JISTaP http://www.jistap.org
Journal of Information Science Theory and Practice
eISSN : 2287-4577 pISSN : 2287-9099

Research Paper
J Inf Sci Theory Pract 6(2): 46-61, 2018
https://doi.org/10.1633/JISTaP.2018.6.2.4

Content Description on a Mobile Image Sharing Service: Hashtags on Instagram

Isabelle Dorsch*

Department of Information Science, Institute of Linguistics and Information Science, Heinrich Heine University Düsseldorf, Düsseldorf, Germany
E-mail: isabelle.dorsch@hhu.de

ABSTRACT

The mobile social networking application Instagram is a well-known platform for sharing photos and videos. Since it is folksonomy-oriented, it provides the possibility for image indexing and knowledge representation through the assignment of hashtags to posted content. The purpose of this study is to analyze how Instagram users tag their pictures regarding different kinds of picture and hashtag categories. For such a content analysis, a distinction is made between Food, Pets, Selfies, Friends, Activity, Art, Fashion, Quotes (captioned photos), Landscape, and Architecture image categories as well as Content-relatedness (ofness, aboutness, and iconology), Emotiveness, Isness, Performativeness, Fakeness, "Insta"-Tags, and Sentences as hashtag categories. Altogether, 14,649 hashtags of 1,000 Instagram images were intellectually analyzed (100 pictures for each image category). Research questions are stated as follows: RQ1: Are there any differences in relative frequencies of hashtags in the picture categories? On average the number of hashtags per picture is 15. Lowest average values received the categories Selfie (average 10.9 tags per picture) and Friends (average 11.7 tags per picture); for highest, the categories Pet (average 18.6 tags), Fashion (average 17.6 tags), and Landscape (average 16.8 tags). RQ2: Given a picture category, what is the distribution of hashtag categories; and given a hashtag category, what is the distribution of picture categories? 60.20% of all hashtags were classified into the category Content-relatedness. Categories Emotiveness (about 4.38%) and Sentences (0.99%) were less often frequent. RQ3: Is there any association between image categories and hashtag categories? A statistically significant association between hashtag categories and image categories on Instagram exists, as a chi-square test of independence shows. This study enables a first broad overview on the tagging behavior of Instagram users and is not limited to a specific hashtag or picture motive, like previous studies.

Keywords: image indexing, Instagram, knowledge organization, folksonomy, user behavior, tagging behavior

Open Access

Accepted date: June 3, 2018
Received date: March 12, 2018

***Corresponding Author:** Isabelle Dorsch
Research Associate
Department of Information Science, Institute of Linguistics and Information Science, Heinrich Heine University Düsseldorf, Universitätsstraße 1, D-40225 Düsseldorf, Germany
E-mail: isabelle.dorsch@hhu.de

1. INTRODUCTION

"A picture is worth a thousand words"; but how would one describe a picture with a handful of terms? What elements of the picture would one want to represent? What kind of terms would one choose?

These questions always arise in the context of knowledge representation, or to be more precise, in indexing as an application of knowledge representation. In order to capture the content of a document like a picture, in professional information services indexing deals with the representation of single objects through controlled concepts (Stock & Stock, 2013). Besides a controlled keyword assignment by human or machine indexers, the free allocation of keywords by everyone has obtained a huge impact since the beginning of the Web 2.0. Nowadays, folksonomies (Peters, 2009) have become indispensable for the web. They can be found in many areas, like for example Twitter (in the field of microblogging), Mendeley (as a social reference management system), Flickr (as a photo management and sharing application), or Instagram (for sharing pictures and videos within a social networking mobile application). Instagram is becoming more and more popular as a social photo and video sharing application. Here, the free allocation of keywords is done by the assignment of hashtags.

How do users index a picture on Instagram with at most 30 hashtags?

What is new in this article? We found out for the first time that different picture categories (as, for instance, Architecture, Fashion, or Food) exhibit both, different numbers of hashtags as well as—more important—different kinds of hashtags (as Content-related tags or tags expressing Emotiveness, Fakeness, "Insta" aspects, Performativeness, and entire Sentences). Furthermore, it is demonstrated that there is an association between image categories and hashtag categories.

1.1. Indexing of Pictures by Folksonomies

Indexing of pictures can be conducted through a content-based or a concept-based approach (Rasmussen, 1997). Extracting features like color, shape, and texture belong to content-based indexing, whereas concept-based indexing requires a textual description (Lancaster, 2003; Rasmussen, 1997). Several methods and models were created for both approaches and the topic was widely discussed (Jörgensen, 2002, 2003; Enser, 2008). Above all, concept-based indexing refers to the ofness and aboutness of a picture (Shatford, 1986), which in turn are based on Panofsky's levels of meaning in art (Panofsky, 1955).

The first definition of the concept of "folksonomy" (being descended from "folk" and "taxonomy") goes back to Vander Wal and was stated in a blog entry of Smith (2004). Vander Wal (2007) chose this term for the free allocation of keywords—in a folksonomy called "tags"—by users in Web 2.0 services like Flickr or Del.icio.us. A folksonomy (Peters, 2009) is the result of the total quantity of all assigned tags in an information service.

Besides tags, the concept of "hashtags" exists. Hashtags are a composition of # and a character string, like for example #summer. On Instagram, it is even possible to use a hashtag emoji like #☺. "[B]y using the # character to mark particular keywords, …users communicate a desire to share particular keywords folksonomically" (Halavais, 2014, p. 36). As Halavais (2014) stated, some suggest the originator of the hashtag is Messina (2007), since he tweeted in August 2007 on Twitter "how do you feel about using # (pound) for groups. As in #barcamp [msg]?" (Messina, 2007). After this posting, the use of hashtags was established and was implemented on several platforms (Halavais, 2014). "Hashtags represent a way of indicating textually keywords or phrases especially worth indexing" (Halavais, 2014, p. 36). So hashtags are "user-generated metadata," too.

Tagging behavior in folksonomies was already observed several times (Golder & Huberman, 2006; Daer, Hoffman, & Goodman, 2014). Flickr, launched in February 2004 (Kremerskothen, 2012), applied one of the first well-known folksonomies for photos and therefore also for image and tag research (Beaudoin, 2007; Nov, Naaman, & Ye, 2008; Hollenstein & Purves, 2010; Rorissa, 2010; Stvilia & Jörgensen, 2010). For Facebook images, Denton, Weston, Paluri, Bourdev, and Fergus (2015) also developed and analyzed models for user conditional hashtag prediction. Besides Flickr, further photo services like Pinterest or Instagram were developed over time.

1.2. Instagram

During the last few years, Instagram has become more and more popular as a mobile social networking application for sharing photos and videos in various ways. It was launched in October 2010 with 25,000 signed-up users for the first day (Instagram, 2010). The number of users and functions has grown over time. The app's monthly active users numbered more than 800 million as of March 2018 (Instagram, 2018). The central point is still to share photos and videos.

Fig. 1 shows a typical posting of a publicly posted Instagram picture. Beside the photo, the creator of the posting (in this case *kajaf*) has the possibility to add

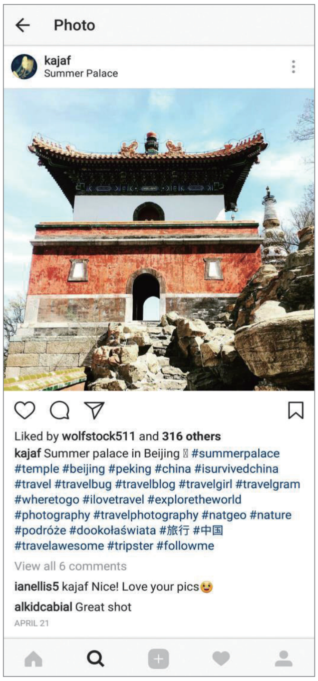

Fig. 1. Posted architecture picture by kajaf on Instagram.

a description text and up to 30 distinct hashtags. The uploaded photo can be liked, commented on (also with hashtags), shared, or favored by Instagram users. The hashtags enable users to make the posting searchable under the respective chosen term. For example, the searching for #temple lists all content tagged with #temple.

Having a closer look at the hashtags in Fig. 1, we are able to identify different sorts of hashtags. #temple or #bejing clearly describe the image's content. However, #Isurvivedchina is a complete sentence, #photography puts the document into a media category, #followme is a request to do something and, finally, #travelgram refers to Instagram. Obviously, there are different categories of hashtags on Instagram.

1.3. State of Research on Instagram

Research on Instagram lays its focus on different aspects like for example content analysis or hashtag use, which are summarized in the following paragraph.

Sheldon and Bryant (2016) surveyed 239 college students about their motives for using Instagram. Five distinct post types like advertising or information were analyzed by Coelho, de Oliveira, and de Almeida (2016) with respect to their likes and comments. The investigated posts depicted one of five business segments (food, hairdressing, ladies' footwear, body design, and fashion gym wear). Giannoulakis and Tsapatsoulis (2015) investigated through an online survey whether other Instagram users would use the same hashtags like the picture owner (based on a set of 30 Instagram images and their most respectively descriptive hashtags). How hashtags are used to retrieve information in social media like Instagram was the subject of research by Buarki and Alkhateeb (2018).

Nashmi (2018) conducted a content analysis with 1,000 Instagram pictures which were posted up to four days after the Charlie Hebdo incident of January 2015 in order to investigate visual post changes. Two content analyses about posted Instagram photos by the 10 largest fast food companies as well as photos from users who posted about the companies were performed by Guidry, Messner, Jin, and Medina-Messner (2015). Holmberg, Chaplin, Hillman, and Berg (2016) categorized food images in order to analyze in which way and what kinds of food 14-year-old adolescents present on Instagram. To receive only pictures from 14-year-olds, they investigated appropriated user profiles by using the hashtag #14år ("14 years"). In order to characterize the dietary trend of cheat meals, Pila, Mond, Griffiths, Mitchison, and Murray (2017) conducted a thematic content analysis of 5,600 tagged #cheatmeal Instagram postings. Both photographic and textual elements of the cheat meal postings were analyzed. Also referring to the subject of food, Ye, Hashim, Baghirov, and Murphy (2018) investigated the content of 1,382 Instagram postings tagged with #Malaysianfood. Therefore, the hashtag descriptions were placed into the categories informative or emotional as well as positive or negative.

Marcus (2016) performed an image content analysis about pro-anorexic and fat acceptance communities in Instagram. She identified several popular hashtags regarding those topics in order to obtain 400 suitable pictures for each community. 1,967 Instagram pictures tagged with #swisher (a popular cigar brand in the United States) were thematically analyzed by Allem, Escobedo, Chu, Cruz, and Unger (2017) to support future tobacco control efforts and education campaigns as well as to understand health behavior through social media data. Sensitive self-disclosures and their responses on Instagram were analyzed by Andalibi, Ozturk, and Forte (2017) through a three-phase methodology: "Phase I establishes the content of depression-tagged posts on Instagram; Phase II investigates the kinds of responses these posts attract; and Phase III examines relationships between the kinds of posts and the kinds of responses they attract" (Andalibi et al., 2017, p. 1489). Depression tagged postings were identified through the hashtag #depression. Their final sample consists of

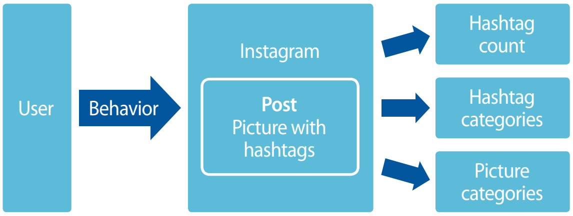

Fig. 2. Research model.

788 images and captions. Souza et al. (2015) studied Instagram selfies in context of their characteristics (e.g., demographic information, distribution of post frequency, likes, comments, etc.). An examination of a variety of selfie hashtags with respect to their popularity was conducted, too. Oh, Lee, Kim, Park, and Suh (2016) investigated Instagram users' "participatory hashtag practices" regarding the Weekend Hashtag Project. "Participatory hashtag practices" describes a "hashtagging phenomenon, where a certain user account suggests a hashtag to its followers and promotes them to upload photos suitable to the hashtag" (Oh et al., 2016, p. 1281). Public Instagram photographs tagged with #funeral were analyzed by Gibbs, Meese, Arnold, Nansen, and Carter (2015). In the field of sports, Pegoraro, Comeau, and Frederick (2018) also investigated Instagram pictures which were tagged with certain hashtags. They conducted a content analysis for images tagged with #SheBelieves (n=629) or #FIFAWWC (n=706). Likewise in the context of sports, but with respect to postings tagged with the hashtag #fitspo, Carrotte, Prichard, and Lim (2017) analyzed 415 postings from Instagram, Tumblr, Facebook, and Twitter, whereas with 360 posts the majority were found in Instagram. Veszelszki (2016) investigated 400 Instagram postings tagged with #time, #truth, or #tradition in respect of the relationship between the image and text(hashtag).

The following study is one of the bases for this research project. In an empirical study about content and users on Instagram, Hu, Manikonda, and Kambhampati (2014) determined 8 popular photo categories (friends, food, gadget, captioned photo, pet, activity, selfie, and fashion) for Instagram and 5 different types of user clusters. Furthermore, they found that the number of a user's followers does not depend on the users' posted photos on Instagram (Hu et al., 2014).

1.4. Tagging Behavior on Instagram

Previous studies on Instagram hashtags investigate one specific hashtag or some hashtags referring to a certain topic. This study analyzes the tagging behavior of Instagram users in a wider field by use of a content analysis (Krippendorff, 2004). Therefore the authors' hashtags, assigned to Instagram images of 10 different picture categories, were analyzed regarding different hashtag categories (Dorsch, Zimmer, & Stock, 2017). The main research questions are:

RQ1 : Are there any differences in relative frequencies of hashtags in the picture categories?
RQ2 : Given a picture category, what is the distribution of hashtag categories; and given a hashtag category, what is the distribution of picture categories?
RQ3 : Is there any association between image categories and hashtag categories?

The research aspects of this study are represented in Fig 2. The tagging behavior will be investigated for Instagram users regarding their Instagram picture postings. The factors of a posting that are to be analyzed are the hashtag counts of a picture and the categories of the assigned hashtags, as well as the subject categories of the pictures. The study deepens our understanding about tagging behavior on social media with the example of Instagram. Additionally, we are going to gain insights regarding knowledge representation by layman indexers.

2. METHODS

Content analysis (Krippendorff, 2004) provides the option to analyze content systematically. To analyze the tagging behavior of Instagram users, pictures and hashtags were coded—through content analysis—into specific picture and hashtag categories.

2.1. Selection of Picture Categories

The collected 1,000 Instagram pictures conform to the following selected 10 image subject categories (Table 1). Thereby, each picture category contains 100 pictures. The categories are based on Hu et al. (2014) who detected empirically by cluster analysis and qualitative analysis (done by two human coders) 8 popular general Instagram photo categories in their study about Instagram photos and users. Categories 1, 4 to 7, and 9 to 10 originate from them, but their category descriptions were complemented or partly modified for this work after our pretests, as can be seen in Table 1.

It should be noted that the category Gadget from Hu et al. (2014) is missing, because it was omitted after the pretest of this study. The test demonstrated that it is too

Table 1. Analyzed picture categories with the descriptions of their image subjects

Picture category	Description: The picture depicts…
Activity	… "both outdoor & indoor activities, places where activities happen, e.g., concert" (Hu et al., 2014, p. 597). This does not include landscape and architecture related pictures.
Architecture	… all architecture related content, except if places where activities happen are in the foreground. Art-related work is excluded.
Art	… drawings, paintings, sculptures, land art, crafted stuff, tattoos, and any other art-related content. Photo shots are not included, except where they show the art, like for example a photo of a sculpture. The original author of the artwork can be the profile owner of the posted picture or a foreign person.
Captioned Photo	… "pictures with embed text, memes, and so on" (Hu et al., 2014, p. 597). Besides this, the picture also has to show graphical content (e.g., persons, items, landscapes, etc.), because pure text does not allow analysis of the picture content in terms of the chosen hashtags.
Fashion	… "shoes, costumes, makeup, personal belongings, etc." (Hu et al., 2014, p. 597).
Food	… "food, recipes, cakes, drinks, etc." (Hu et al., 2014, p. 597).
Friends	… "users posing with others friends" (Hu et al., 2014, p. 597) or only the friends of a user. At least one person must be depicted in the picture. Contrary to the definition of Hu et al. (2014, p. 597), faces do not have to be seen.
Landscape	… all nature related content, except if places where activities happen are in the foreground. Art-related work is excluded.
Pet	… "animals like cats and dogs which are the main objects in the picture" (Hu et al., 2014, p. 597).
Selfie	… "self-portraits; only one human face is present in the photo" (Hu et al., 2014, p. 597). It must be identifiable that the photo was taken by the person in the picture ("mirror pictures" are also allowed).

Table 2. General coding rules for categorizing Instagram posts

#	Rule
R1	The basic requirement for each picture is that they contain their respective category hashtag. Without this hashtag, they are not admitted for the categorization.
R2	All category pictures were chosen so that they can be treated as prototypes for their respective categories and are therefore always classified into one category only. With respect to the diversity of picture motifs in social media like Instagram, it cannot always be excluded that motif borders blur. For example, a picture predominantly depicts the content of its respective category, but it can also contain some subsidiary further content (like for example a pet in the background of a selfie). For that reason preference rules were developed. They apply for those cases and ensure a clear picture categorization.

Table 3. Used top picture category hashtags and alternative hashtags with the total count of their assigned hashtags in Instagram (valid as of December 18, 2016)

Picture category	Top picture category hashtags	Number of tagged media for the top hashtag	Alternative hashtag(s) and the number of tagged media
Activity	#activity	0.7 m	#activities (0.4 m)
Architecture	#architecture	45.8 m	#architectures (0.2 m) #architectureporn (3.2 m) #cityscape (4.2 m) #cityscapes (0.4 m)
Art	#art	211.3 m	#artwork (26.6 m)
Captioned Photo	#quote	38.2 m	#quotes (33 m) #meme (14.7 m) #memes (10.4 m)
Fashion	#fashion	324 m	#fashions (0.6 m)
Food	#food	198 m	#foods (8.5 m) #foodporn (106.5 m)
Friends	#friends	247.7 m	#friend (47 m) #friendship (28.9 m)
Landscape	#landscape	46.5 m	#landscapes (3.1 m) #countryside (5.6 m) #countrysides (3.4 m)
Pet	#pet	39.7 m	#pets (25.5 m) #animal (29.1 m) #animals (25.8 m)
Selfie	#selfie	281.7 m	#selfies (17.1 m)

m, million.

broadly defined for collecting pictures according to the chosen method of this study. Partially, too many pictures were tagged which had nothing to do with a gadget or had violated the rules of the codebook (especially general picture category codebook rule 2, R2) (Table 2). Since those #gadget pictures comprised more picture motifs, a noticeable amount of these pictures (compared to the other pretest picture categories) could not be sorted precisely into the picture category Gadget and therefore were not used for a pretest.

In particular, the categories Fashion and Gadget exhibit strong overlaps during the pretest. To select narrower (hashtag) terms of Gadget to serve as picture categories could be a possible solution for the stated problems. Such a solution was discarded for this study because this category approach would stand in contrast to the other broadly diversified categories.

Categories 2, 3, and 8 were created for this study and are also widespread subjects on Instagram, as their hashtag number in Table 3 displays. However, why had any categories to be chosen? The categories represent a selection of possible subjects; and the study does not claim to cover all possible types of picture subjects. Rather, they provide a balanced distribution of data. Furthermore, a strict picture segmentation into categories, with one (main) picture motif, enables the possibility to make clear statements about the categories and possible category relations. Due to limited coder capacities, only these 10 categories could be selected.

2.2. Selection of Hashtag Categories

In knowledge representation, different concept categories can be derived for image indexing. Concepts describing aboutness and ofness are most popular in indexing non-textual documents. Those categories serve as a basis for the analysis of the Instagram hashtag tagging behavior in this study. The difference to the regular usage in indexing is that these categories were considered from a retro perspective. They do not help an indexer to decide whether a certain index term had to be chosen or not (and consequently to guarantee an appropriate capture of a document), but to categorize empirically the collected hashtags. Existing tagging categories influenced the creation of the categories, too. The following categories were developed for categorization: Content-relatedness (including ofness, aboutness, and iconology), Emotiveness, Isness, Performativeness, Fakeness, "Insta"-Tags, and Sentences.

2.2.1. Content-related tags

Content-related tags involve everything a picture directly or abstractly depicts. This category refers to the definitions of aboutness and ofness in pictures which in turn are based on Panofsky's (1955) three levels of meaning in the visual arts. These levels are called per-iconographic, iconographic, and iconologic. They relate to different meaning aspects in an artwork. The pre-iconographic level refers to practical experience and can be factual or expressional. For example, factual would be the representation of natural objects like "human beings, animals, plants, houses, tools and so forth," or expressional "by identifying their mutual relations as events; and by perceiving such expressional qualities as the mournful character of a pose or gesture, or the homelike and peaceful atmosphere of an interior" (Panofsky, 1955, p. 28). In addition, the iconographic level comprises specific themes and concepts instead of basic objects and events. Here one is on the level of literary knowledge, as for instance

> …by realizing that a male figure with a knife represents St. Bartholomew, that a female figure with a peach in her hand is a personification of veracity, that a group of figures seated at a dinner table in a certain arrangement and in certain poses represents the Last Supper, or that two figures fighting each other in a certain manner represent the Combat of Vice and Virtue (Panofsky, 1955, pp. 28-29).

Finally, iconology declares the intrinsic meaning or content (e.g., *Weltanschauung*) and addresses therefore symbolic values: "It is apprehended by ascertaining those underlying principles which reveal the basic attitude of a nation, a period, a class, a religious or philosophical persuasion-qualified by one personality and condensed into one work" (Panofsky, 1955, p. 30). For example, to understand Leonardo da Vinci's fresco "Il Cenacolo"/"L'Ultima Cena" (The Lord's Supper) as culture of the Italian high renaissance, is an association to the level of iconology.

This classification of artworks became transferred to a general consideration of subject kinds in a picture (Shatford, 1986; Shatford Layne, 1994). A picture can be *of* something and also *about* something. In information science, this means that the first two of Panofsky's three levels (1955) correspond particularly to an ofness and aboutness. For the first level, the factual pre-iconographical objects comprise the ofness (also called "generic Of") whereas the expressional pre-iconography denotes the aboutness. Ofness of iconography (also called "specific Of") includes more "specifically what a picture is Of" (Shatford, 1986, p. 44); iconographical aboutness, on the other hand, denotes allegories, personifications, and symbols. Since Shatford (1986) discusses Panofsky's levels in the context of image indexing, iconology is excluded, because it addresses the interpretational aspects of a picture. It would not be possible to index it consequently. It is also possible to subdivide the different kinds of ofness and aboutness with respect to the facets time, space, activities, events, and objects to get a more specific classification (Shatford Layne, 1994).

In her explanations about the subject access to art images Shatford Lane's (2002) remarks that it is helpful to distinguish between ofness and aboutness in order to determine which index terms provide which kind of subject access (e.g., the difference between pictures that depicts death and are therefore of it versus pictures depicting death symbolically and thus are about death). Since the terms are already indexed and nobody has to decide whether an ofness or aboutness term is better, this is not necessary for an analysis of hashtags. Even so, it would be interesting to distinguish between them.

Furthermore, Shatford Lane (2002) explains that there is a wide reach between generic ofness and specific ofness. It is not possible to consider them dichotomously. Also, aboutness is not always determinable, obviously. In this case, she refers only to art images, but she already addresses this difficulty in earlier works for all kinds of images except abstract artworks (Shatford, 1986).

Subjectivity always occurs when pictures are indexed by ofness and aboutness concepts (Shatford, 1986). According to Turner (1995), the distinction between ofness and aboutness for the purpose of indexing is often impossible. Already Panofsky (1955) stated that the boundaries between pre-iconography and iconography could blur.

Besides, aboutness also depends on the point of view (Maron, 1977). For example, Maron (1977) distinguishes between S-about (subjective about; the relation between a document and its reader's experiences), O-about (objective about; "an external or observer's point of view" [Maron, 1977, p. 41]), and R-about (retrieval about; "the information searching behavior of a class of individuals" [Maron, 1977, p. 41]). Ingwersen (2002) also stated different kinds of aboutness: Author aboutness provides the content of an author's document, indexer aboutness is the interpretation of the content by an indexer, request aboutness is the information need formulated as a search argument, and user aboutness is the interpretation of the content by a user. This study (Rondeau, 2014) addresses the aboutness, assigned from different subjective users' perspectives.

The codebook was developed to be as objective as possible, so that coding is repeatable for anyone. Therefore we did not differentiate between all kinds of introduced ofness and aboutness. They were merged together into the category Content-relatedness. Moreover, and as stated above, such a distinction is especially important for deciding whether a term should be indexed or not for a certain subject area and information need (Shatford, 1986). This difficulty is dropped when categorizing hashtags because the terms are already assigned. The fact that some of the hashtags refer (rather) to the author description of the picture is another reason for the combination to only one category.

Although the subjective interpretational level of iconology (Panofsky, 1955) is excluded for indexing in information science (Shatford, 1986), it can be tagged by users of a folksonomy (Stock & Stock, 2013). During the coding, we detected some hashtags that can be best described as iconological. Those hashtags were also sorted into the category Content-relatedness since they are relating to some "content." "User-specific tags [describing] or [evaluating] a document only from the user's very own perspective," so that some tags "are virtually meaningless to anybody except their creators" (Pluzhenskaia, 2006, p. 23) are likewise counted as content-related hashtags.

2.2.2. Emotiveness

The category Emotiveness comprises emotional hashtags. But how is the concept of emotions defined? As summarized by Knautz (2012) and Siebenlist (2013) there exists a variety of definitions. Several researchers tried to define a set of basic emotions (Ortony & Turner, 1990). Basic emotions used in this study are love, happiness, fun, surprise, aspiration, sadness, anger, disgust, fear, and shame, and were adopted from Siebenlist (2013). Besides this, all possible manifestations of emotions a user could formulate were considered as emotive hashtags.

Looking at the categories, one could argue against a separation of emotional hashtags. Formally, they belong to the content-related tags and as a consequence to the category Content-relatedness. Likewise Shatford (1986) assigns emotions as aboutness. However, in Folksonomies, those tags were analyzed separately (Kipp, 2006; Beaudoin, 2007; Yanbe, Jatowt, Nakamura, & Tanaka, 2007; Schmidt & Stock, 2009; Mohammad & Kiritchenko, 2015; Ye et al., 2018). For Twitter, Mohammad and Kiritchenko (2015) also pointed out that emotional hashtags can function as labels of emotions in tweets. To treat emotional hashtags separately in order to get distinctive information and because such a distinction is not problematic for the objectivity of the codebook, emotional hashtags like #love, #happy, or #fun were separated to the category Emotiveness.

2.2.3. Sentences

If the separation of emotiveness tags from content-related hashtags is questionable, the same applies to sentences. Sentences represent content, but with the function of information condensation and not as information filters. They cannot be counted as part of an indexing process like single concepts and phrases, but are a part of abstracts (Stock & Stock, 2013). For that reason, they had to be separated from the content-related terms and phrases. Only whole sentences (containing subject, verb, and object) and their abbreviations (e.g., wiwt = what I wore today) were considered for this category. If a Sentences hashtag contains overlaps with other hashtag categories, the hashtag will always be categorized into the category Sentences, since the tag primarily represents a sentence.

2.2.4. Isness

The category Isness originates from Ingwersen (2002). It comprises all non-topical features of a document, e.g., for pictures all "technical aspects of the photograph (e.g., camera type, length of exposure, aperture)" (Stock & Stock, 2013, p. 616). Further specific examples for this study, corresponding to Ingwersen (2002), are the (user)name of the person who took the photo (only if she/he is not depicted), date, location (if not depicted), or type of the picture (e.g., selfie). Special consideration was given to art images. In contrast to isness features like whether the photo *is* a selfie, *is* a black and white picture, or *is* a photography overlayed with a photo filter, photos usually depicts artworks instead of being an artwork. According to Shatford Lane (2002), hashtags like #art or #artwork were assigned to the category aboutness, when they are related to a photo depicting art. If the entire image is an artwork (e.g., computer graphics, photographic art), such tags describe the isness. The same applies for people, objects, and places which could be—according to context—content-related or isness-related. For example, #polishgirl is content-related if the girl is depicted in the picture, but it has to be categorized into Isness if no person is visible and one had to assume a Polish girl is the author of the posting or is the photographer of the image.

2.2.5. Performativeness

"Performative tags" (Peters & Stock, 2007), also called "time and task related tags" (Kipp, 2006; Kipp & Campbell, 2006) or "signaling tags" (Dennis, 2006), call for an action like #followmearound, #discover, or #like4like. Performativeness goes back to Austin (1962), as there are sentences which do not describe a proposition but a call for or a promise of an action. They were classified into the category Performativeness. Especially for Instagram, so-called hashtag contests like "participatory hashtag projects" (Oh et al., 2016) or features exist. One participates by simply tagging a picture with a certain hashtag. In doing so, the chance of a price or feature could exist in some cases. Feature means an account posts (and thus promotes) a picture or username of someone, like for example the Instagram account hypetoronto displayed in Fig. 3. Often, those tags are named similar to their related accounts like #hypetoronto for hypetoronto (Fig. 3), so that the call to action is not always exactly recognizable. However, such hashtags still belong to the performative ones and were categorized to this category, too. Some cases received an exemption clause in the codebook.

(a) (b)

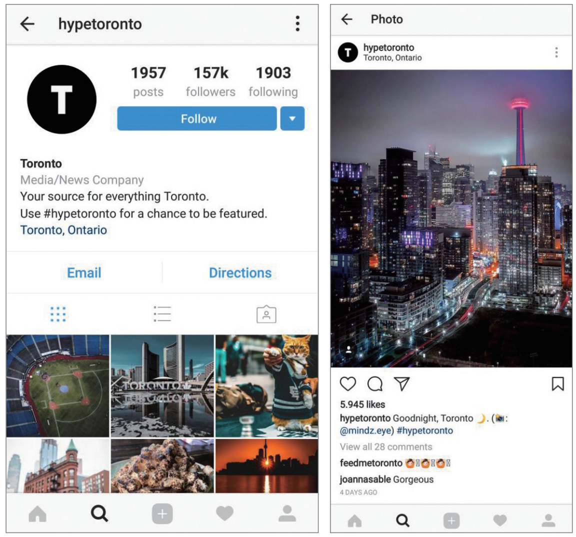

Fig. 3. Featured performative hashtag. (a) Instagram account *hypetoronto* calls for using their hashtag #hypetoronto, to get the chance to be featured in their feed. (b) *hypetoronto* features a photograph by @ mindz.eye who previously used the hashtag #hypetoronto.

Tags like #pictureoftheday or #petoftheday have a "contest character," because the tagged photos slightly compete with each other to be the best picture or the best depicted pet of the day. For that reason, they possibly could be categorized as performative tags. However, they cover up aboutness (e.g., "pet of the day," when displayed in the picture) or isness (e.g., "photo of the day," regarding aspects of the image and if no photo is depicted in the image). Since only one category could be selected, such general performative hashtags combined with "of the day" were not categorized into the category Performativeness. The categories Content-relatedness or Isness were chosen depending on the case.

2.2.6. Fakeness

The category Fakeness includes all stated hashtags that are not valid for the respective image or posting description in any way and are therefore deliberately incorrectly assigned, for instance, a picture tagged with #dog but depicting a single cat and nothing else. If the decision for a hashtag provides room for interpretation and it can not be clearly determined if the hashtag is true or false, it will not be categorized into Fakeness. The same applies to hashtags containing typos or the false singular/plural form. It was considered to categorize those latter hashtags into Fakeness or to create a category "Falseness." However, it is well known that folksonomy vocabulary is not formally proved (Peters, 2009) and such mistakes may happen. Besides, some tags are more common in their singular or plural form. It may be not correct to assign a picture representing a single person with #friends, but it is not a false statement in total.

2.2.7. "Insta"-Tags

A special focus lays on the category "Insta"-Tags. This category contains the following hashtags or hashtag components: "Insta," "gram," "Instagram," or an abbreviation of these expressions (like #IGers for Instagramers). Such components contain hashtags like for example #instagood, #instadaily, #instapic, #instaart, #instacat, #webstagram, or #webstagramers and would be classified into this hashtag category. Hashtags which include one or more stated components, but do not refer to Instagram, were excluded from this category.

Both content analyses for the stated picture and hashtag categories are predominantly based on a direct content analysis approach (Hsieh & Shannon, 2005) also known as deductive content analysis (Elo & Kyngäs, 2008). In those approaches, analysis and initial codes are based on previous knowledge of a theory or research findings (Hsieh & Shannon, 2005; Elo & Kyngäs, 2008). Additionally, new categories emerged from the observed data which are lean on conventional content analysis (Hsieh & Shannon, 2005), also known as inductive content analysis process (Elo & Kyngäs, 2008). For the hashtag category codebook, these newly created categories are Fakeness, "Insta"-Tags, and Sentences, where Content-relatedness (aboutness/ofness/iconology), Emotiveness, Isness, and Performativeness were found in the literature. For the picture category codebook, we followed mainly deductive analysis and referred to the work by Hu et al. (2014). The new categories Architecture, Art, and Landscape are widespread on Instagram and thus inductively considered as categories.

2.3. Codebooks

In total, two codebooks were developed for this study: the "Instagram picture category codebook" for the selection of the 1,000 Instagram picture postings, and the "Instagram hashtag category codebook" for the analysis of the assigned picture posting hashtags. They contain a short introduction about their purpose and all common coding rules which apply on every investigated Instagram posting, as well as their specific picture or hashtag codes. The creation and structure of both codebooks are based on MacQueen, McLeallan, Kay, and Millstein (2009).

The rules (R) in Table 2 apply to every investigated Instagram post and are taken from the Instagram picture category codebook. All pictures which do not predominantly show one category or suit the preference rules have been discarded. The rules (R) in Table 4 apply to every code and hashtag (and are taken from Instagram hashtag category codebook).

Based on MacQueen et al. (2009) every specific picture or hashtag code contains: a code name, a definition, and also a short definition for the coding category, information about when to apply or when not to use the code (as well as some preference rules), and coding examples.

It was necessary to continually update the books due to the pretest and coding process as well as to recode parts of data in the end, because some rules changed during the process.

2.4. Data Collection

All Instagram pictures were collected during November 2016 to January 2017. The final data comprises 50 Instagram pictures for a pretest and at least 1,000 Instagram pictures for the analysis. Pretest pictures were chosen manually. Category Architecture replaced category Gadget after the pretest and was thus not included into the pretest. The collection of the examined 1,000 Instagram pictures can be subdivided in an automatic and a manual part. The automatic part includes downloads of raw datasets via the Instagram web viewer Imgrum Application Program Interface (API).[1] The images could not be downloaded directly from the Instagram API, because this API was not freely accessible at this time. Based on an adequate top hashtag (Table 3) for each picture category, a dataset was downloaded.

Criteria for selecting the top hashtags were the thematic reference to the respective picture category in combination with the number of tagged media. Regarding the thematic aspects, all hashtags were chosen in order to be very

[1] http://www.imgrum.net/
[2] e.g., https://de.tagdef.com/

Table 4. General coding rules for categorizing Instagram hashtags

#	Rule
R1	Every hashtag can only be assigned to one category. Thereby, the category is chosen which (according to the coding rules and in combination with the depicted content on the picture) most closely applies to the hashtag.
R2	Only all author assigned hashtags in the picture description are considered for the coding. Hashtags in comments (also when the author of the picture commented on his/her posted picture) are excluded from the analysis.
R3	Analysis language of hashtags is English. Every hashtag will be excluded from analysis, if it is not formulated in English. Excluded from this rule are proper names (e.g. for cities, mountains, meals, names of regional holidays). They can be retained in their language. If the coder does not know them, he or she had to translate those proper names by himself or herself (only for understanding).
R4	The categorization of a hashtag primarily depends on the content of the associated posted Instagram picture. Additionally, further information can be obtained from the author's picture description. Similar as in R2, all other information in comments is excluded for the coding process, as well as the author's/commenter's profile information.
R5	Typos in hashtags are not considered.
R6	Special characters within hashtags or hashtags which consist only of one or more special characters are not taken into account for the analysis. The same applies for emojis.
R7	The coder has to figure out the meaning of hashtag abbreviations. Therefore hashtag meanings and definition websites[2] can be used. It is important that the coder annotates all used definition websites. In addition, all coders keep a record of the abbreviations and their written-out form. If the meaning is ambiguous or vague, the coder has to choose the most appropriate category.
R8	"Divided tags" (e.g., #private #lessons; #black #cat) are always considered and categorized separately.
R9	Same hashtags, but double or multiple assigned tags are only considered as one (e. g. #food, #food, #food counts as one hashtag).

general and close to their picture category, such that they cover their category optimally. For instance, the category Fashion comprises not only clothes, but rather accessories and makeup too. Therefore, it would not have been useful to choose a hashtag like #outfit, or #clothes. Relating to the count, the hashtag had to be one of the most common hashtags. Table 3 shows that every chosen hashtag had the highest number of tagged media in contrast to the alternative hashtags.

Such an approach enables a random picture selection which usually matches the selected hashtag aboutness and is adopted from Marcus (2016). However, it should be noted that not every picture subject conformed to the tagged hashtag, since the author could make false statements (e.g., using #selfie despite the picture not depicting a selfie). Besides, some appropriate subjects were not in accordance with the previously enumerated rules of the category codebook: for example, a pet picture subject tagged with #friends. The pet may be a friend in this case, but as the codebook denotes, only human friends are valid picture subjects for this category. Or, the picture shows elements of two or more categories in equal shares. In doing so, it fulfills not only the category the hashtag pretends, but rather more categories. For that reason, the automatically generated datasets covered a multiple of the respective required 100 category pictures, so that they could be cleaned up. It should be noted that some data had to be downloaded in several stages, since after the manual check only a few crawled pictures fulfilled the conditions for this study (language, category, etc.), so that the first data set did not always suffice. The datasets were collected in parallel to the coding process.

Every final and adjusted downloaded data set contains the following information: user name of the posting author; URL to the posting on Imgrum; the full image description of the posting; all used hashtags in the image description; the total hashtag number of one image description; the picture category name (e.g., Activity, Architecture, Art, etc.); a unique picture ID; and a separate file, containing each picture file named with its ID.

As stated in coding rule R2 (Table 4), further posting comments (whether from the author or the users) were not considered for this study.

2.5. Coding Process

The coding process involves the assignment of the picture description hashtags to the introduced hashtag categories and with regard to the code rules. In November 2016, the coding pretest took place, followed by the categorization of the 1,000 image dataset during December 2016 to February 2017. According to a 4-eyes principle, 2 coders were involved in the process. First, both of them had to code individually each data set in accordance with the hashtag codebook. What followed was a check as to whether the assigned codes matched. If they did not match, both coders had to discuss and to agree for one category (based on the codebook rules).

Coding is an iterative process. After the pretest and during the coding process some (special) cases appeared for the first time or have been defined as too vague, making it necessary to adjust or supplement some of the rules. Even if there was a pretest, coders had to adapt themselves to the coding rules at the beginning of the main coding process. This might cause additional reasons for recoding, too. For instance, hashtags containing "of the day" like in #pictureoftheday or #photooftheday were widespread in Instagram. "Picture of the day" and similar hashtags generally refer to Isness. According to that and thus the former codebook rules, all kinds of "oftheday"-hashtags had to be coded into Isness at the beginning. During the main coding phase, this rule was revised in consequence of more and more "of the day"-hashtags addressing an element like a living being or object in the picture (e.g., "catoftheday" and a cat was depicted, "foodoftheday" and food was depicted). In this case and also in the other cases, the pretest and the pre-developed hashtag codebook were not sufficient to handle those cases correctly and the codebook had to be continually updated. In the end, a follow-up check and further recoding of coding data parts (regarding changed rules as well as the category content-relatedness) took place. As stated, the category content-relatedness comprises ofness, aboutness, and iconology in the end. At first, the coders had only distinguished between ofness and aboutness hashtags. After the entire coding process, both categories were merged into the category content-relatedness.

3. RESULTS

The following paragraph presents the outcome of this study following the three research questions.

RQ1 : Are there any differences in relative frequencies of hashtags in the picture categories?

Overall, 14,649 hashtags for a total 1,000 Instagram pictures were coded into their respective hashtag categories (Table 5). Thereby, the average number of hashtags per picture in the respective picture category (Fig. 4) varies from nearly 11 to about 19 hashtags with an average of 15 hashtags. Especially, the person-related categories Selfie (average 10.9 tags per picture) and Friends (average 11.7 tags per picture) received the lowest average values. Pet (average 18.6 tags), Fashion (average 17.6 tags), and Landscape (average 16.8 tags) are the picture categories with the highest average hashtag count for their pictures. The standard deviation is similar for nearly every picture category (between 8 and 9).

RQ2 : Given a picture category, what is the distribution of hashtag categories; and given a hashtag category, what is the distribution of picture categories?

With 60.20%, the majority of all hashtags were classified into the category Content-relatedness, followed by the hashtag category Isness with almost 14.87%. "Insta"-Tags (7.32%) were third most; tags of the category Performativeness (7.20%) were the fourth most assigned. A minority of hashtags was classified into the categories Fakeness (5.03%), Emotiveness (4.38%), and Sentences (0.99%).

Pictures of all 10 categories are also predominantly tagged with content-related hashtags, but high values for "Insta"-Tags in the category Pet (20.24%) and Fakeness for the category Captioned Photo (11.80%) are striking. In contradiction to the total values of all assigned hashtags,

Table 5. Relative frequency of hashtag categories by picture categories (N=1,000 posts; 100 posts per picture category)

	Content-relatedness	Emotiveness	Fakeness	"Insta"-Tags	Isness	Performa-tiveness	Sentences	%
Activity	74.29	4.35	3.75	1.72	9.90	5.25	0.75	100.00
Architecture	63.20	2.54	2.54	6.67	15.41	9.35	0.28	100.00
Art	68.54	1.05	5.62	4.19	14.68	5.69	0.22	100.00
Captioned Photo	61.67	5.16	11.80	3.46	7.33	8.75	1.83	100.00
Fashion	59.51	3.52	7.38	5.91	16.70	6.02	0.97	100.00
Food	51.34	1.81	6.71	6.31	25.84	7.11	0.87	100.00
Friends	57.75	8.74	3.94	7.88	14.40	6.17	1.11	100.00
Landscape	61.68	3.98	1.66	5.47	16.16	10.75	0.30	100.00
Pet	53.93	8.02	3.82	20.24	7.59	4.31	2.10	100.00
Selfie	51.05	4.57	2.38	7.96	23.70	9.06	1.28	100.00
Total	60.20	4.38	5.03	7.32	14.87	7.20	0.99	100.00

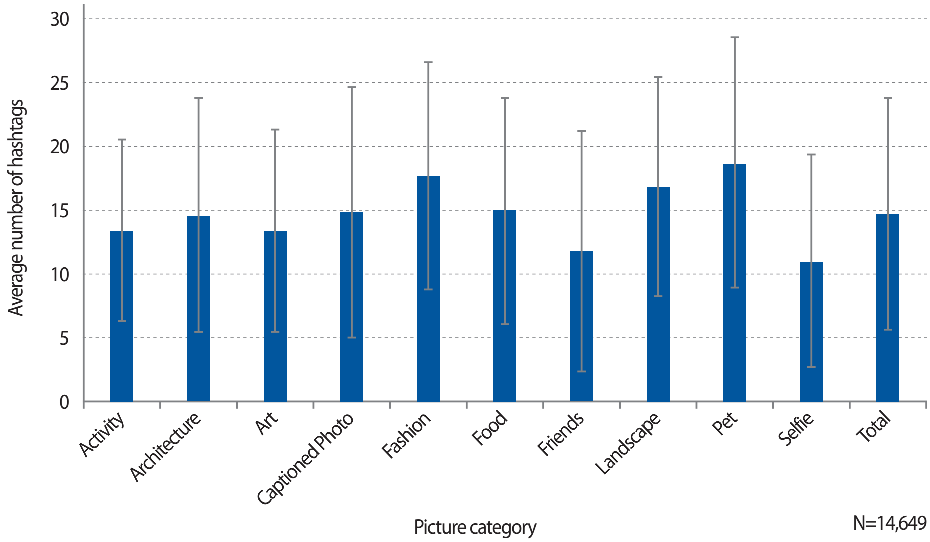

Fig. 4. Average number of hashtags per picture and standard deviation within its respective picture category.

they are the second most assigned hashtag categories for their stated picture category. For all other picture categories, the second most assigned category is Isness. Even if the total amount of "Insta"-Tags is the third most assigned, this hashtag category is not the third most for any of the remaining 9 picture categories. For them, "Insta"-Tags are mostly fourth or fifth most, and for the categories Activity and Captioned Photos they were even sixth most assigned. The overall high distribution is therefore resulting due to the frequent occurrence in the category Pet. Considering the last ranks, for each picture category Sentences hashtags were tagged at least most, which is in accordance with the overall hashtag category distribution. It is noticeable, however, that the category Emotiveness is only three times in the sixth rank (for categories Art, Fashion, and Food).

The distribution of the hashtag category Content-relatedness varies between half and three quarters of all hashtags per each image category. With approximately 51%, Food and Selfie related pictures are tagged fewest of all with Content-related hashtags. The highest percentages record the categories Activity (74.29%) and Art (68.54%). Even though Food and Selfie have the lowest rate of content-related tags, the occurrence of Isness tags in these two categories is the highest. The majority of the other picture categories have around about 15%, apart from the categories Activity (9.90%), Pet (7.59%), and Captioned Photo (7.33%), with under 10%. The distribution for "Insta"-Tags amounts for all picture categories less than 8%, except for the category Pet (20.24%). With 1.72%, Activity pictures have an especially low frequency of "Insta"-Tags. Regarding Performativeness, the categories Landscape (10.75%), Architecture (9.35%), and Selfie (9.06%) have the most hashtags of this category. Generally, Fakeness tags are not strongly represented. Similar to "Insta"-Tags, there is only one category exceeding the 8% for Fakeness, namely the category Captioned Photos (11.80%). The fewest Fakeness tags were received by Landscape pictures (1.66%). Pictures about Friends (8.74%) or Pets (8.02%) were tagged most emotionally, whereas Art (1.05%) and Food (1.81%) images contained nearly no emotional hashtags. Within the hashtag category Sentences, Pet pictures (2.10%) record the highest percentage of those hashtags. It is notable that 35 out of 39 Pet Sentences hashtags are about emotional facts like for example #ilovemydog or #ihatemondays. So the category Pet received not only the most Sentences hashtags, but also the most emotional.

RQ3 : Is there any association between image categories and hashtag categories?

A chi-square test of independence was conducted between hashtag categories and picture categories with the following hypotheses:

H_0 : There is no association between hashtag categories and hashtag pictures.

H_A : An association exists between hashtag categories and hashtag pictures.

All expected cell frequencies were greater than five. There was a statistically significant association between hashtag categories and picture categories, $\chi 2(54)=1557.860$, $P<.0005$. The association was small (Cohen, 1988), Cramer's V=.133. Therefore, we can reject the null hypothesis and accept the alternative hypothesis.

4. DISCUSSION

This research study connects aspects of the research fields social media, knowledge representation, and image indexing in order to investigate the outcomes of tagging behavior of Instagram users. In respect of the research model, 1,000 Instagram user postings, consisting of a picture and a textual author description with hashtags, were analyzed in terms of their hashtag frequencies, hashtag categories, and picture categories.

On average, the Instagram users applied about 15 hashtags per picture. A closer consideration shows that the categories Pet, Fashion, and Landscape received between 1 and 4 more hashtags on average. Instagram users assigned the fewest tags to the person-related categories Selfie and Friends. The average hashtag number goes in line with previous research. An average 12.41 hashtags per picture (based on 400 pictures and 4,966 hashtags) was identified by Veszelszki (2016) in her study.

The representations of ofness and aboutness are important aspects for Instagram users. They tagged their pictures predominantly with Content-related hashtags (about 60%). Likewise as in our results, such picture describing hashtags (called thematizing context-marker hashtags) were the most frequent (83%) in the sample of Veszelszki's (2016) study. Beside ofness and aboutness, indexing refers in the classical sense of knowledge representation also to aspects regarding to the Isness of a document; however, in professional information services, there are specific bibliographic fields including Isness aspects. This is also reflected in this study since the relative frequency of Isness hashtags was the second most. It is noticeable that only a few emotional hashtags were assigned, although social media services such as Instagram are characterized by social and emotional interactions. The usage of Sentences hashtags is generally weak. Especially, "Insta"-Tags are often assigned to pictures of the category Pet. The same applies for Fakeness tags and pictures of the category Captioned Photo. A small statistical association between hashtag categories and picture categories was found.

Content analysis on Instagram was already realized in several studies as mentioned in the state of research. The key strength of this study is its diversity. It is the first study which evaluates Instagram not only with respect to a single topic or hashtag, but rather in a wider field of both picture categories and hashtags. In doing so, it reveals a general insight into the tagging behavior of Instagram users.

Data collection and sampling of social media data can be challenging. Therefore, the generalizability of these results is subject to the following limitations. Since Instagram offered only limited access to the data at the time of data collection, the data set is based on a non-probability sampling instead of a sample which represents the whole population. However, this problem is well-known in social media research and is not limited to this study. Furthermore, the study only investigated pictures and the respecting account owners' picture descriptions. Besides pictures, videos are part of Instagram content as well.

This study shows how Instagram users tag their photos. By the results of this work, further studies regarding users' intentions are enabled. At this point, we can draw some conclusions from the results. The proportion of Content-related tags is possibly so high since users finally want to find their photos by content. If someone wants to see something about, for example, travel places, he or she would use appropriate hashtags of this topic as search arguments. In order to make pictures detectable, the description of their content by hashtags is necessary. Which role do the rest of the hashtag categories play? An explanation for the generally weak hashtag use of Sentences could be that users consider those hashtags as too specific. Perhaps they think that nobody else searches for those tags and prefer more frequently used tags in order to get more attention for their postings.

The phenomenon of "Insta"-Tags is platform-specific and—to the best of the authors' knowledge—not on any other platform as a similar phenomenon in use. Why and how was it established by users? Especially, "Insta"-Tags in the category Pet are frequent. Does this have particular reasons? Considering all 100 Pet postings, it is striking that for similar picture motives chosen hashtags were consimilar. Websites and apps listing top Instagram hashtag lists (in general and for specific topics in order to receive more interactions) are well known.[3] Those lists also include "Insta"-Tags. Maybe the account owners of the observed postings for the category Pet especially used those pre-built hashtag sets. This could explain the huge amount of similar hashtags as well as the huge amount of "Insta"-Tags in the category Pet. However, the use of "Insta" is not limited to pet related matters, so it is questionable what the exact reasons are. Why are Emotiveness hashtags so less frequently assigned, although social media is full of emotional content? For the category Art, this could be explained by the fact that the majority of analyzed artwork pictures are original works, created by the posting authors. Foreign art could evoke more emotions. Picture postings about Friends and Pets could have received the most emotional hashtags because they represent a relationship between entities causing emotions in us. However, the next step should be to prove these hypotheses.

To gather further insights about tagging behavior, the parameters of this study could be broadened. Besides pictures, videos could be analyzed, too. Are there any differences in tagging behaviors of pictures and videos on Instagram? Could it be interesting to distinguish between three main aspects; tags which only apply to the picture or video, tags which only apply to the descriptive content, and tags which apply to both? Are there any gender-specific or cultural differences concerning tagging behavior on Instagram?

Furthermore, it would be interesting to investigate how tagging behavior affects the success of an Instagram account. Do "successful" users (i.e., users with many followers) tag their pictures differently than less successful users? To what extent does tagging behavior influence success on Instagram in general? Therefore, it is necessary to define which factors are responsible for success on Instagram. Possible factors could be the count of followers, likes per postings, comments, favored pictures, editing of pictures and videos, or social interaction with other users, stating only a few.

The findings of this paper show how Instagram users tag their pictures; however, the analysis of the reasons for their tagging behavior was not part of this study. In order to understand the tagging behavior more deeply, a holistic theory that also accounts for tagging motivations is required.

ACKNOWLEDGMENTS

I would like to thank Anika Mehlem for creating and providing a program for the data collection as well as Franziska Zimmer for her work as coder during the content analysis.

REFERENCES

Allem, J.-P., Escobedo, P., Chu, K.H., Cruz, T. B., & Unger, J. B. (2017). Images of little cigars and cigarillos on Instagram identified by the hashtag #swisher: thematic analysis. *Journal of Medical Internet Research*, 19(7), e255.

Andalibi, N., Ozturk, P., & Forte, A. (2017). Sensitive self-disclosures, responses, and social support on Instagram: the case of #depression. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (pp. 1485-1500). New York: ACM.

Austin, J. L. (1962). *How to do things with words*. Oxford: Clarendon.

Beaudoin, J. (2007). Folksonomies: Flickr image tagging: patterns made visible. *Bulletin of the American Society for Information Science and Technology*, 34(1), 7-11.

Buarki, H., & Alkhateeb, B. (2018). Use of hashtags to retrieve information on the web. *The Electronic Library*, 36(2), 286-304.

Carrotte, E. R., Prichard, I., & Lim, M. S. C. (2017). "Fitspiration" on social media: A content analysis of gendered images. *Journal of Medical Internet Research*, 19(3), e95.

Coelho, R. L. F., de Oliveira, D. S., & de Almeida, M. I. S. (2016). Does social media matter for post typology? Impact of post content on Facebook and Instagram metrics. *Online Information Review*, 40(4), 458-471.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York: Psychology Press.

Daer, A. R., Hoffman, R. F., & Goodman, S. (2014). Rhetorical functions of hashtag forms across social media applications. *Communication Design Quarterly Review*, 3(1), 12-16.

Dennis, B. M. (2006). Foragr: Collaborative tagged photographs and social information visualisation. In *Proceedings of the Collaborative Web Tagging Workshop at WWW 2006*, Edinburgh, UK.

Denton, E., Weston, J., Paluri, M., Bourdev, L., & Fergus, R. (2015). User conditional hashtag prediction for images. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1731-1740). New York: ACM.

Dorsch, I., Zimmer, F., & Stock, W. G. (2017). Image indexing through hashtags in Instagram. *Proceedings of the Association for Information Science and Technology*, 54(1), 658-659.

Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of Advanced Nursing*, 62(1), 107-115.

Enser, P. G. B. (2008). Visual image retrieval. *Annual Review of Information Science and Technology*, 42(1), 1-42.

Giannoulakis, S., & Tsapatsoulis, N. (2015). Instagram hashtags as image annotation metadata. In R. Chbeir,

[3] e.g., https://www.tagsforlikes.com/

Y. Manolopoulos, I. Maglogiannis, & R. Ralhajj (Eds.), *Artificial Intelligence Applications and Innovations. IFIP Advances in Information and Communication Technology* (Vol. 458, pp. 206-220). Cham: Springer.

Gibbs, M., Meese, J., Arnold, M., Nansen, B., & Carter, M. (2015). #Funeral and Instagram: death, social media, and platform vernacular. *Information, Communication & Society*, 18(3), 255-268.

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

Guidry, J. D., Messner, M., Jin, Y., & Medina-Messner, V. (2015). From #mcdonaldsfail to #dominossucks: an analysis of Instagram images about the 10 largest fast food companies. *Corporate Communications: An International Journal*, 20(3), 344-359.

Halavais, A. (2014). Structure of Twitter: Social and technical. In K. Weller, A. Bruns, J. Burgess, N. Mahrt, & C. Puschmann (Eds.), *Twitter and society* (pp. 29-41). New York: Peter Lang.

Hollenstein, L., & Purves R. (2010). Exploring place through user-generated content: using Flickr tags to describe city cores. *Journal of Spatial Information Science*, 1(1), 21-48.

Holmberg, C., Chaplin, J. E., Hillman, T., & Berg, C. (2016). Adolescents' presentation of food in social media: an explorative study. *Appetite*, 99, 121-129.

Hsieh, H.-F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, 15(9), 1277-1288.

Hu, Y., Manikonda, L., & Kambhampati, S. (2014). What we Instagram: A first analysis of Instagram photo content and user types. In *Proceedings of the 8th International Conference on Weblogs and Social Media, ICWSM 2014* (pp. 595-598). Palo Alto, CA: AAAI Press.

Ingwersen, P. (2002). Cognitive perspectives of document representation. In H. Bruce, R. Fidel, P. Ingwersen, & P. Vakkari (Eds.), *Proceedings of the 4th International Conference on Conceptions of Library and Information Science* (pp. 285-300). Greenwood Village, CO: Libraries Unlimited.

Instagram. (2010). *Instagram launches*. Retrieved April 14, 2017 from https://instagram-press.com/blog/2010/10/06/instagram-launches-2/.

Instagram. (2018). *Instagram statistics*. Retrieved Mar 5, 2018 from https://instagram-press.com/our-story.

Jörgensen, C. (2002). Image access, the semantic gap, and social tagging as a paradigm shift. In J. Lussky (Ed), *Proceedings of the 18th Workshop of the American Society for Information Science and Technology Special Interest Group in Classification Research*, Milwaukee, WI.

Jörgensen, C. (2003). *Image retrieval: Theory and research*. Lanham, MD: Scarecrow Press.

Kipp, M. E. I. (2006). @toread and cool: Tagging for time, task and emotion. In *Proceedings of the 17th Annual ASIS&T SIG/CR Classification Research Workshop, Austin, TX*.

Kipp, M. E. I., & Campbell, D. G. (2006). Patterns and inconsistencies in collaborative tagging systems: an examination of tagging practices. *Proceedings of American Society for Information Science and Technology*, 43(1), 1-18.

Knautz, K. (2012). Emotion felt and depicted: consequences for multimedia retrieval. In D. R. Neal (Ed.), *Indexing and retrieval of non-text information* (pp. 343-375). Berlin: De Gruyter Saur.

Kremerskothen, K. (2012). *An amazing 8 years*. Retrieved April 16, 2017 from http://blog.flickr.net/en/2012/02/10/an-amazing-8-years/.

Krippendorff, K. (2004). *Content analysis: an introduction to its methodology* (2nd ed). Thousand Oaks, CA: Sage Publications.

Lancaster, F. W. (2003). *Indexing and abstracting in theory and practice*. Champaign, IL: University of Illinois, Graduate School of Library and Information Science.

MacQueen, K. M., McLellan, E., Kay, K., & Milstein, B. (2009). Codebook development for team-based qualitative analysis. In K. Krippendorf, & M. A. Bock (Eds.), *The content analysis reader* (pp. 211-219). Los Angeles: Sage Publications.

Marcus, S.-R. (2016). Thinspiration vs. thicksperation: comparing pro-anorexic and fat acceptance image posts on a photo-sharing site. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, 10(2).

Maron, M. E. (1977). On indexing, retrieval and the meaning of about. *Journal of the American Society for Information Science*, 28(1), 38-43.

Messina, C. (2007). *how do you feel about using # (pound) for groups. As in #barcamp [msg]? @chrismessina (Twitter)*. Retrieved May 2, 2017 from https://twitter.com/chrismessina/status/223115412.

Mohammad, S. M., & Kiritchenko, S. (2015). Using hashtags to capture fine emotion categories from Tweets. *Computational Intelligence*, 31(2), 301-326.

Nashmi, E. A. (2018). From selfies to media events: from selfies to media events: how Instagram users interrupted their routines after the Charlie Hebdo shootings. *Digital Journalism*, 6(1), 98-117.

Nov, O., Naaman, M., & Ye, C. (2008). What drives content tagging: The case of photos on Flickr. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1097-1100). New York: ACM.

Oh, C., Lee, T., Kim, Y., Park, S., & Suh, B. (2016). Understanding participatory hashtag practices on Instagram: a case study of weekend hashtag project. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (pp. 1280-1287). New York: ACM.

Ortony, A., & Turner, T. J. (1990). What's basic about basic emotions? *Psychological Review*, 97(3), 315-331.

Panofsky, E. (1955). *Meaning in the visual arts*. Garden City, NY: Doubleday Anchor Books.

Pegoraro, A., Comeau, G. S., & Frederick, E. L. (2018). #SheBelieves: the use of Instagram to frame the US women's soccer team during #FIFAWWC. *Sport in Society*, 21(7), 1063-1077.

Peters, I. (2009). *Folksonomies: Indexing and retrieval in Web 2.0.* Berlin: De Gruyter Saur.

Peters, I., & Stock, W. G. (2007). Folksonomy and information retrieval. *Proceedings of the American for Information Science and Technology*, 44(1), 1-28.

Pila, E., Mond, J. M., Griffiths, S., Mitchison, D., & Murray, S. B. (2017). A thematic content analysis of #cheatmeal images on social media: Characterizing an emerging dietary trend. *International Journal of Eating Disorders*, 50(6), 698-706.

Pluzhenskaia, M. (2006). Folksonomies or fauxsonomies: How social is social bookmarking? In *Proceedings of the 17th ASIS&T SIG/CR Classification Research Workshop* (pp. 23-24), Austin, TX.

Rasmussen, E. M. (1997). Indexing images. *Annual Review of Information Science and Technology* (ARIST), 32, 169-196.

Rondeau, S. (2014). The life and times of aboutness: A review of the library and information science literature. *Evidence Based Library and Information Practice*, 9(1), 14-35.

Rorissa, A. (2010). A comparative study of Flickr tags and index terms in a general image collection. *Journal of the American Society for Information Science and Technology*, 61(11), 2230-2242.

Schmidt, S., & Stock, W. G. (2009). Collective indexing of emotions in images: a study in emotional information retrieval. *Journal of the Association for Information Science and Technology*, 60(5), 863-876.

Shatford, S. (1986). Analyzing the subject of a picture: a theoretical approach. *Cataloging & Classification Quarterly*, 6(3), 39-62.

Shatford Layne, S. (1994). Some issues in the indexing of images. *Journal of the American Society for Information Science*, 45(8), 583-588.

Shatford Layne, S. (2002). Subject access to art images. In M. Baca (Ed.), *Introduction to art image access: Issues, tools, standarts, strategies* (pp. 1-20). Los Angeles, CA: Getty Research Institute.

Sheldon, P., & Bryant, K. (2016). Instagram: Motives for its use and relationship to narcissism and contextual age. *Computers in Human Behavior*, 58, 89-97.

Siebenlist, T. (2013). Emotionale suche. In D. Lewandowski (Ed.), *Handbuch Internet-Suchmaschinen* (pp. 299-327). Heidelberg: Akademische Verlagsgesellschaft.

Smith, G. (2004). Folksonomy: *Social classification*. Retrieved May 9, 2018 from http://atomiq.org/%0Aarchives/2004/08/folksonomy_social_classification.html.

Souza, F., de Las Casas, D., Flores, V., Youn, S., Cha, M., Quercia, D., & Almeida, V. (2015). Dawn of the selfie era: the whos, wheres, and hows of selfies on Instagram. In *Proceedings of the 2015 ACM on Conference on Online Social Networks* (pp. 221-231). New York: ACM.

Stock, W. G., & Stock, M. (2013). *Handbook of information science*. Berlin: De Gruyter Saur.

Stvilia, B., & Jörgensen, C. (2010). Member activities and quality of tags in a collection of historical photographs in Flickr. *Journal of the American Society for Information Science and Technology*, 61(12), 2477-2489.

Turner, J. M. (1995). Comparing user-assigned terms with indexer-assigned terms for storage and retrieval of moving images: Research results. In *Proceedings of the 58th ASIS Annual Meeting, Vol 3.* (pp. 9-12). Medford, NJ: Information Today.

Vander Wal, T. (2007). *Folksonomy coinage and definition*. Retrieved Apr 16, 2017 from http://www.vanderwal.net/folksonomy.html.

Veszelszki, Á. (2016). #time, #truth, #tradition: an image-text relationship on Instagram: Photo and hashtag. In A. Benedek, & A. Veszelszki (Eds.), *In the beginning was the image: the omnipresence of pictures: time, truth, tradition* (pp. 93-113). New York: Peter Lang.

Yanbe, Y., Jatowt, A., Nakamura, S., & Tanaka, K. (2007). Can social bookmarking enhance search in the web? In *Proceedings of the 7th ACM/IEEE-CS joint Conference on Digital Libraries* (pp. 107-116). New York: ACM.

Ye, Z., Hashim, N. H., Baghirov, F., & Murphy J. (2018). Gender differences in Instagram hashtag use. *Journal of Hospitality Marketing & Management*, 27(4), 386-404.