. Isolate key theoretical ideas.
3. Categorized the citations as general classification descriptors and specifically folksonomy descriptors.
4. Create tables in which the classification and folksonomy codes were placed.
5. Conduct the walkthrough for each hashtag pool documenting observations with screen captures to collect visual datasets.
6. Analyze by comparing visual datasets with the codes from the literature.
7. Make observations in the tables detailing correspondences and contradictions between collected data and codes.
9. Answer research questions based on observation of the comparison.

During the process, memos noted whether the code seemed an apt descriptor of hashtagging on *Instagram* (see table 3, below). The process of coding was iterative to accommodate new connections. The following sections discuss in greater detail the walkthrough method and its relationship to the theories just discussed in chapter two.

3.3. Walkthrough Method

The walkthrough method applied in this thesis was developed by Light, Burgess, and Duguay (2016) to study social media applications. The objective of their method is a combination "science and technology studies with cultural studies, through which researchers can perform a critical analysis of a given app" (Light, Burgess, & Duguay, 2016, 1). This method allows the researcher to consider the cultural significance of using an application. In this thesis the walkthrough allowed a critical consideration of how the application and platform of *Instagram* affords user generated entity creation, user classification of the entities using hashtags, and the practices informing hashtag usage. The application as well as the platform was considered during the walkthrough because the application interoperates with the platform (Tiwana, 2014). The *Instagram* application and platform work together to afford usage.

Burgess, et al., (2016) note that social media applications are the site of "significant sociocultural and economic transformations across many domains, from health and relationships to entertainment and everyday finance (Light, Burgess, & Duguay, 2016, 1). In general, social media applications are worth studying because they can be responsible for new developments in the way people communicate, make money, engage with audiences, use language to communicate, produce content, and present content, including how they label it, share it, organize it. Specific to this thesis, the *Instagram* application

affords the creation and organization of user generated entities (as defined in the literature review). As discussed earlier, the entities examined here are images and the classification practices related to hashtagging—hashtagging as a classificatory practice. *Instagram* represents one of a few applications where the shift from tagging that was established by earlier social media platforms such as *Flickr* to hashtagging, can be observed. As will be seen, the walkthrough method helps understand the relationship between user generated entities and how users classify them using hashtags.

Burgess, et al., (2016) are correct when they note that "apps configure relations among people through mediator characteristics: user interface arrangements, functions and features, textual content and tone, and symbolic representation" (Light, Burgess, and Duguay, 2016, 13-14). Applications also configure relations between users and the entities they upload to the platform. One of the relationships between *Flickr* users and their entities is the classificatory practice of tagging those entities. The affordances of the application allow users to associate words to the images they upload thereby configuring a relationship of classification as well as public sharing between user and entity whereby *Instagram* reconfigured the relationship between users and entities by affording the pairing of hashmarked (i.e. #) words to images, but the relationship remains one of classification and public sharing. The key difference is the addition of the hashmark (i.e. #) to the practice to facilitate algorithmic recognition of entities (see section 2). In this way, the perspective of Burgess et al., (2016) differs from the perspective of this research but the focus on a user relationship remains the same. This will be discussed further in the following section about the cognitive walkthrough method.

The walkthrough method is a structured approach to the study of mediator characteristics, which afford an active relationship between the user and the platform. These are interface windows, drop down menus, automatic suggestions, spaces in which to write text—to name a few of the many ways users interact with application's interfaces. In this thesis, I documented the interface arrangements, functions afforded to users, and textual content as part of the *Instagram* application in general to understand the general relationship between user and platform via the application, especially when it comes to hashtagging entities. I also documented the interface arrangements, functions afforded to users, and textual content presented to users as part of hashtagging on the platform via the *Instagram* application and via a popular third-party photo editing application. The three walkthroughs I conducted allowed for an examination of the mediator characteristics that afford the relationship between users and the platform: in general, using the official app of the platform, and using another popular app to interact with the platform.

Burgess et al., (2016) lay out the practices and purpose of the walkthrough method: "The core of this method involves the step-by-step observation and documentation of an app's screens, features, and flows of activity—slowing down the mundane actions and interactions that form part of normal app use in order to make them salient and therefore available for critical analysis" (Light, Burgess, Duguay, 2016, 3).

Screens and features summarize mediator characteristics while flows of activity demonstrate how users move between application interfaces, interact with drop down menus, and other activities that constitute usage of the application. For example, one of *Instagram*'s mediator characteristics is a dropdown menu that suggests hashtags to users; a flow of activity would be the user scrolling through the menu and choosing a hashtag to

pair with their image. Crucially, the authors note that the walkthrough slows down "mundane actions and interactions" and offers a critical examination of them (Burgess, et. al. 2016).

Mundane actions are the ones a user takes for granted and does almost mindlessly as part of application usage. An example of this is scrolling through a *Facebook* feed until one is not paying attention. Metaphorically, this is the social media equivalent of highway hypnosis. Slowing down allows the researcher to think carefully about what they do. The documentation process of taking screen captures allows the researcher to return to the site of interaction, such as an interface window, to examine it again later. In the moment of interaction, it is not always clear what is and is not a salient part of application usage even if one pays attention to what is happening in the application, so visual documentation is a valuable way to preserve mediator characteristics and interactions with them for later analysis. It also provides a record of that interaction.

In addition to the methods developed by Burgess et. al. (2016), I will discuss the cognitive walkthrough method developed by Lewis and Rieman (1993). Although this research was conducted primarily using the walkthrough method of Burgess et. al. (2016), Lewis and Rieman (1993) provide additional and useful knowledge about the process. In the next section I will discuss the importance of bracketing off knowledge when conducting a walkthrough.

3.4. Cognitive Walkthrough Method

As discussed, Burgess et. al. (2016) focus on the relationships between users because of the mediator characteristics of applications while the research being conducted for this thesis focused on the relationships afforded by mediator characteristics between users and the

entities they classify, specifically photographs but more so on the interface and platform and the patterns left behind by users. Furthermore, the perspective of an application as a tool used by users aligns with Lewis and Rieman (1993) who approach applications from a software design perspective rather than the socio-technical constructivist approach offered by Burgess et. al. (2016). Rather than looking at the way technology helps construct social relations, the authors approach applications as tools used by users, which aligns with the focus on application usage that informs this thesis.

Lewis and Rieman (1993) developed the cognitive walkthrough to evaluate the ease of use of an interface because they acknowledge that, as software designers, the interfaces they design may be easy to use for people with software design knowledge but difficult for users who lack that knowledge. A user with software design knowledge might be able to navigate a complex interface, but users without that knowledge require a simple and intuitive interface. Lewis and Rieman (1993) develop a set of heuristic criteria which a designer can apply during their walkthrough to assess the ease of use of an interface. Their heuristic criteria do not apply to this research because I did not focus on ease of use, but the additional guidance they give to designers conducting a walkthrough is relevant.

Like Burgess et. al. (2016) the interface is approached from the perspective of a user. Unlike Burgess et. al. (2016), the authors note that knowledge of the back-end workings of an interface should be suspended (Lewis and Rieman, 1993). For software designers, suspending what they know about how their interface works makes them experience the interface as an average user. The process intends to make the designers self-reflexive. Suspending the back-end workings of the walkthrough was relevant to this research because I had conducted some research about the back-end workings of the *Instagram*

application and platform prior to conducting my walkthroughs. I needed to keep that knowledge separate from the walkthrough, so I could focus on documenting the mediator characteristics and their affordances without the distraction of trying to explain the technical components at play. For example, I might have missed important details of uploading if I had walked through the uploading process and tried to explain the interactions between myself and the interfaces using back-end knowledge of the *Instagram* platform at the same time.

Instead, I bracketed off the knowledge of underlying technology as much as possible and focused on the experience of using the application in general and to hashtag. This allowed me to objectively and reflexively focus on the walkthrough experience and to document the walkthrough from an interface and workflow perspective. After the walkthroughs were complete and documentation organized, I involved back-end knowledge such as the fact that algorithms cluster images. Although a small modification to the walkthrough method of Burgess et. al. (2016) by Lewis and Rieman (1993) informed the approach to the walkthrough by emphasizing in-depth experience and documentation. Only afterwards was knowledge of the back-end workings applied to attempt understanding the interaction with the mediator characteristics and interfaces of *Instagram*.

3.5. Applied Methodological Approach

In this thesis walkthrough methods were adopted, platforms were considered as per Tiwana's (2014) platform ecosystem framework discussed in section 2.1 of this thesis. First, I considered Tiwana's (2014) materialist description of platforms, applications, and their ecosystems. The walkthrough method approached the study of the application as a discursive script informed by the socio-technological constructivist approach developed by

Light, Burgess, and Duguay (2016). This approach provides a structured methodology to conduct social media research especially for the study of the discursive practices of a digital artefact such as an app or an ecosystem such as a platform. Light, Burgess, and Duguay (2016) provide guidelines to mindfully study the mediator characteristics of a digital object which are: "user interface arrangements, functions and features (affordances), textual content and tone, and symbolic representation" (Light, Burgess, Duguay, 2016, 13-14). The characteristics include icons, drop down menus, and uploading windows. Other guidelines include documenting the walkthrough, so that the images can be analyzed later. In this project, screenshots were collected as seen in appendix A.

The second methodological approach was the cognitive walkthrough proposed by Lewis and Rieman (1993). These scholars assess the user friendliness of an interface. This involves suspending knowledge of the back-end workings of an interface and focusing on the experience of using it. In this case, data collection involved suspending knowledge of algorithms and other computational workings to focus on the user experience of *Instagram*. The cognitive walkthrough of Lewis and Rieman (1993) informed the walkthrough developed by Light, Burgess, and Duguay (2016). Together, they allowed data collection that focused on the user experience as intermediated by the platform's interface. Specifically, the experience of using *Instagram* to the hashtag and following the hashtagged content as seen in figures 8 and 9 included in chapter four.

The methodological approaches were applied to the collection of visual data sets. From the corpus of collected data key criteria regarding classification theories in the form of citations acted as codes (Hsieh and Shannon, 2005). Codes were separated into two categories: formal classifications and folksonomy. The codes were tested against the data

sets to see if they were "summative, salient, essence-capturing" assessments of the user experience in accordance with descriptive coding (see chapter 4) (Saldana, 2013). The relevance of the scholarly theories was based on their comparison to the walkthrough experience through the coding process.

3.6. Suitability of the Methodology

The methodological approach I chose aims to explain the relationship between users and the platform by providing a structured approach to user engagement in *Instagram* hashtagging practices and to study classifications. Overall, my modified walkthrough, screen capture documentation, and the comparison of visual datasets to existing classificatory theories has proven useful as a method for this hashtag research when approaching hashtags as classifications. The following will now address the reasons for a hybrid walkthrough method rather than that of Burgess, et. al. (2016) alone.

The cognitive walkthrough of Lewis and Rieman (1993) informed the walkthrough developed by Light, Burgess, and Duguay (2016). The combined method allowed empirical data collection that focused on the user experience by suspending knowledge of the back-end workings of an interface such as algorithms and other computational workings. Lewis and Rieman (1993) encourage the researcher to bracket off the computational workings and focus on the experience of interacting with the interface, in this case, the *Instagram* application and the *Instagram* platform. Burgess et al., (2016) provide a useful framework for studying the experience as a researcher rather than a user. Specifically, the experience of using *Instagram* to hashtag and following the hashtagged content. This methodological approach for studying *Instagram* hashtagging resulted in an objective data collection

process. The author of this thesis is an *Instagram* user, therefore the subjective user needed to become an objective researcher.

The walkthrough of the *Instagram* platform allowed the documentation of the mediator characteristics of the application to assess affordances (Light, et. al., 2016). Here the researcher critically documented the limitations and rights given to users by the terms and conditions (Instagram, 2018) to show the conditions under which users create content. Affordances of the platform were documented to show the option available to users such as profiles for users, how profiles are used, and liking content and collecting content. This provided a general situation in the platform from a research perspective.

The second walkthrough focused on mediator characteristics of hashtagging specifically. The principle behind this walkthrough was to determine a deep understanding of hashtagging practices. To do this, the researcher uploaded photos and explored hashtag pools to situate himself as a critical member of the community of practice. I performed walkthroughs on *Instagram* and related third-party application VSCO (a popular photo editing application) to see if this affected the classification process. The purpose of the walkthroughs was to slow down and pay attention to the details (Light, Burgess, and Duguay, 2016, 3). To answer the research questions, the literature about classification systems was compared to the experience of producing classes and an examination of pools of images that were hashtagged.

This was not without challenges. One of the challenging aspects of the theoretical literature about classification systems are their varied focuses, terminology, and examples. Thomas Vander Wal (2007) detailed practices and corresponded them to a specific term: folksonomy. Scholars such as Geoffrey Bowker and Susan Leigh Star (1999) detail

everyday classification systems but do not label them. Scholars such as Ian Hacking (2006, 1986) write about the way classification systems work, why they are used especially by institutions with authoritative power. The purpose of this thesis was to compare and contrast the literature about classification systems to the practice of hashtagging on *Instagram* to assess the descriptive relevance of the literature to the practices. The diversity of the literature made it ideal for the purposes of this thesis, but also presented a challenge to conduct a qualitative comparative analysis.

A variety of scholarly articles were chosen as a representative sample of different perspectives on classifications to understand hashtagging on *Instagram* as classification. The comparison is useful but not comprehensive but useful for theory testing.

A table of key ideas was assembled of the concepts and definitions distilled from the literature review of formal classifications and informal classification processes such as folksonomies (see chapters four and five). Relevance was determined based on how well the concepts acted as "summative, salient, essence-capturing" assessments of the walkthrough experience (Saldana, 2013). For example, whether "mutually exclusive" as a characteristic of formal classification describes *Instagram* hashtags (Bowker and Star, 1999). Descriptive coding afforded an understanding of the relevance of the literature to hashtagging practices on *Instagram*.

As will be seen in chapter 4, the table of folksonomic literature codes (see table 9) corresponded to the first research question by comparing the walkthrough of hashtagging to early conceptions of user generated classifications. While the tables for formal classifications (see tables 4 through 9) correspond to the second research question by comparing the walkthrough of hashtagging to formal theories.

CHAPTER 4: DATA COLLECTION

4.1. Introduction

In the methodology, the walkthrough method for data collection and the descriptive coding method were described. In this chapter, the application of the walkthrough method to collect data is provided including the application of the descriptive coding to analyze the data the results of which will be discussed chapter five. During the walkthroughs, I documented application usage with the screen capture tool of my phone to provide empirical documentation to support my analysis and conclusions. The remainder of this chapter will discuss how this was done.

4.2. The Walkthrough Flowline

The flowline illustrated in figure 10 below describes the walkthrough stages of hashtagging an image. From the smartphone, I accessed the *Instagram* application, which took me to the home screen. At the bottom of the home screen is a "+" inside a square signifying the upload option. Pressing the button opens the upload interface. The first step is choosing a photo from the phone, which demonstrates the interoperation of the *Instagram* application

and the gallery application of the smart phone. A user chooses an image from their gallery and automatically moves to the next step. The next step is adding a filter to change the color tone of the image. A user can switch between the filter and edit interfaces within the same window to tone and adjust their image. Now the user presses "next" and arrives at the final interface. Here they can caption their photo, add hashtags, choose to include a location tag, and toggle automatic sharing between *Instagram* and linked platforms. A user can link their *Facebook* or Tumblr account to *Instagram* to automatically share their image to those platforms as soon as it is uploaded to *Instagram*. The caption box allows the user to write a string of text if they choose, add emojis if they choose, and add hashtags. An image can have all three components or any combination of the three. Emojis were however not part of this analysis as these were not part of classifying since these are not related to hashtagging, rather, to captioning an image. Adding a hashtag to an image is an example of human-computer interaction. As a user types "#" and the first few letters of the hashtagged text, the application suggests hashtags to the user. These appear as a popup menu that shows suggested hashtags in order of most popular to least popular. The user can choose as many hashtags they want. Once satisfied the user presses "upload." The images appear on the home feed where other users in the uploader's network can see it. It appears in the profile of the user. Finally, it appears alongside other images in pools.

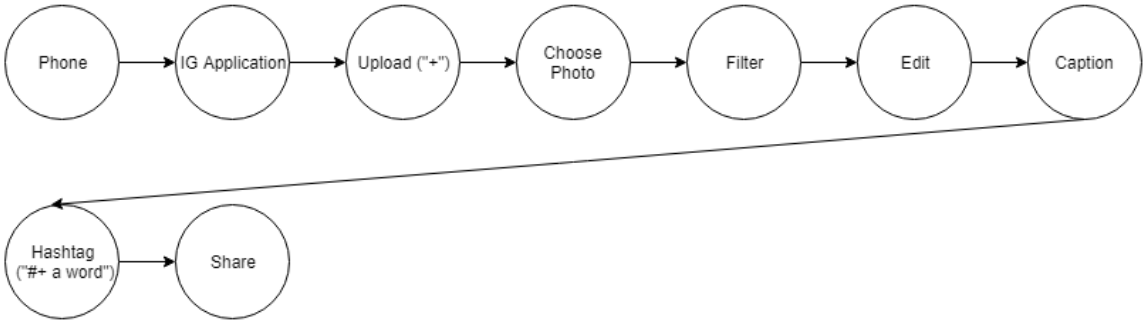


Figure 8: *Instagram* Image Hashtagging Flow Line

The next flowline illustrates the steps to access and explore a hashtag pool. From the smartphone, I accessed the *Instagram* application, which took me to the home screen. From the symbols across the bottom of the home screen, a user presses the magnifying glass symbol to assess the search interface. This takes the user to a search interface where they can type a word and have hashtags suggested to them. The user does not have to type # when searching for a hashtag, only a word. As the user types, the platform suggests hashtags based on the query in a drop-down menu. Once the user chooses a hashtag from the dropdown menu they are taken to the pool where they scroll down to start exploring the pool. Users see some recent images highlighted at the top of the pool as they scroll down and can scroll down to explore other images in the pool created by the hashtag. This can be a time-consuming activity because hashtag pools often contain thousands of images.

Figure 9: Accessing an *Instagram* Hashtag Pool Flow Line

During each walkthrough of each hashtag pool, screen captures were taken as data to document the content of the pools. I followed a data collection schedule (see table 3 below) and at the appointed time accessed the hashtag pools and scrolled through them. During the scrolling process I did not decide which images to screen capture. I chose samples as randomly as possible to ensure actual representation of the hashtag pools at any a given moment. Once I stopped scrolling I immediately took a screenshot. Each screen shot contains fifteen images and two were taken per hashtag pool resulting in a sample of 30 images per pool, with a total of 1,050 screen captures. I sampled hashtag pools each day, for one hour, within a one-month period (see table 3 below). The screen shots provided empirical visual data sets against which to compare the theoretical literature. To compare the literature, I organized the tables by author and subject. I then summarized in one word each of the key characteristic of classification systems as theorized by the each of the authors. Each characteristic was accorded its own column. Each characteristic was compared to the visual data sets to see if it applied to the hashtag pool. These are represented in the tables 4 through 9 below and will be discussed further in this chapter 5. Note that X means the characteristic applies and a blank box means that it does not.

4.3. Data Collection

Table 3 below illustrates the data collection schedules, such as the time when data were collected, the number of screen captures taken, and the total number of photos in the screen captures. This table is included to show that 35 hashtag pools were walked through and a total of 1,050 images collected. These data were collected with the aim of being as

objective as possible across a large and varied number of hashtag pools that represented general things. I also included some observations.

Table 3: Data Collection Details

Categories	Pool Name	Day of Collection	Time of Collection	No. of Screen Captures	Total No. of Photos	Observations
Nature	#tree	1	6-7 p.m.	2	30	Includes trees and people with trees.
	#cat	2	9-10 p.m.	2	30	Strict exclusivity of images—purely cats.
	#fish	3	12-1 a.m.	2	30	Two communities of practice obvious: fishers and chefs.
	#lakedistrict	4	3-4 a.m.	2	30	Exclusive application of labels to location and images of the landscape resemble each other.
Celebrities	#janebirkin	5	6-7 p.m.	2	30	Jane in each image with the addition of a hanger on, i.e. Serge.
	#seanconnery	6	9-10 p.m.	2	30	Sean in each image alone or with another person.
	#jonimitchell	7	12-1 a.m.	2	30	Mostly images of Joni, but also includes art or memes inspired by her. Not as exclusive but related.

Categories	Pool Name	Day of Collection	Time of Collection	No. of Screen Captures	Total No. of Photos	Observations
Tech	#samsung6edge	8	3-4 a.m.	2	30	Largely exclusive to images taken with the phone.
	#acer	9	6-7 p.m.	2	30	Exclusive to images of their products or of people using their products.
Places	#ottawa	10	9-10 p.m.	2	30	Very little to suggest that these photos are from Ottawa save a few landmarks (one of the sculptures at NGA and one with parliament).
	#portugal	11	12-1 a.m.	2	30	Portuguese, but shows a diversity of representative subjects (architecture, football, holidays, etc.)
	#hollywood	12	3-4 a.m.	2	30	Looks surprisingly unlike Hollywood—no celebrities or landmarks associated with it.
	#france	13	6-7 p.m.	2	30	A few landmarks such as the Eiffel Tower but also

Categories	Pool Name	Day of Collection	Time of Collection	No. of Screen Captures	Total No. of Photos	Observations
						irrelevant images such as St. Basil's Cathedral (Russia) and "Follow for Follow Back."
	#canada	14	9-10 p.m.	2	30	No like-ness between images and images do not suggest some sort of Canadian-ness. They could have been taken anywhere.
	#greece	15	12-1 a.m.	2	30	People on holiday and architecture distinctly Greek. The split is roughly half.
	#montreal	16	3-4 a.m.	2	30	As with #canada, there is little to suggest that these images were taken in Montreal.
Mental Health	#panicattack	17	6-7 p.m.	2	30	Some images are about supporting people in the situation or eliminating stigma. The others depict people or activities unrelated

Categories	Pool Name	Day of Collection	Time of Collection	No. of Screen Captures	Total No. of Photos	Observations
						(e.g. someone surfing).
	#mentalhealth	18	9-10 p.m.	2	30	Similar to #panicattack, a loose connection between support and education. A broad variety of subjects in the images.
Arts	#picasso	19	12-1 a.m.	2	30	The first capture has little to do with Picasso, but 8/30 in the second show his paintings. In both shots there is a Van Gogh painting, why is unclear.
	#britishart	20	3-4 a.m.	2	30	A diversity of British arts, mostly paintings from no particular time.
Lifestyle	#book	21	6-7 p.m.	2	30	It is uncertain why some images include tarot cards and a woman on the beach in silhouette. About half of the images in total are of books.

Categories	Pool Name	Day of Collection	Time of Collection	No. of Screen Captures	Total No. of Photos	Observations
	#exercise	22	9-10 p.m.	2	30	Like the scientists in that exercise is represented clearly but in this case as an activity.
	#minimalism	23	12-1 a.m.	2	30	A variety of images symbolic of that way of life, i.e. "less is more."
	#vinyl	24	3-4 a.m.	2	30	Features LPs mostly, some vinyl collections as well.
	#walking	25	6-7 p.m.	2	30	Some images of walkers, but most do not suggest walking.
	#thrifting	26	9-10 p.m.	2	30	The activity of thrifting was not depicted as often as the items found and worn after the activity (4/30).
Learning	#carletonuniversity	27	12-1 a.m.	2	30	Images of students, locations on campus, unknown ads in Chinese.
	#gradschool	28	3-4 a.m.	2	30	Largely similar images around life in grad school and supporting grad students.

Categories	Pool Name	Day of Collection	Time of Collection	No. of Screen Captures	Total No. of Photos	Observations
	#chemist	29	6-7 p.m.	2	30	Related largely to chemists in the field, chemistry as a field (diagrams, experiments), and molecular behavior.
	#library	30	9-10 p.m.	2	30	4/30 show actual libraries. Most of the others show books.
	#astronomy	31	12-1 a.m.	2	30	Focused mostly on the subjects of astronomy. An identifiable like-ness between images.
	#physics	28	3-4 a.m.	2	30	Similar to the chemists. Related largely to physics in the field, physics as a field (diagrams), and physics experiments.
Misc.	#morning	29	3-4 a.m.	2	30	Shows people in the morning, or related things like breakfasts and sunrises.
	#night	30	12-1 a.m.	2	30	Unlike #morning, many images do not represent

Categories	Pool Name	Day of Collection	Time of Collection	No. of Screen Captures	Total No. of Photos	Observations
						night or activities done at night. About two thirds show no relation to the time of day.
	#holiday	31	9-10 p.m.	2	30	All the images share a similar holiday-ness that is out of the ordinary, relaxed, and tropical.

4.3.1. Sorting things Out: Bowker and Star characteristics of a classification, (1999)

Table 4 below shows how the characteristics of classifications in general compared to screen capture data resulting from the walkthrough. The six characteristics represent what Bowker and Star (1999) theorize to be common in classification processes.

1. They are mutually exclusive meaning that words like tree and cat apply only to images of trees and cats.
2. Visibility refers to the quality of being able to see the entity which is classified, in this case images on a social media platform (Bowker and Star, 1999).
3. Relativity to communities of practice refers to the classification being related to a specific group of classifiers, broadly *Instagram* users but possibly niche communities are active (Bowker and Star, 1999). I will address the possibility of

niche communities of practice within the community of users in the analysis chapter.

4. Comparability refers to the usage of the classification the same way between individuals (Bowker and Star, 1999). For example, user A hashtags an image cat with #cat and user B hashtags an image of a cat with #cat. We can compare the way user A classified and the way user B classified and conclude that they used the classification the same way based on what was seen in *Instagram*.
5. Consistent and unique classification principles refer to the methods of classification remaining the same so that user A classifies along the same principles as user B (Bowker and Star, 1999).
6. Finally, the classification is supposed to be complete. Complete classification systems organize all that they are designed to organize, so that all images assigned #cat are collected under that hashtag by the algorithms of the platform (Bowker and Star, 1999).

Table 4: General Classification Characteristics Bowker and Star, (1999)

Categories	Pool Name	Mutually Exclusive	Visibility	Relative to Communities of Practice	Comparability	Consistent /Unique	Complete System
Nature	#tree	X	X	X	X	X	X
	#cat	X	X	X	X	X	X
	#fish		X	X	X		X
	#lakedistrict	X	X	X	X	X	X
Celebrities	#janebirkin	X	X	X	X	X	X
	#seanconnery	X	X	X	X	X	X

Categories	Pool Name	Mutually Exclusive	Visibility	Relative to Communities of Practice	Comparability	Consistent /Unique	Complete System
	#jonimitchell		X	X	X		X
Technology	#samsung6edge	X	X	X	X		X
	#acer	X	X	X	X	X	X
Places	#ottawa		X	X			X
	#portugal	X	X	X	X	X	X
	#hollywood		X	X	X		X
	#france		X	X	X		X
	#canada		X	X	X		X
	#greece		X	X	X		X
	#montreal		X	X	X		X
Mental Health	#panicattack		X	X	X		X
	#mentalhealth		X	X	X		X
Arts	#picasso		X	X			X
	#britishart	X	X	X	X	X	X
Lifestyle	#book		X	X			X
	#exercise	X	X	X	X	X	X
	#minimalism	X	X	X	X		X
	#vinyl	X	X	X	X	X	X
	#walking		X	X	X		X
	#thrifting		X	X	X		X
Learning	#carletonuniversity		X	X	X		X
	#gradschool	X	X	X	X		X
	#chemist	X	X	X	X	X	X

Categories	Pool Name	Mutually Exclusive	Visibility	Relative to Communities of Practice	Comparability	Consistent /Unique	Complete System
	#library		X	X	X		X
	#astronomy	X	X	X	X	X	X
	#physics	X	X	X	X		X
Misc.	#morning	X	X	X	X		X
	#night		X	X	X		X
	#holiday	X	X	X	X	X	X

4.3.2. Characteristics of a classification standards, Bowker and Star (1999)

Table 5 below shows how the three characteristics of classification standards compared to the walkthrough screen captures.

1. Standards are the regulations for the creation and implementation of classifications (Bowker and Star, 1999). In this case, the classifications are hashtags implemented to cluster images according to the hashtag pools.
2. Agreed upon regulations refer to the common principles by which a classification is implemented. A classification is supposed to be consistent and regulations determine, among other things, what belongs in a classification and what does not. Bowker and Star (1999) suggest that legal bodies and established institutions enforce standards of classifications. In this research, I broadened the characteristic to consider whether there were are any enforcers in *Instagram* to manage classification standards.

The last characteristic refers to the difficulty of changing classification standards because they have been solidified by repetitive use (Bowker and Star, 1999). I considered whether the classification could be changed easily in any regard, or whether Bowker and Star

(1999) are correct in suggesting that the change would be difficult and expensive. I will discuss this further in the analysis.

Table 5: Bowker and Star's Classification Standards (1999)

Categories	Pool Name	Agreed upon Regulations	Enforcers	Difficult to Change
Nature	#tree			X
	#cat			X
	#fish			X
	#lakedistrict			X
Celebrities	#janebirkin			X
	#seanconnery			X
	#jonimitchell			X
Technology	#samsung6edge			X
	#acer			X
Places	#ottawa			X
	#portugal			X
	#hollywood			X
	#france			X
	#canada			X
	#greece			X
	#montreal			X
Mental Health	#panicattack			X
	#mentalhealth			X
Arts	#picasso			X
	#britishart			X
Lifestyle	#book			X
	#exercise			X
	#minimalism			X
	#vinyl			X
	#walking			X
	#thrifting			X
Learning	#carletonuniversity			X
	#gradschool			X
	#chemist			X
	#library			X
	#astronomy			X
	#physics			X
Misc.	#morning			X
	#night			X
	#holiday			x

4.3.3. Characteristics of Infrastructure, Bowker and Star, (1999)

Table 6 below shows the characteristics of the infrastructure in which classifications are embedded compared with the walkthrough screen captures. The infrastructure for this research refers to the environment that facilitates the creation and use of hashtags, ergo the *Instagram* platform here is the infrastructure. I considered each of the five characteristics as follows when looking at the screen captures of the walkthroughs.

1. A "practical match among routines of work practice, technology, and wider scale organizational and technical resources" which refers to the way infrastructure and its resources work with classifiers to help them practice the classification (Bowker and Star, 1999). Each element works with the other to facilitate classification.
2. Multifunctionality refers to the infrastructure's ability to support a variety of agendas. For example, social media platforms have become venues for selling products, advertising services, promoting causes, sharing news, sharing images, and contributing to citizen journalism rather than their remaining a venue where one socializes with friends.
3. Members of the infrastructure learn how to use the infrastructure as part of their membership, which is like signing up for *Instagram* because one becomes a member according to the terms and conditions of service (Terms, 2018). As a member of the infrastructure, one learns how to use it to classify entities. I will discuss this further in the analysis chapter below.
4. Rather than existing side by side as seen in Tiwana's (2014) platform ecosystem framework, Bowker and Star (2014) suggest that infrastructures are embedded in other infrastructures.

5. Finally, ready to hand refers to the ease of accessing and using the infrastructure (Bowker and Star, 1999). If an infrastructure is meant for people to use, its principles should be easy to understand and its interfaces intuitive enough to minimize relearning. In this way, the user spends more time using the infrastructure and the classification system.

Table 6: Bowker and Star's Characteristics of Infrastructure Housing Classification (1999)

Categories	Pool Name	Match between Routines	Multifunctionality	Usage Learned by Members	Embedded in Other Structures	Ready to Hand
Nature	#tree	X	X	X		X
	#cat	X	X	X		X
	#fish	X	X	X		X
	#lakedistrict	X	X	X		X
Celebrities	#janebirkin	X	X	X		X
	#seanconnery	X	X	X		X
	#jonimitchell	X	X	X		X
Technology	#samsung6edge	X	X	X		X
	#acer	X	X	X		X
Places	#ottawa	X	X	X		X
	#portugal	X	X	X		X
	#hollywood	X	X	X		X
	#france	X	X	X		X
	#canada	X	X	X		X
	#greece	X	X	X		X
	#montreal	X	X	X		X
Mental Health	#panicattack	X	X	X		X
	#mentalhealth	X	X	X		X
Arts	#picasso	X	X	X		X
	#britishart	X	X	X		X
Lifestyle	#book	X	X	X		X
	#exercise	X	X	X		X
	#minimalism	X	X	X		X
	#vinyl	X	X	X		X
	#walking	X	X	X		X
	#thrifting	X	X	X		X
Learning	#carletonuniversity	X	X	X		X

Categories	Pool Name	Match between Routines	Multifunctionality	Usage Learned by Members	Embedded in Other Structures	Ready to Hand
	#gradschool	X	X	X		X
	#chemist	X	X	X		X
	#library	X	X	X		X
	#astronomy	X	X	X		X
	#physics	X	X	X		X
Misc.	#morning	X	X	X		X
	#night	X	X	X		X
	#holiday	X	X	X		X

4.3.4. Ian Hacking Dynamic Nominalism (1986, 1996)

Table seven below compares the characteristics of dynamic nominalism to the walkthrough screen captures. Ian Hacking's (1986, 2006) framework included five clearly defined characteristics. Each of the five characteristics is part of the interactive cycle through which classification occurs.

1. The first characteristic is a mutually exclusive classification. Mutual exclusivity here refers to exclusive categories as well as exclusive groups of entities. Each group is on its own and it contains entities that belong only to that group.
2. Entities refer to that which is classified, in this case images. In Hacking's (1986, 2006) case, his entities are people being classified.
3. The institution in this framework is similar to the infrastructure of Bowker and Star (1999) as the site that allows classification to occur.
4. Related to the institution are knowledge and experts. I have designated experts as technical to reflect that the experts are the users of the technical affordances of *Instagram*. Furthermore, the experts are analogous to Bowker and Star's (1999) members.

5. The difference between Bowker and Star's (1999) members and Hacking's (1986, 2006) experts is that experts mobilize knowledge to classify. I compared the characteristics of dynamic nominalism to the visual data from the walkthroughs to create this table.

Table 7: Hacking's Dynamic Nominalist Framework (1986, 2006)

Categories	Pool Name	Classification (mutual/ exclusive)	Entities	Institution	Knowledge (technical)	Experts (technical)
Nature	#tree	X	X	X	X	X
	#cat	X	X	X	X	X
	#fish		X	X	X	X
	#lakedistrict	X	X	X	X	X
Celebrities	#janebirkin	X	X	X	X	X
	#seanconnery	X	X	X	X	X
	#jonimitchell		X	X	X	X
Technology	#samsung6edge	X	X	X	X	X
	#acer	X	X	X	X	X
Places	#ottawa		X	X	X	X
	#portugal	X	X	X	X	X
	#hollywood		X	X	X	X
	#france		X	X	X	X
	#canada		X	X	X	X
	#greece		X	X	X	X
	#montreal		X	X	X	X
Mental Health	#panicattack		X	X	X	X
	#mentalhealth		X	X	X	X

Categories	Pool Name	Classification (mutual/ exclusive)	Entities	Institution	Knowledge (technical)	Experts (technical)
Arts	#picasso		X	X	X	X
	#britishart	X	X	X	X	X
Lifestyle	#book		X	X	X	X
	#exercise	X	X	X	X	X
	#minimalism	X	X	X	X	X
	#vinyl	X	X	X	X	X
	#walking		X	X	X	X
	#thrifting		X	X	X	X
Learning	#carletonuniversity		X	X	X	X
	#gradschool	X	X	X	X	X
	#chemist	X	X	X	X	X
	#library		X	X	X	X
	#astronomy	X	X	X	X	X
	#physics	X	X	X	X	X
Misc.	#morning	X	X	X	X	X
	#night		X	X	X	X
	#holiday	X	X	X	X	X

4.3.5. Hacking's Dynamic Nominalist Engines (1986, 2006)

Table 8 below shows how the characteristics of dynamic nominalism's engines compared to the walkthrough screen captures. I had some uncertainty trying to apply scientification and taking action. These are modifications to the original engines of resistance made by Lauriault (2012) in her dissertation examining the social construction of geographic spaces because the original engines biologizing, medicalizing, and geneticizing did not apply to

things and places. Lauriault's (2012) analogous engines are applicable to this research because it is entities classified and not people. Although the term formalization of a class might be a better term, the term scientification was kept as the ideas are analogous, namely, once something becomes scientifically affirmed it becomes difficult to change it, here, if a class becomes a popular cluster, it becomes difficult to change. For the time being, I will concentrate on the engines that were applicable to the screen captures collected during the walkthroughs. Three engines were applicable: counting, quantifying, and norms.

1. Counting refers to a body of entities that is assembled, while
2. quantifying refers to the representation of body of entities. During the walkthrough of hashtagging an image, the user is presented with suggested hashtags that represent the body of entities within each hashtag as a numerical value.
3. Norms represent the expected outcomes of counting and quantifying. In this case, I saw the similarity between images and the hashtag as a norm when I walked through the pools of hashtagged images. The expected quality of a hashtag is that it contains images related to the words of the hashtag. The table below shows where images largely matched the hashtag. I chose to look for norms elsewhere because I could not establish a normative outcome related to the suggested hashtags, for example that users consistently chose the hashtags at the top of the list because they were perceived as popular due to the large quantity of images they contain. That would require an interview with users or additional data from *Instagram*.
4. Likewise, a correlation did not emerge during the comparison of *Instagram*

walkthrough data to this characteristic. I will discuss this further in the analysis chapter.

5. Action implies the isolation of a problem, development of a solution, and the application of the solution to the perceived problem.
6. Normalization implies the act of trying to bring deviants in line with what is considered normal (Hacking, 2006, 11).
7. Bureaucracy here implies system of administration (Hacking, 2006, 11)
8. Furthermore, I will address the applicability of the engines of resistance because of the unique challenges they posed in this research. In the analysis I also take up the question of whether biologizing, medicalizing, and geneticizing or the alternative scientification are applicable or whether they need to be modified for social media research.

Table 8: Hacking's Dynamic Nominalist Engines (1986, 1996)

Categories	Pool Name	Count (volume)	Quantify (numeral)	Norm	Correlate	Action	Scientifica -tion	Normalize	Bureau -cracy	Resis- tance
Nature	#tree	X	X	X	?	?	?			
	#cat	X	X	X	?	?	?			
	#fish	X	X		?	?	?			
	#lakedistrict	X	X	X	?	?	?			
Celebrities	#janebirkin	X	X	X	?	?	?			
	#seanconnery	X	X	X	?	?	?			
	#jonimitchell	X	X		?	?	?			
Tech	#samsung6edge	X	X	X	?	?	?			
	#acer	X	X	X	?	?	?			
Places	#ottawa	X	X		?	?	?			

Categories	Pool Name	Count (volume)	Quantify (numeral)	Norm	Correlate	Action	Scientifica -tion	Normalize	Bureau -cracy	Resis- tance
	#portugal	X	X	X	?	?	?			
	#hollywood	X	X		?	?	?			
	#france	X	X		?	?	?			
	#canada	X	X		?	?	?			
	#greece	X	X		?	?	?			
	#montreal	X	X		?	?	?			
Mental Health	#panicattack	X	X		?	?	?			
	#mentalhealth	X	X		?	?	?			
Arts	#picasso	X	X		?	?	?			
	#britishart	X	X	X	?	?	?			
Lifestyle	#book	X	X		?	?	?			
	#exercise	X	X	X	?	?	?			
	#minimalism	X	X	X	?	?	?			
	#vinyl	X	X	X	?	?	?			
	#walking	X	X		?	?	?			
	#thrifting	X	X		?	?	?			
Learning	#carletonuniversity	X	X		?	?	?			
	#gradschool	X	X	X	?	?	?			
	#chemist	X	X	X	?	?	?			
	#library	X	X		?	?	?			
	#astronomy	X	X	X	?	?	?			
	#physics	X	X	X	?	?	?			
Misc.	#morning	X	X	X	?	?	?			

Categories	Pool Name	Count (volume)	Quantify (numeral)	Norm	Correlate	Action	Scientifica -tion	Normalize	Bureau -cracy	Resis- tance
	#night	X	X		?	?	?			
	#holiday	X	X	X	?	?	?			

4.3.6. Folksonomy Literature

Table 9 below shows the characteristics of folksonomy literature and how they apply when compared with the walkthrough screen captures.

1. Ohkura, et. al. (2006) note that folksonomies seem "able to deal with a large amount of content," which means that they are capable of collecting large volumes of information. Large volumes of entities are evidenced by the numerical quantification of hashtags encountered during the walkthrough of the hashtagging process.
2. Choy and Lui suggest that (2006) folksonomies lead to an "emergent categorization of web resources in terms of tags and creates a different kind of web directory." A traditional directory either structures a database like a table of contents or allows a user to search within the database, but a hashtag pool allows neither and the user must scroll through it to explore the contents. I chose the term pool to reflect the expansive size yet inability to explore in a structured way.
3. Golder and Huberman (2005) suggest that there are no central authorities responsible for hashtagging and that there is too much content to classify. The volume of entities was assessed again by comparing the walkthrough screen captures to the characteristic and determining if a person could organize a thousand or more images in real time. Images appear in the hashtag pools

immediately. There is a debate as to whether hashtagging is a collaborative or non-collaborative process. Vander Wal (2007) suggested that they are not collaborative while

4. Golder and Huberman (2005) suggested that they are collaborative. To assess this, I examined whether users were talking about how to use hashtags in the comments of their images and I did not find any conversations.
5. Vander Wal (2007) suggests that they are not collaborative, which was demonstrated by the absence of conversations noted in point four above.
6. Vander Wal (2007) suggests that users infer the label text from the image subjects, but the mismatch between hashtags and images in of most hashtag pools challenged that assumption.
7. Vander Wal (2007) suggests that images are classified for user retrieval, but this was challenged when reflecting on the volume of photographs contained in a hashtag pool combined with the inability to search the pool for a specific image. Each user has their own profile in which their images are collected.
8. Vander Wal's (2007) three data points bear a resemblance to Hacking's (1986, 1996) experts, entities, and classification because they are 1) the person tagging; 2) the object being tagged; and 3) the tag being used on that object.
9. Finally, Vander Wal's suggested that folksonomies occur in a public environment. As a member of the *Instagram* platform, I was able to see other user's photos and potentially interact with them by leaving comments. Seeing and potentially interacting with strangers was taken as symbolic of being in a public environment analogous to walking in a public city space.

Table 9: Folksonomy Literature

Categories	Pool Name	Large Volumes	New Directory	No Overseers	Collaborative	Not collaborative	For Retrieval	Inferred Label	Three Data Points	Public Environ.
Nature	#tree	X	X	X		X			X	X
	#cat	X	X	X		X		X	X	X
	#fish	X	X	X		X		X	X	X
	#lakedistrict	X	X	X		X			X	X
Celebrities	#janebirkin	X	X	X		X		X	X	X
	#seanconnery	X	X	X		X		X	X	X
	#jonimitchell	X	X	X		X		X	X	X
Tech	#samsung6edge	X	X	X		X			X	X
	#acer	X	X	X		X		X	X	X
Places	#ottawa	X	X	X		X		X	X	X
	#portugal	X	X	X		X			X	X
	#hollywood	X	X	X		X		X	X	X
	#france	X	X	X		X			X	X
	#canada	X	X	X		X			X	X
	#greece	X	X	X		X			X	X
	#montreal	X	X	X		X			X	X
Mental Health	#panicattack	X	X	X		X			X	X
	#mentalhealth	X	X	X		X			X	X
Arts	#picasso	X	X	X		X			X	X
	#britishart	X	X	X		X			X	X
Lifestyle	#book	X	X	X		X		X	X	X
	#exercise	X	X	X		X			X	X
	#minimalism	X	X	X		X		X	X	X
	#vinyl	X	X	X		X		X	X	X
	#walking	X	X	X		X		X	X	X
	#thrifting	X	X	X		X			X	X
Learning	#carletonuniversity	X	X	X		X			X	X
	#gradschool	X	X	X		X			X	X
	#chemist	X	X	X		X		X	X	X
	#library	X	X	X		X		X	X	X
	#astronomy	X	X	X		X			X	X
	#physics	X	X	X		x		X	x	X
Misc.	#morning	X	X	X		X		X	X	X
	#night	X	X	X		X		X	X	X
	#holiday	X	X	X		X			X	X

4.4. Summary of the Data Collection Process

In this chapter I outlined the walkthrough methods and the characteristics of classifications derived from the theoretical literature discussed in chapter two. I listed the results of my observations in the six tables above and explained the characteristics of each. I explained how the tables were constructed and how the observations collected from the screen captures (see appendix A) were analyzed according to the classification characteristics of the theories. The following chapter will include the analysis of these observations and will provide responses to the two research questions.

CHAPTER 5: ANALYSIS OF RESULTS

5.1. Analysis of Results

In chapter 4 I provided the results of the walkthrough methods (tables 4-9) and included the data collection schedule as seen in table 3. I also very briefly touched on some of the results of the comparison of the classification characteristics to the screen captures to highlight research problems or clarify the meaning of a characteristic that was otherwise difficult to define. In this chapter I will discuss the results of the comparison of classification characteristics to the screen captures in greater depth. I did this by tallying the tables introduced in the data collection chapter. Tallies provide numerical data that illustrate the relevance of the classification characteristics. I will not discuss all the tallies, but I will highlight the most notable results, focusing especially on those that show connections between different characteristics and those that produced unexpected outcomes such as those that challenged the characteristics deemed essential to formal classifications (e.g. mutual exclusivity). I will also address some of the research challenges encountered when collecting data and discuss these as limitations.

5.2. Results: Bowker and Star, (1999)

As seen in table 4 the most notable results of the analysis are that 18/35 of the hashtag pools were mutually exclusive classifications. All images were visible publicly, which is why I was able to screen capture them. In the hashtag pools, 35/35 suggested a community of practice behind the hashtagged images. Broadly, the community of practice are *Instagram* users. Specifically, #fish showed two niche communities of practice. Chefs and fishers both contributed to the hashtag images of fish. Both belong to the community of *Instagram* users and to their respective communities based on their occupation. Only 7/35 or about 1 in 5 of the hashtag pools showed consistent hashtagging principles meaning that there was a noticeable likeness between the images in the pools. For example, all screen captured images from #chemist were similar in that they depicted chemists, chemistry experiments, or molecular diagrams. When analyzing against Bowker and Star's (1999) characteristics of classification standards, as seen in table 5, none of the screen captures from the hashtag pools demonstrated regulation or enforcement. Regulations are agreed upon principles by which classifications are applied consistently and enforcers ensure the application of the standards to the classification (Bowker and Star, 1999). The criteria of being difficult to change a standard was also not observed either, primarily because no standards were observed beyond common practice.

When analyzing against Bowker and Star's (1999) characteristics of infrastructures as seen in table 6, in which classifications occur, most of the characteristics described parts of the *Instagram* platform. Routines of practice by users as they hashtag images or explore hashtag pools are assisted by the technical resources of the platform. A notable example of this is the user being given suggested hashtags when they are about to upload an image.

This is similar to a standard, but since it is not overseen, it is more a normalized practice and might be more reflective of nudging. Technical resources inform the practice of the user. Hashtag pools showed a variety of uses corresponding to Bowker and Star's (1999) multifunctionality. For example, in the #carletonuniversity pool of images there were a range of photos from graduation photos, to the library services strike, to event promotion. These examples represent celebration, social justice issues, and promotion. Learned usage applied across all hashtag pools because the presence of images within them suggests that the users know how to upload photos and to situate them in a shared context. They understand how to navigate the platform, choose images, edit images, and apply a hashtag to images.

Embeddedness in other infrastructures was not observed. As noted in the data collection chapter, it is more likely that *Instagram* as a platform interoperates with other platforms rather than being nested within them which is the reason for the hashtag in the first place. This is because users can share images from *Instagram* to other platforms suggesting a link between social media ecosystems as discussed through Amrit Tiwana's (2014) framework in the platform economies section (see chapter 2). Finally, the platform was easy to access through the application and the hashtag pools which correspond to Bowker and Star's (1999) suggestion that the infrastructure is "ready to hand." The platform is easy to access because smartphones are typically close to their owners. Accessing the platform means unlocking the phone and pressing the *Instagram* icon. The application keeps users logged in, so they need not sign in each time they want to use it. Furthermore, hashtag pools are easy to access because of the navigation bar at the bottom of each *Instagram* interface. This navigation bar allows users to upload images and allows

them to search for a hashtag pool (see flowlines). Since the navigation bar never goes away, the practice of hashtagging or the exploration of hashtag pools is always at hand.

5.3. Results: Hacking (1986, 1996)

The most notable results of the analysis are the largely descriptive characteristics of the dynamic nominalist framework as seen in table 7. Of the hashtag pools 18/35 showed mutually exclusive classification, which according to Hacking (1996) is a category exclusive to a given kind of entity. The characteristic of mutual exclusivity applied equally between Hacking's (1996) framework and Bowker and Star's (1999) criteria and was not consistent across all pools. Each of the other four components of the dynamic nominalist framework were observed in the hashtag pools. Entities corresponded to photographs as the object being classified by the hashtag.

Experts corresponded to members of the platform who mobilize knowledge to classify. In this research, the knowledge mobilized is usage of the infrastructure which includes how a hashtag is created. As noted with Bowker and Star (1999) (table 6), the presence of the pictures in the pools suggests that the users know how to use the infrastructure. Institutions affirm the classifier's "legitimacy, authenticity, and status as experts" (Hacking, 1996). I equated this characteristic with the terms and conditions that users accept when they sign up for *Instagram*. Each user must accept the terms to use the platform. In doing so, they become members of the community practicing classification. This corresponds to Bowker and Star's (1999) community of practice (table 4). The legal document such as the terms of use refers to them as members. Accepting the terms when signing up legitimates the person as a member of the *Instagram* community. Each user whose image appears in the screen captures went through the same process of legitimation.

If they did not accept the terms, they would not be able to use the platform.

The most notable results of the analysis of the dynamic nominalist engines was corresponding to counting and quantifying to hashtagging. The hashtag pools represent a body of entities. Hashtags suggested to users when they are uploading a photo (e.g. # holiday) include a numerical representation of the body of entities beside it. The numerical representation is a count and a quantification. Furthermore, the numerical quantification of a hashtag pool is seen at the top of each pool. Norms represent an expected outcome of counting and quantifying. The normal outcome of counting and quantifying may be the use of the recommended hashtag provided by the platform. I also applied norms to understanding whether the photos in the screen captures had similarities between (see table 7) them and found that 18/35, or a little more than half of the hashtag pools contained images that shared some like-ness. This matches the observed 51 percent exclusivity of the pools as a classification. The correspondence suggests that like-ness of images and exclusivity of classification are related in creating a classification that is somewhat consistent.

Analyzing some of the other dynamic nominalist engines posed a challenge. Hacking (1996) originally proposed additional engines: correlation, medicalization, biologization, geneticizing, normalization, and bureacratization, and resistance. Hacking (1986, 1996) examined the classification of people in human science, so correlation, medicalization, biologization, geneticizing, normalization described the process. In studies of the classification of non-human entities, these engines do not apply as readily. Lauriault (2012) in her study of cartography and the construction of spaces and modified these engines accordingly. Instead of medicalization, she developed taking action, namely if land was

classified as ecologically sensitive, it was not medicalized, but policy makers or planners would apply remedies or take actions to protect of develop that land according to the class (Lauriault, 2012). Instead of biologization and geneticizing she developed scientification, as the making of a land class for example is derived from a scientific process (Lauriault, 2012). Taking action represents the isolation of a problem, development of a solution, and the application of the solution to the perceived problem. Scientification is the body of knowledge mobilized to achieve the course of taking action. In this sense, Lauriault (2012) notes that scientification determines "a new course of action" in the same way that Hacking intended.

The data collected from the walkthroughs suggest a relationship between the images and their classification because of the correspondence between the norms and the exclusivity of the classification. The presence of a correlation or a statistical relationship between things in *Instagram* would require a more quantitative analysis which is beyond the scope of this thesis, however, some of the work being done by machine learning processes provides some insight (Giannoulakis and Tsapatsoulis, 2016). Since medicalization, biologization, and geneticization did not a apply, I attempted to see if taking action and identification could be seen in the data. The presence of either engine was not found because there were no observations of problems being isolated and actions taken to correct them. Specifically, some users were misclassifying their images by applying hashtags that did not match the image but no one else saw that as a problem or tried to correct it. Misclassification is not addressed in the terms and conditions of the platform, so *Instagram* does not take action either to maintain the boundaries of a class. Other users did not leave comments asking the misclassifier to hashtag their image properly. The result of

taking action is normalization: making "unfavourable deviants as close to normal as possible" (Hacking, 2006). This was not observed because neither scientification nor taking action were observed. Bureaucracy is, for Hacking (1986, 1996) and Lauriault (2012), an engine of administration. Administration is the system that "sees itself as an objective way to determine who needs help" (Hacking, 2006). But for someone or a place to be determined as needing help, there is a need to be first identified as problematic in some way. In this case, none of the images or users, including misclassifiers showed any signs of being identified as problematic or being administered in corrective ways, at least in terms of hashtagging. For the purposes of this project, I left correlation, taking action, scientification inconclusive. Neither normalization nor bureaucracy were observed although the platform does determine how things get done but not the process of classifying.

Resistance was not observed in the hashtag pools, because the identifiable forms of resistance are the voluntary removal of the user from the public community of practice. Users that do not follow *Instagram*'s code of conduct do not appear in the hashtag pools, and that is less a practice related to classification and more an action taking against those who break content rules. One method of resistance is making a profile that was public to a private one, so that the user's images are not made public but the user can still use the platform. I had to look for private profiles elsewhere on the platform. The second form of resistance is deleting one's profile, which eliminates the user from the community of practice and removes all images from the public eye. Self-removal from the community of practice means that the user is not bound by the terms of use or any other aspect of the platform. Again, this is about users and not resistance to a hashtag class.

5.4. Results: Folksonomy Literature

The most notable results of the analysis are that the folksonomy literature largely described hashtags pools on *Instagram*. There are many common entities in the hashtag pools (see table 9). The pool itself constitutes a new form of directory as mentioned in the data collection chapter because they are neither searchable nor structured for user exploration like a table of contents or an index. The user can only scroll down to explore although it is possible to find tree images if one searches the #tree. As observed with Bowker and Star's standards (1999) and Hacking's (1986, 1996) engines as just discussed no overseers govern the classifications. Collaboration was not observed in terms of classifying although contributors are influenced by the platform and each other, and this lack of collaboration is part of Vander Wal's (2007) characteristic and does not align with Golder and Huberman (2005). Vander Wal (2007) suggested that folksonomies are used for user retrieval of the classified entity, but this was not observed, mostly because *Instagram* as a platform does not provide the same functionality as *Flickr*. Moreover, the images a user uploads are present in their profile which is easily accessed using the navigation bar that is present throughout the *Instagram* interfaces but not necessarily in chronological order. Rather than scrolling through a hashtag pool, the user can find their images on their profile. This is easier and less time consuming for the user. Instances of photos with inferred labels, hashtags that are connected explicitly to the image, occurred in 23/35 of the cases of most of the time in the hashtag pools. Most pools contained some images whose presence in that hashtag could be connected even loosely to the subject of the image. For example, photos of breakfast food are loosely connected to #morning in which it was found. All of the hashtag pools exhibited Vander Wal's (2007) (table 9) three data points because each

image was uploaded by a user who assigned the image a hashtag thereby corresponding to 1) a classifier, 2) and entity to classify, and 3) an assigned classification label. Finally, all the hashtag pools were accessible for this research because they occur in a public space, which is the *Instagram* platform. As a member of the community of practice, I was able to access and interact with the hashtag pools.

5.5. Summary

In this chapter I discussed the results of analyzing the classification theories using the visual data sets gathered during my *Instagram* walkthroughs. I discussed which characteristics of the classifications theories are applicable descriptors of hashtagging on *Instagram*. I discussed challenges to this research posed by the formal classification literature, particularly Hacking's (1986, 1996) engines of discovery. In this way, I acknowledged where the literature is applicable, where it does not apply, and which parts of the literature are inconclusive in their applicability in this case. The following chapter discusses conclusions and answers the two research questions this thesis set out to answer.

CHAPTER 6: CONCLUSION

6.1. Introduction

My master's thesis research set out to answer two distinct and, yet, interrelated questions:

1. Are early conceptions of user generated classifications useful descriptors of hashtagging practices on Instagram?
2. Do older classification theories, developed prior to hashtagging practices still apply in a user generated context?

Throughout the thesis, I have explored and grounded these questions in empirical contexts and analyzed my empirical materials using relevant theories. Chapter 5 allowed me to analyze the applicability of classification theories to hashtagging on *Instagram* through descriptive coding and, by so doing, prompted me to answer the two research questions. Overall, some characteristics of formal classification systems helped conceptualize hashtagging practices as formal classification, but more of informal classification characteristics found in folksonomy literature described hashtagging practices consistently.

Of the nine folksonomy characteristics six described each hashtag pool. Ergo, early conceptions of user generated classifications remain useful descriptors of hashtagging practices despite being developed before *Instagram* and the practice of hashtagging. Formal classification theories are partially applicable to hashtagging practices in a user generated context.

Before discussing the contributions of this research to studies of classifications, it is important to reiterate and clarify that the focus of this project was not the user experience but the affordances of the interface when it comes to classification and theory testing. The hybrid walkthrough method provided empirical documentation through the *Instagram* interface to understand classifying through hashtag use. I was able to document the steps and end results of hashtagging as an act of classification, then test established classification theories with documentation. This fulfills the project's motivation to understand where the theory and the reality as it pertains to classification intercept and where they don't. I will address the most relevant observations in the section on Thinking Differently about Classifications.

The underlying rationale of this project relates to the importance of theory testing. Theories are based on what is observed in a historical moment. Time change, as does context as does the advent of new technologies that may change practices which may have an impact on accepted theories. The theories themselves risk becoming fixed in the past, which is problematic especially if those theories are the basis for understanding in the present. Theory testing does not seek to disprove theories but to see where they explain reality and where reality might inform change. A project such as this is meant to open a discussion around theoretical matches and mismatches rather than trying to posit major

amendments to established theories. It does however provide a nuanced discussion about existing theories. With that in mind, I will move from the justification of the project to discussion of its contributions. I will accomplish this by presenting notable mismatches between theory and the observed reality of this study of classification practices on *Instagram* followed by some concluding observations.

6.2. Thinking Differently about Classifications

This section refers to the literature examined in this thesis (see Chapter 2) and points to where I believe some rethinking of classification can be done, especially in the following four ways.

1. *Classification practices are now and always were every day activities, but informal ways to classify were not always recorded as they are now in social media platforms, and documentation leans toward formalization.*

By formality, I refer to the traditional perspective of classification as the consequence of institutional administration. Classifications are not always the bureaucratic processes described by Bowker and Star (1999) and Hacking (1986, 1996) but are also mundane sorting activities that are part of our daily lives and recorded in social media platforms. Although administrative classification continues to exist elsewhere, in *Instagram* hashtagging is incredibly popular, and much classification occurs but it lacks rigorous oversight or formalization. Hacking's (1986, 1996) engines of normalization and bureaucracy were not observed and scientification and taking action developed by Lauriault (2012) from Hacking's (1986, 1996) dynamic nominalism proved inconclusive in their application to hashtagging practices, although there is nudging from the platform which can be conceptualized as an institution in Hacking's framework. The inability to

observe these engines of administration suggests that classifications in *Instagram* are not overseen in a social media context, as stated by Golder and Huberman (2006). Moreover, classification of this type is an every day behavior because social media has become an every day form of engagement for millions of people, which includes recording classes with the use of hashtags when sharing content. It is a banal activity only revealed when thinking, as Bowker and Star (1999) encourage, about classifications as daily practices of sorting things out.

2. *Classifications are used for promotion/visibility, not just archiving and retrieval.*

Classification processes observed with *Instagram* hashtags speaks back to Vander Wal's (2007) suggestion that folksonomies are used to help users retrieve their content. But here, it is also a way to increase the visibility of a user's content. Using a hashtag extends the reach of the user's content beyond their immediate social circle, so their content is clustered with other images that can be seen by other people who do not follow the user. Encouraging the pursuit of visibility and marketability are part of the suggested hashtags users encounter during the upload process. This relates to the environment of presumption and digital labour, which I discuss when rethinking platforms below.

3. *Classifications are more individualized not just institutional/organizational.*

That is, the process of classification is the pursuit of everyone in the platform whereby content is selected, hashtagged, and uploaded. Each person relies on their subjective relationship to the content rather than an external body of standards. A consequence of this was that the hashtag pools lacked the mutual exclusivity that standards tend to produce. In *Instagram*, a small group of experts establishing standards for the rest of the population was not observed. Furthermore, I argued that users are experts in hashtag use, thereby

reconsidering Hacking's (1986, 1996) suggestion that experts are a relatively small body of accredited people. While this is true in some fields such as the sciences, hashtagging opens the definition to include every day people who are not accredited but have a body of common knowledge that constitutes their expertise in a particular form of classification. Classification using hashtags is a folk activity. The individualization of classification means that a new type of expert emerges who mobilizes their technical knowledge of hashtagging according to a personal relationship with their content rather than deferring to an external body of standards.

4. *Classifications are increasingly informal and less bounded.*

Another consequence of the folk activity of using hashtags is that classifications recorded in a platform become increasingly informal and their boundaries become indistinct. Hashtag pools were the opposite of the taxonomy developed by Linnaeus (1735) because they were sprawling repositories of photographs that showed little commonality between images. Generally, there was little order among the contents of the hashtag pools except that all of them are organized chronologically. Users are bound by the terms and conditions of *Instagram* that forbid nude images, for example. Within that perimeter there is greater liberty in classifying than seen with formal experts aligned with formal institutions as exemplified in the dynamic nominalist framework. The use of hashtag classifications for sharing means that images can be hashtagged to appear in multiple pools on *Instagram*, and they can be shared across hashtag-centric platforms such as *Tumblr*. Engaging in classification for sharing revises the traditional boundaries of the classification such as mutual exclusivity, standardization, and expert enforcement (Bowker and Star, 1999). Instead, classifications become unbounded. Although this speaks back to Hacking's (1986,

1996) dynamic nominalism and Bowker and Star's (1999) characteristics, it affirms the suggestion that folksonomies produce a different type of web directory (Choy and Lui, 2006). It is arguable, however, that the notion of a directory suggests a level of organizational structure and searchability not found in hashtag pools. But it does provide another form of knowledge production as the process of formal classification recedes to the background.

6.3. Thinking Differently about Platforms

Chapter 2 established *Instagram* as a platform within the framework of Amrit Tiwana's (2013) platform ecosystem where the ecosystem consists of a platform with complimentary applications that provide access for the end user via an interface such as a smartphone (Tiwana, 2014). In this case, the *Instagram* application allowed me to access the *Instagram* platform via my smartphone. When thinking about the platform ecosystem, I referred to Iansiti and Levien (2004) to modify Tiwana's (2013) framework to consider value creation within the platform and therefore prosumption.

Value creation ensures that platforms attract and keep users. If the platform did not constantly innovate and provide new interaction affordances, or if its core affordances were not in some way unique, users would leave the platform and go elsewhere (Iansiti and Levien, 2004). I noted that the changes brought to *Instagram* by the *Facebook* acquisition, such location tagging, constituted value because they gave the users increased options when sharing their photos. Compared to a formal classification system, a folksonomic classification system privileges the individual giving them the liberty to hashtag based on their subjective relationship with their content. Furthermore, the user can choose a few to a dozen hashtags to add to an image, and they do not have to be in the English language (see

figure 8). The liberal room for decision making makes hashtagging ideal for a keystone firm like *Facebook*. As a core affordance of the platform hashtagging is a strong source of value. Moreover, hashtagging also affords the work of prosumers.

As noted earlier, *Instagram* users are prosumers: they produce information and consume information. Arguably, it is possible for prosumer behavior to occur on a small scale such as a user's immediate social circle. However, another value in hashtagging is that it extends the reach of a user's image beyond their immediate social circle and into the hashtag pool, so other users can see it, like it, and comment on the image. Hashtagging is a classification system not for retrieval but for promotion and engagement, as noted above. Because it extends visibility, the use of hashtags as classifications allows prosumer behavior to occur on a much larger scale than otherwise possible within a platform such as *Instagram*.

However, prosumer behaviour is not bounded by the platform. *Instagram* posts are sharable to other hashtag-centric platforms such as *Tumblr* and the hashtags are kept with the image. If the same hashtags exist on *Tumblr*, the *Instagram* photo is collected in those pool also. Sharing across platforms when using hashtags increases the visibility of a user's content by exposing it to an audience on a different platform. Practices of production, visibility, and consumption are central to each platform, but can extend over more than one platform because of the affordance of being able to share hashtagged content. Cross-platform sharing is therefore an additional value for users of hashtags because they can promote their content beyond their immediate social group, and beyond the public of their home platform. This is especially valuable to any user trying to create a revenue stream by promoting their business through social media. Hashtagging encourages a rethinking of

classifications, but it also encourages a rethinking of value creation and prosumer behavior within and across platforms.

6.4. Limitations/Future Directions

Time was a key limitation of this project. One of its constraints was the sample size. Although 35 hashtag pools yielded 1,050 images, this represents a small portion of the hundreds of thousands of hashtag pools available to users of *Instagram*. It would be beneficial in another study to test the findings with a larger sample size.

Another key limitation of this project was the methodological focus. Since this research aimed to test theory against observed hashtagging practices, the methodological approach was sociological and was oriented toward the interface of *Instagram* and the instrumental usage of the hashtag to classify. Although I used the same interface as did the users, I could only generalize about what is afforded to users. I was able generalize about suggested hashtags and the nature of hashtag pools, but I am unable to speak to the subjective process of hashtagging except from my experience as a researcher. Hashtagging is an individualized practice and may be a communal one in which users rely on their subjective relationship with their content and the content of others. A future project could undertake a survey or even ethnography of users to better understand how they sort things out. A generalization of the hashtagging process from the user perspective could then be established. The result would be a more complete picture of hashtagging as an instrumental and technological classification practice as well as a user-generated classification practice undertaken by members of a community of practice.

Another advantage to speaking to users is understanding whether classifications "make up" entities the way Hacking (1986,1996) suggests. Hacking (1986,1996)

understands making up people, which is the way classifications shape the public perceptions of classified people. Lauriault (2012) revised dynamic nominalism slightly to show how mapping makes up spaces, for example, how mapping can shape the public perception of a geographic space. Perhaps hashtag pools shape the perceptions of users around certain subjects. The pool of images represents a body of data about the subject signified by the hashtag, so #britishart is a collection of images that shapes the public perception around British art. Administrative classifications make up people, mapping classifications make up spaces, and perhaps hashtags make up subjects. Rather than being the activity of a formal organizations as seen with Hacking (1986,1996) and Lauriault (2012), hashtagging is an informal user-generated practice. Hashtagging might represent an informal process of users making up subjects to other users, so focus groups or surveys might present "making up" in a user-generated context. Understanding the implications of that conclusion would be beneficial to better understand knowledge and subject production.

Instagram as a platform provides a limitation because it updates frequently. This means that hashtagging in 2018 as represented in this research may not be the same as hashtagging on *Instagram* in a year or two. None of the platform updates that occurred during this research impacted the study, but the platform could provide updates that influence how hashtags are suggested to users, for example, he ability to search hashtag pools did not exist at the time of writing, but that could change in the future. Platform architecture changes in ways unanticipated because platform ecosystems are in competition with each other to attract and hold users (Iansiti and Levien, 2004). Revisiting this research would test its findings against what is becomes normal in the future to re-assess the results of this thesis. This is the same reason I tested theories written in the early 2000s against

what is contemporary practice today. Testing research is an important mode of conducting research. Self-reflexivity within a field of study ensures that the body of knowledge about a subject, especially one that is new the way *Instagram* is a topic to Communications and Media Studies, and to assess how knowledge is produced in these large social media platforms has social implications, beyond the scope of this study, but suggests, that the crowd as opposed to formalized classification is creating new 'subjects'. Perhaps that is more consensus based or perhaps it is diminishing science. Those would be important philosophical questions worthy of further study.

APPENDIX

Enclosed CD with visual data sets.

BIBLIOGRAPHY

Barthes, R. (1983). Mythologies (A. Lavers, Trans.). New York: Hill and Wang.

Bowker, G., & Star, S. (1999). Sorting things out (1st ed.). Cambridge, Massachusetts: MIT Press.

Bryman, A., Bell, E., & J. Teevan, J. (2012). Social research methods (3rd ed., p. 133). Oxford University Press Canada

Choy, S., & Lui, A. (2006). Web Information Retrieval in Collaborative Tagging Systems. 2006

Cottle, L (2018). Retrieved from https://play.google.com/store/apps/details?id=io.storysave.android

Light, B., Burgess, J., & Duguay, S. (2016). The walkthrough method: An approach to the study of apps. New Media & Society, 146144481667543. http://dx.doi.org/10.1177/1461444816675438

Gruber, T. (2007). Ontology of Folksonomy. International Journal on Semantic Web and Information Systems, 3(1), 1-11. http://dx.doi.org/10.4018/jswis.2007010101

Haklay, M., 2013, Citizen Science and Volunteered Geographic Information – overview and typology of participation in Sui, D.Z., Elwood, S. and M.F. Goodchild (eds.), 2013.Crowdsourcing Geographic Knowledge: Volunteered Geographic Information (VGI) in Theory and Practice. Berlin: Springer. p. 105-122

Hacking, I. (2006). "Kinds of people: moving targets". In British Academy Lecture. London: The British Academy. Retrieved 18 September 2017, from https://www.britac.ac.uk/sites/default/files/hackingdraft.pdf

Hacking, I. (1986). "Making up People." In T. Heller et al., Reconstructing Individualism. Stanford: Stanford University Press.

Happy Birthday, *Instagram* Stories! (2018). Retrieved from https://*Instagram* press.com/blog/2018/08/02/happy-birthday-*Instagram*-stories/

Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. Qualitative health research, 15(9).

Golder, S., & Huberman, B. (2006). Usage patterns of collaborative tagging systems. Journal of Information Science, 32(2), 198-208.

Gruber, T. (2007). Ontology of Folksonomy. International Journal on Semantic Web and Information Systems, 3(1), 1-11. doi: 10.4018/jswis.2007010101

Iliadis, A. J. (2016). A black art: ontology, data and the Tower of Babel problem (PhD). Purdue University.

Iansiti, M., & Levien, R. (2004). Strategy as Ecology. Retrieved from https://hbr.org/2004/03/strategy-as-ecology

Instagram. Retrieved from https://www.*Instagram*.com/

Instagram monthly active users 2018 | Statistic. (2018). Retrieved from https://www.statista.com/statistics/253577/number-of-monthly-active-*Instagram*-users/

Kenney, M., & Zysman, J. (2016). The Rise of the Platform Economy. Issues in Science and Technology, Spring 2016.

Kuhn, T. (1970). The structure of scientific revolutions (2nd ed.). Chicago, Ill.: The University of Chicago Press.

Lampland, M., & Star, S. (2009). Standards and their stories. Ithaca: Cornell University Press.

Lauriault, T. (2012). Data, Infrastructures and Geographical Imaginations (Ph.D.). Carleton University.

Lewis, C., & Rieman, J. (1993). Task-centered user interface design. Boulder, Colo.: University of Colorado, Boulder, Dept. of Computer Science.

Light, B., Burgess, J., & Duguay, S. (2016). The walkthrough method: An approach to the study of apps. New Media & Society.

Linnaeus, C. (1735). Systema naturae (1st ed.). Leiden: Lugduni Batavorum. (Haak).

Maloof, J. (2018). Vivian Maier. Retrieved from http://www.vivianmaier.com/

MacArthur, A. (2018). Where Hashtags Really Came from and What They Do. Retrieved from https://www.lifewire.com/history-of-hashtags-3288940

Nielsen, J., & Molich, R. (1990). Heuristic evaluation of user interfaces. Conference on HumanFactors in Computer Systems. New York: ACM

Ohkura, T., Kiyota, Y., & Nakagawa, H. (2006). Browsing system for weblog articles based on automated folksonomy. Proceedings of The WWW 2006 Workshop on The Weblogging Ecosystem: Aggregation, Analysis and Dynamics, At WWW.

Oikarinen, J., & Reed, D. (1993). RFC 1459 - Internet Relay Chat Protocol. Retrieved from https://tools.ietf.org/html/rfc1459#section-4.2.1

The PDP-11 Assembly Language. (2011). Retrieved from https://programmer209.wordpress.com/2011/08/03/the-pdp-11-assembly language/

Thomsen, S., Straubhaar, J., & Bolyard, D. (1998). Ethnomethodology and the study of online communities: exploring the cyber streets. Information Research, 4(1).

Peterson, A. (2014). *Snapchat* agrees to settle FTC charges that it deceived users. Washington Post.

Seyfert, R., & Roberge, J. (2016). Algorithmic cultures. New York: Routledge.

Saldaña, J. (2013). The coding manual for qualitative researchers (2nd ed.). Sage Publications.

Sosna, T. (2018). Tina Sosna (@tinasosna) • *Instagram* photos and videos. Retrieved from https://www.*Instagram*.com/tinasosna/?hl=en

Stallkamp, J., Schlipsing, M., Salmen, J., & Igel, C. (2012). Man vs. computer: Benchmarking machine learning algorithms for traffic sign recognition. Neural Networks, 32, 323 332. doi: 10.1016/j.neunet.2012.02.016

Striphas, T. (2015). Algorithmic culture. European Journal of Cultural Studies, 18(4-5), 395-412.

Systrom, K. (2011). What is the genesis of *Instagram*? Retrieved from https://www.quora.com/*Instagram*-company/What-is-the-genesis-of-*Instagram*

Tiwana, A. (2014). Platform ecosystems. Waltham, MA: Morgan Kaufmann.

Toffler, A. (1980). The Third Wave. New York: Bantam Books.

Topic: Smartphones. (2018). Retrieved from https://www.statista.com/topics/840/smartphones/),

Upbin, B. (2012). *Facebook* Buys *Instagram* for $1 Billion. Smart Arbitrage. Retrieved from https://www.forbes.com/sites/bruceupbin/2012/04/09/*Facebook*-buys *Instagram*-for-1 billion where's-the-revenue/#271290a74b8a

Vander Wal, T. (2007). Folksonomy. Vanderwal.net. Retrieved 19 September 2017, from http://www.vanderwal.net/folksonomy.html

VSCO (2018). Retrieved from https://play.google.com/store/apps/details?id=com.vsco.cam

Worm, O. (1655). Museum Wormianum (1st ed.). Amsterdam: Lvdovicvm & Danielem Elzevirios

Zhang, L., Wu, X., & Yu, Y. (2006). Emergent Semantics from Folksonomies: A Quantitative Study. Journal on Data Semantics, VI (LNCS 4090).