

Phans, Stans and Cishets: Self-Presentation Effects on Content Propagation in Tumblr

Michael Miller Yoder
Carnegie Mellon University
Pittsburgh, Pennsylvania
yoder@cs.cmu.edu

Qinlan Shen
Carnegie Mellon University
qinlans@cs.cmu.edu

Yansen Wang
Carnegie Mellon University
yansenwa@andrew.cmu.edu

Alex Coda
Carnegie Mellon University
acoda@andrew.cmu.edu

Yunseok Jang
University of Michigan
Ann Arbor, Michigan
yunseokj@umich.edu

Yale Song
Microsoft Research
Redmond, Washington
yalesong@microsoft.com

Kapil Thadani
Yahoo Research
New York, New York
thadani@verizonmedia.com

Carolyn P. Rosé
Carnegie Mellon University
cprose@cs.cmu.edu

ABSTRACT

Research on content propagation in social media has largely focused on features from the content of posts and the network structure of users. However, social media platforms are also spaces where users present their identities in particular ways. How do the ways users present themselves affect how content they produce is propagated? In this paper, we address this question with an empirical study of interaction and self-presentation data from Tumblr. We use a pair-wise learning-to-rank framework to predict whether a given user will reblog (share) another user's post from features comparing self-presented textual and visual identity information. We find evidence that alignment in identity presentation is associated with content propagation, as these features increase performance over a baseline of content features. Interpreting learned feature weights on self-presented text identity labels, we find that users who present labels that match or indicate shared interests and values are generally more likely to propagate each other's content.

CCS CONCEPTS

• **Human-centered computing** → **Collaborative and social computing**; • **Computing methodologies** → *Natural language processing*.

KEYWORDS

social media, identity, Tumblr, content propagation, identity labels

ACM Reference Format:

Michael Miller Yoder, Qinlan Shen, Yansen Wang, Alex Coda, Yunseok Jang, Yale Song, Kapil Thadani, and Carolyn P. Rosé. 2020. Phans, Stans and Cishets: Self-Presentation Effects on Content Propagation in Tumblr. In *12th ACM Conference on Web Science (WebSci '20)*, July 6–10, 2020, Southampton, United Kingdom. ACM, New York, NY, USA, 10 pages. <https://doi.org/10.1145/3394231.3397893>

1 INTRODUCTION

Social media is a space where users not only share content, but also construct identities and position themselves in relation to others. However, it is difficult to directly measure how presentations of identity come across to others at a large scale. In this paper, we develop a framework to identify such self-presentation effects on a primary form of interaction on social networks: content (information) propagation.

Prior work quantifying patterns of social media content propagation has not addressed a relationship to self-presentation. Most existing studies rely primarily on content and network features [29, 42, 43]. However, content propagates through a social network from individual decisions to share others' posts. At this local level, content propagation is an interaction between users. Users with similar attributes are known to have stronger network connections (the property of homophily). In this way, self-presented identity attributes may affect content sharing as a form of network connection. This paper investigates this connection for two reasons: in order to broaden understanding of factors associated with content propagation, and to provide an experimental paradigm in which reactions to self-presentation practices can be investigated.

We use data from Tumblr, a blogging and social media platform. Sharing content is one of the primary modes of interaction on Tumblr; more than 90% of posts are "reblogs" of other posts [39]. Identity construction is also an important part of the participatory culture on Tumblr. Users on Tumblr each have a personal blog, an individualized artifact that reflects a user's identity [22]. Tumblr's multimedia content, personalized blog layouts, and affordances for users to maintain multiple blogs without being tied to a real name

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

WebSci '20, July 6–10, 2020, Southampton, United Kingdom

© 2020 Association for Computing Machinery.

ACM ISBN 978-1-4503-7989-2/20/07...\$15.00

<https://doi.org/10.1145/3394231.3397893>

have created a unique environment for identity expression without many of the social pressures found on other social media sites [13]. Talk about identity issues such as gender, sexuality, and ethnicity—as well as their intersection with media, culture, and fandom—is common on Tumblr [15]. This makes the identity positioning of users who create content especially relevant on the platform. For these reasons, we expect users’ self-presentation of identity to play a role in how content is propagated on Tumblr. We investigate this role through two research questions:

RQ 1: Is there evidence of an association between identity alignment and content propagation?

To explore the possibility of this effect, we construct a classification task predicting whether a user will reblog another user’s posts given how both users express their identity in profile images and text blog descriptions. We find that incorporating identity-related features improves classification performance when used in addition to post content features, and that profile images and text blog descriptions provide complementary effects. This shows evidence that identity presentation is associated with content propagation.

However, this result does not specify the nature of this effect: features signifying similarities, differences, or some other interaction between text or image features could be especially predictive. We therefore explore a further research question:

RQ 2: What is the nature of the association of identity alignment with content propagation?

We investigate this question with logistic regression models trained on identity comparison features between text blog descriptions. Tumblr users often use these free-text blog description fields to provide identity labels such as ‘girl’, ‘canadian’, or ‘intj’ [30]. We find that providing identity labels that indicate shared values or experiences generally increases the likelihood of users reblogging.

2 THEORETICAL MOTIVATION AND PREVIOUS WORK

In this section, we introduce and motivate the identity concepts investigated in our work, as well as situate our contribution among work in content propagation and Tumblr.

2.1 Identity and Interaction

We conceive of online self-presentation as the construction of an artifact that reflects identity in the particular space of Tumblr [22]. This artifact, a Tumblr blog, contains both linguistic and visual elements of self-presentation: written blog descriptions, profile images, and multimedia posts. Much of this self-presentation consists of symbols that point to culturally specific understandings of types of people. For example, being a female fan can be expressed in a blog description explicitly with labels (*i’m a super fangirl*) or more implicitly (♥♥OBSESSED♥♥ *with exo*) [16, 24].

The framework of “performed” identity is often used to analyze self-presentation [7, 9, 17], including in social media analysis. For example, Bullingham and Vasconcelos [8] applied Goffman’s dramaturgical metaphor to personas on Second Life and found that users often present their offline selves in avatar form with selective edits. In this work, we examine only the “performed” self presented in Tumblr blogs. We intentionally do not infer identity information

about users. Self-presented information is most relevant to the users themselves, and there is evidence that Tumblr users have built a unique culture around self-presentation. Working with Tumblr blog descriptions, Oakley [30] found that LGBTQIA users have developed labeling practices that challenge dominant binary conceptions of gender and sexuality.

What separates this paper from many other analyses of self-presentation online is a focus on the *effects* of self-presentation on interaction, specifically on how content is propagated. This contrasts with the common task of predicting user attributes from behavioral data [11]. There have been a few other computational studies that predict the effects of online self-presentation. Wang and Jurgens [37] examined how the presentation of gender on Reddit, StackExchange, and Wikipedia affects reactions of support. Bareket-Bojmel et al. [5] examined how self-enhancing and self-derogating posting strategies affected responses to users with different goals on Facebook. Vedres and Vasarhelyi [35] find that behaving in ways that are typical of male or female users on GitHub has more of a relationship with success than the “actual” gender of users. In this paper, we use content propagation as a measure of how particular forms of self-presentation come across in the context of Tumblr.

The well-known property of homophily suggests that users who share attributes are more likely to be connected within social networks. Gong et al. [18], for example, find that inferring user attributes improves link prediction performance in Google+. However, identifying what relationships between identity labels are meaningful in context is not trivial. Bucholtz and Hall [7], for example, identify authenticity and legitimacy as axes other than sameness/difference on which speakers position themselves. Piergallini et al. [31] discuss challenges with modeling the complex alliance systems in online street gang discourse. Our work builds a framework for identifying combinations of identity labels that hold social currency in their association with content propagation.

2.2 Content Propagation

Content propagation and virality on social media are often predicted from features of individual post content or patterns in how that content spreads through a network [29, 36, 43]. Naveed et al. [29], for example, predicted whether a tweet was going to be retweeted based on text features from the tweet. Zhang et al. [42] studied how the network feature of reciprocity is associated with reblogging on Sina Weibo. On Tumblr, where content can take diverse forms (text, photos, videos, audio, etc.), work more often focuses on network-based features. To predict whether a Tumblr post will become viral, Xie et al. [38] centered their analysis on “early adopters” (the first users to share a post) in combination with content-specific features. Others have examined network features in relation to reblog cascades. Chang et al. [10] examined Tumblr reblog cascades based on structure of the follower network, while Alrajebah et al. [1] examine both structural and temporal aspects of reblog cascades on a set of popular posts on Tumblr.

Our work extends this research by additionally viewing content propagation as a social interaction in which the identities of the users involved may play a role.

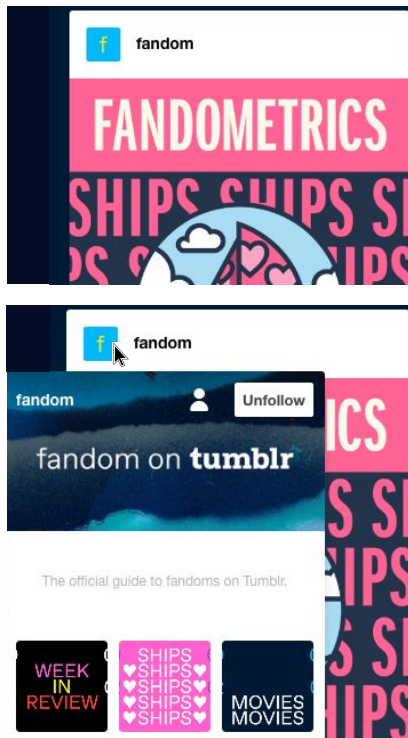


Figure 1: Example of information displayed about a Tumblr blog when mousing over the profile image on a post. This includes the text blog description.

2.3 Previous Research on Tumblr

Prior research on Tumblr can roughly be divided into social-scientific approaches, often focusing on identity expression, and computational research including network analysis, recommendation systems, and search tools for identifying relevant or problematic content [3, 10, 27, 39, 40]. Our work attempts to bring themes of identity that are well-studied in qualitative Tumblr work into a computational analysis of the effects of self-presentation. Specifically, we adopt a focus on the discursive *construction* of identity from qualitative identity research [7].

In their survey of research on Tumblr, Attu and Terras [2] find that sexuality and other identity issues are commonly studied. Many of these studies use qualitative methods to examine how identity is expressed through Tumblr content. Fink and Miller [15], for example, use auto-ethnographic dialogue to relate how trans and queer Tumblr users created an artistic space that challenged dominant straight, cisgender norms. In their interview- and content-based research on Tumblr users posting NSFW selfies, Tiidenberg [34] similarly finds that Tumblr users create a space for body and sex positivity outside what is deemed “sexy” by mainstream society. Haimson and Hayes [20] focus on language used in gender transition blogs before and after the divisive 2016 US presidential election. With the goal of informing sex educators and clinicians, Zeglin and Mitchell [41] choose Tumblr as a lens to see how public understandings of sexuality contrast with theoretical models of sexuality. They find a particular emphasis on sexual identity issues on Tumblr.

As for computational work that addresses identity in Tumblr, Grbovic et al. [19] build classifiers for user (binary) gender. To predict “correct” gender labels from names matched to a baby name database, they use profile features such as blog description and title n -grams, as well as behavioral features including tag use, liking, following and reblogging. In contrast, our work uses self-presented identity information to predict interaction.

3 RESEARCH CONTEXT: TUMBLR

With its design and user culture, Tumblr provides an ideal platform for studying the relationship between identity and content propagation. As of December 2018, when data was collected for this study, Tumblr had 167 billion posts over 450 million blogs. The site is advertised as a place to “express yourself”, “be yourself”, and “connect with your people”.¹ Self-presentation is important on Tumblr as there are no provided fields for listing personal information or centralized communities to officially belong to. Thus, the information other users can determine about a user from their blog is what is consciously expressed in that user’s own terms. Finally, Tumblr users are known for being interested in social justice issues which often intersect with identity [30].

The main form of content propagation on Tumblr is reblogging posts, analogous to retweeting on Twitter. Reblogged content appears on that user’s blog and often on the individualized “dashboards” (feeds of current activity) of followers of that blog.

3.1 Self-presentation on Tumblr

In this work we focus on blog descriptions and profile images, but Tumblr users present themselves through a variety of means on their personal blogs:

- **Blog Descriptions** (*used in this work for prediction and interpretation*). Users often provide short spans of identity information such as age, gender, sexual orientation, or interests in the free-text blog description field [30]. This description appears on the blog as well as with a mouseover of a user’s profile image next to posts (Figure 1).
- **Profile Images** (*used for prediction*). Each blog has either a profile image chosen by the owner or a default profile image. This image appears on the blog and also with posts appearing on other users’ dashboards.
- **Blog Names**. Blog names are often descriptive of the content or user behind the blog. They appear in the URL of the blog, as well as with a mouseover of a profile image in a dashboard.
- **Posts and Reblogs**. Recently posted material, displayed at the top of a blog, likely plays a role in the characterization of a blog to other users. This also motivates the influence of self-presentation on content propagation: reblogged content becomes part of users’ own self-presentation.
- **Other Blog Aesthetics**. Tumblr users can choose from a myriad of “themes” to customize the appearance of their blogs. Aesthetic choices include background images, wallpaper, fonts, borders around posts, and even mouse changes and default music for blog visitors.

¹<https://www.tumblr.com/about>

Users have the opportunity to make decisions about sharing content either when viewing another user’s blog or when viewing content on their dashboard. Information about the user is readily available when viewing a blog, and some information is available with a mouseover of the profile image next to a post on the dashboard (Figure 1). However, our analysis does not assume that users necessarily check this self-presentation information before choosing to reblog. We certainly do not assume that this information has more of an effect on reblog choices than the content of posts. Rather we test whether identity features have any statistical association with content propagation when we control for post content.

For this analysis, we use two prominent affordances for identity presentation: profile images and text blog descriptions. In a sample of 1 million blogs that made at least 10 reblogs from June through November 2018, 61.2% had filled in blog descriptions. We sampled profile images for 810,800 blogs from this set of 1 million (the remainder could not be accessed). Of these 810,800, 69.6% provided a non-default profile image.

Note that unlike the conventions in many other social media sites such as LinkedIn, Facebook, Twitter, and Instagram, many of these profile images on Tumblr do not contain images of the user. In a sample of 1,000 profile images from blogs who made at least 10 reblogs from June to November 2018 annotated by two of the authors, only 29.3% supplied a profile image of a person who is likely the user, whereas 42.5% supplied another type of image. This would make conventional gender, age and ethnicity image detection systems not as effective; we use continuous, dense, neural representations of profile images instead (see Section 5).

4 EXPERIMENTS

We design an experimental paradigm to evaluate whether similarities and differences in the identity presentation of two users are associated with patterns in their decisions to propagate each others’ content (**RQ 1**). This framework allows us to identify patterns in how self-presentation comes across in interaction, specifically associations with content propagation. We train logistic regression, SVM, and neural network models on features that represent identity alignment between users (matches, mismatches, and other interactions) and look for changes in accuracy over a baseline of post content features. A significant increase in accuracy suggests that comparisons and contrasts in self-presentation are relevant for predicting content propagation between users.

We treat the choice of reblogging as a selection problem, where a user is exposed to a wide variety of content and chooses to propagate some posts rather than others (see Figure 2). For example, if issues related to sexual orientation are important to a user and they present themselves in those terms, we expect that all else being equal, they would choose to reblog a post from a user who also signals sharing that value compared with another user they follow who does not share that identity framing. We use a pairwise learning-to-rank paradigm to embody such a comparison.

4.1 Learning-to-Rank Formulation

The learning-to-rank method we use in our experiments is a variant of the RankSVM algorithm [23]. The RankSVM algorithm enables the use of traditional classifiers, like support vector machines (SVM),

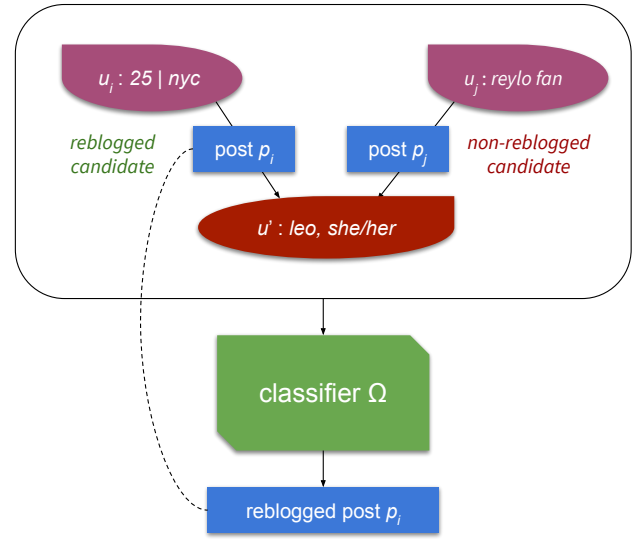


Figure 2: Illustration of the pairwise learning-to-rank formulation. From the perspective of user u' , we want the classifier Ω to decide whether to reblog post p_i or p_j from different users $u_{i,j}$ that user u' is following. Fabricated examples of text blog descriptions are included in the diagram, though profile images are also used in practice.

to make pairwise comparisons by considering items in a comparison feature space. Given a set of pairwise post comparisons P , their corresponding ranking labels R , and a classifier $\Omega(X, Y)$ that can be optimized given feature vectors X and corresponding labels Y , RankSVM performs the following transformation:

- (1) For every pairwise post comparison $\langle p_i, p_j \rangle \in P$:
 - (a) Map p_i and p_j into a common feature space \mathcal{F} using feature function $\Phi(p) : p \rightarrow \mathcal{F}$
 - (b) Create a feature vector representing the comparison by calculating the difference between the feature vectors for p_i and p_j . The resulting comparison feature vector $c_{ij} \in C$ is now in a pairwise comparison space \mathcal{F}'_C .

$$c_{ij} = \Phi(p_j) - \Phi(p_i)$$

- (2) Train a classifier $\Omega(C, R)$.

The trained $\Omega(C, R)$ takes a feature vector in the pairwise comparison space and produces a ranking between the pair of posts: rank 0 for the post the user reblogged, rank 1 for for the post they did not. While SVM with a linear kernel was traditionally used in RankSVM, the algorithm can be extended to other classifiers.

Reblog classification could be formed as a simple prediction task over a sample of posts that a user did or did not reblog. However, users on average reblog fewer than 1% out of all posts from blogs they follow [10], which leads to heavily skewed data. The pairwise learning-to-rank formulation addresses this issue by directly representing the reblogged/non-reblogged post comparisons, allowing us to rebalance the skewed dataset in a meaningful way. This is similar to previous work that handles highly skewed distributions

	Prediction dataset	Interpretation dataset
Identity features	text + image	text
# Users	14,177	34,801
# Reblog prediction instances	228,424	712,670

Table 1: Dataset statistics. The prediction dataset contains users who provide both text blog descriptions and profile images. The interpretation dataset contains users who provide a blog description but not necessarily a profile image.

in information cascade size prediction by constructing balanced binary classification tasks [12, 28].

4.2 Experimental Dataset

We use Tumblr posts and text blog descriptions from a data dump ranging from 1 June 2018 to 30 November 2018². Profile images for users were obtained from the Tumblr API.

To construct our dataset of paired reblogged and non-reblogged posts, we first sample a set of 1,000 blogs³, U , which have reblogged at least 10 posts as a minimum level of activity. For these 1,000 users, we find all users they have followed, F . We collect reblogged posts made by blogs in U from blogs in F after the associated user began following the blog in F . In the case of reblog cascades with content that is reblogged multiple times, we take identity features from the most recent reblogger of the content. These features are from the immediate user who the user follows, whose self-presentation information is more readily apparent than that of the original poster.

For each reblog in this set, we sample candidate non-reblogged posts to act as a comparison in our learning-to-rank framework. Since the details of Tumblr’s dashboard ranking algorithm are not public, we assume recent posts from followers likely appear on a user’s dashboard. We restrict both reblogged and non-reblogged candidate posts to only be from blogs the user follows. To increase the likelihood that paired non-reblogged posts were seen by the user, we select those that were posted within 30 minutes of each paired reblogged post. Since the user reblogged a post within 30 minutes, there is a greater chance they were active and saw these other, non-reblogged posts from blogs they follow. We sample up to five non-reblogged posts, from unique blogs, for every reblog.

From this initial data collection, we extract two datasets (Table 1). For prediction we use identity features from profile images and blog descriptions, and so we filter to users who provide both. For interpretation we only use text blog descriptions, and so we incorporate users who provide blog descriptions but not necessarily profile images into a slightly larger dataset.

Ranking labels are generated by randomly shuffling the order of the posts within each comparison so that the reblogged post appeared as p_i 50% of the time and p_j 50% of the time. However, we want feature weights in the model to be consistent and interpretable

²This period is before the adult content ban was announced in December 2018, so any changes due to this ban are not reflected in our data.

³Hereafter, we will also refer to blogs as *users*. While a Tumblr user may have multiple blogs associated with their account, for simplifying purposes, we consider users to be on the blog-level, as the identity labels in blog descriptions we examined apply to the user and not the blog.

(i.e. positive weights indicate a higher likelihood of reblogging). So in practice, we always treat the reblogged post as p_i but flip the sign on the label and features when it should be considered p_j . Each dataset is randomly split into a training/test split of 90% and 10%.

4.3 Model Hyperparameters

Logistic regression classifiers are trained with ℓ_2 regularization; constants are selected using grid search from 10^{-4} to 10^4 on a base 10 log scale on 10-fold cross-validation on the training set.

SVM models were trained with linear kernels due to the traditional use of linear SVM with RankSVM and the large size of our training set. ℓ_2 regularization constants were chosen using grid search from .01 to 100 on a base 10 log scale.

For our neural network, we use a multi-layer perceptron (MLP) over the same feature set as the logistic regression and SVM models. The MLP consists of three hidden layers of size 100, 50, and 32 with ReLU activation in each layer. We train this model with ℓ_2 regularization with a constant $C = 10^{-4}$. We used the Adam optimizer [26] with $\alpha = 0.001$, $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\epsilon = 10^{-8}$. The MLP was trained with early stopping, where 10% of the training data was randomly set aside as a validation set.

5 FEATURE EXTRACTION

We describe here the feature function $\Phi(p)$ that is applied to candidate posts p_i and p_j and their associated users in each pairwise comparison. Identity features are extracted from both text blog descriptions and profile images.

5.1 Text Blog Descriptions

We extract identity information from blog descriptions at two levels: (1) what specific identity labels, such as ‘trans man’ and ‘british’, are given, and (2) which broader identity categories, such as *gender* or *age*, those labels indicate. Our intuition is that providing similar categories of identity, even if labels are different, may orient users to the platform in similar ways. We use a bootstrapping approach to find labels that indicate identity and to group them into categories for automatic annotation.

Similarly to profile images, neural embeddings could be learned for blog descriptions and various similarity metrics computed for comparison between them. However, for interpreting which kinds of similarities and differences in self-presentation are associated with content propagation, we choose to extract one-hot features.

Bootstrapping delimiters. As a convention, Tumblr users often provide identity labels separated by delimiters (such as commas or pipes) in blog descriptions [30]. For example, note the pipes as delimiters in the fabricated blog description *22yo | she/they | too many fandoms* and the periods as delimiters in *andre . nyc . manga*. To identify such delimiters, we started with a small list of identity labels manually identified from blog descriptions. We then searched for these labels in a separate set of blog descriptions and found characters in between these terms as potential delimiters. Manually reviewing this list of potential delimiters, we kept those that could function as separators between labels in a list, primarily punctuation and emojis. This resulted in a list of 95 delimiters.

Identity Category	Label Examples
age	21, seventeen
ethnicity/nationality	latina, haitian
fandoms	shipping, crossovers, star wars, lotr
gender	woman, husband, mtf, nonbinary
interests	photography, running, makeup
location	australia, london, socat
personality type	intp, slytherin
pronouns	she/her, they
relationship status	married, single
sexual orientation	bi, lesbian, aro-ace
zodiac	virgo, capricorn

Table 2: Our identity categories with examples of labels.

Bootstrapping identity labels. To find additional identity labels, we extracted short texts (maximum 25 characters) in between any of these delimiters on the larger set of blog descriptions. Long spans often indicated quotes or other unrelated material. We extracted and ranked identity label candidate n -grams in these short segments by frequency, discarding stopwords and other terms that were not indicative of identity.

Choosing identity categories. We want categories that are:

- (1) Popular, and thus relevant on Tumblr.
- (2) Largely about the user, not the content. Our goal is to identify how users position themselves, not the main topics of a blog.
- (3) A relatively limited set of possible values so they can be accurately identified outside of a training set.

Guided by Bucholtz and Hall [7], we manually grouped labels into categories that encompassed broad demographics as well as labels more specific to Tumblr (such as fandoms and interests). Some of these popular labels are creative and not well-known outside of Tumblr, such as ‘phans’ for fans of Phil Lester, a YouTube personality, and ‘stans’ for obsessive fans. Others are more specific to Tumblr but intersect with larger demographic categories, such as ‘cishets’ for cisgender heterosexuals. Our final list of 11 identity categories, with example labels, is shown in Table 2.

Annotating identity categories. Two authors manually annotated for the presence of these categories in a random sample of 1200 blog descriptions, from which we pulled 100 samples for a development set and 100 for a test set. From our manual annotations, we built regular expressions to automate annotation. On the development set, we iteratively added or refined regular expression patterns for each category. Sometimes it was unclear which category a label indicated, such as ‘LGBT’ indicating gender, sexual orientation, or both. In these cases we added the pattern to all categories that may be indicated (both gender and sexual orientation for ‘LGBT’).

For a subset of identity categories, we compared our regular expression annotation approach with Naive Bayes and SVM models. We trained on unigrams and character 1-4grams in the blog descriptions to predict which categories are present. We found that the regular expressions performed better on the test set (over 80% F1 average compared with under 50%), likely due to the small amount

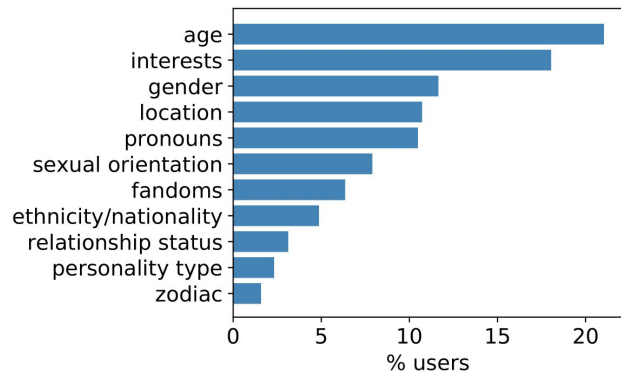


Figure 3: Proportion of users in our interpretation dataset who present each identity category.

of available training data. Percentages of users in our interpretation data training set who present each category are shown in Figure 3.

5.1.1 Category Alignment Features. Let u' be the user making the comparison and u_* be the user associated with the candidate post p_* when applying $\Phi(p_*)$. Features included are listed below.

- **Category Match** (c): A binary variable indicating if u' and u_* both provide identity category c .
- **Category Mismatch** (c): A binary variable indicating if only one of u' or u_* provides identity category c .
- **Directional Category Mismatch** (c, u', u_*): Directional version of **Category Mismatch** (c) indicating if u' provided identity category c but not u_* , or vice versa.

5.1.2 Label Alignment Features.

- **Label Match** (c): A count variable indicating the number of labels used by both u' and u_* in category c .
- **Label Mismatch** (c): A count variable indicating the number of labels that are unique to u' + the number of labels that are unique to u_* for category c .
- **Label Interaction** (l', l_*, c): A count variable of how many times u' used label l' and u_* used l_* for category c .⁴

All text features are normalized to have unit variance over the training set after generating comparison feature set C .

5.2 Profile Images

To extract features from profile images, we used the 1,000-dimension layer before the softmax layer from ResNet-152 [21], a popular computer vision benchmark that was pretrained on 1,000 image categories from ILSVRC-2012 [32]. Since we are interested in alignment between followers and users they follow, we use three different comparisons between profile image embeddings: **Cosine Similarity**, **Euclidean Similarity**, and **Vector Difference** (element-wise subtraction between vectors).

⁴Label interactions across categories were not considered since this would lead to a very large input dimensionality

	LR	SVM	MLP
Content	64.65	64.60	66.28
Content + text	77.66*	77.52*	80.86*
Content + image	81.99*	81.81*	89.97*
Content + text + image	87.72*	87.56*	92.81*

Table 3: Learning-to-rank accuracy with text blog description and profile image identity alignment features. Text identity features include category and label alignment. * $p < 0.05$ compared to the content features (McNemar’s test). A random baseline would achieve 50% accuracy.

5.3 Content Features

Post content features likely capture much of the signal in content propagation. In our experiments, we look for any additional signal provided by self-presentation alignment features above these baseline features. Note that we do not use actual post content of text, images, or videos. The wide variety of formats that post contents take make this inclusion non-trivial, and so we use features that are uniform across all post types.

- **Post Tags:** Post hashtags, similarly used by Naveed et al. [29] to represent post content in predicting content propagation. All post tags are lowercased and only tags that are used by more than one user are considered in the tag vocabulary. Post tag features are binary variables indicating whether the post contains tag in the vocabulary.
- **Note Count:** A count variable indicating the number of notes (likes, reblogs, and comments) attached to the post. This is included to control for popularity.
- **Post Type:** A categorical variable indicating the type of the post. Posts can be of type text, photo, quote, video, audio, chat, link, and answer. Xie et al. [38] also use the type of post as their content features in content propagation prediction.

6 PREDICTION MODEL RESULTS

Results on the reblog prediction task are shown in Table 3 for logistic regression, SVM, and MLP models. We find significant performance increases with the addition of identity alignment features from both blog descriptions and profile images, evidence for an association between identity presentation and content propagation on Tumblr.

Profile image alignment provides a stronger signal for reblog prediction than alignment in the text blog descriptions. This could be due to the visual focus of Tumblr [39], the greater expressiveness of the feature vectors for image alignment, or both. It is also apparent that visual and text self-presentation are complementary, as performance improves with the combination of both signals.

The MLP is the best-performing model on reblog prediction. This may be explained by its ability to exploit non-linear combinations of content features with identity information (e.g. when users list a fandom and a post contains related hashtags).

7 INTERPRETATION

Overall, we find that alignment in self-presented identity labels is associated with content propagation on Tumblr. However, this tells

Category	Category Features	Label Features
<i>Content features only</i>		
	62.69	62.69
+ age	63.25*	66.29*
+ ethnicity/nationality	62.78	63.97*
+ fandoms	62.80	63.49*
+ gender	62.80	64.51*
+ interests	63.35*	65.82*
+ location	63.10*	65.03*
+ personality type	62.69	63.00*
+ pronouns	63.05*	63.89*
+ relationship status	62.82	63.17*
+ sexual orientation	63.10*	63.59*
+ zodiac	62.98*	63.07*
+ all	64.72*	74.30*

Table 4: Learning-to-rank reblog prediction accuracy using logistic regression on text blog description features for interpretation. Category/Label Features refers to content + category or content + label features. Each row refers to a separate model trained only on the features for that identity category. * $p < 0.05$ compared to a baseline of only content features.

us little about the nature of this effect. For example, are users who present similar information more likely to reblog each other’s content? Are users who present dissimilar images and labels less likely to reblog each other’s content? Or do more unexpected interactions play an important role?

To investigate the nature of this effect (RQ 2), we focus on the text blog description features. Our image representations are dense, continuous vectors that are more difficult to interpret. Although there have been efforts to improve the interpretability of visual inference models [25, 33], existing techniques are difficult to apply to our learning-to-rank setup; we leave this to future work. We train logistic regression models separately across categories on a larger dataset including users who provide blog descriptions but not necessarily profile images. See Table 1 for details on this dataset and Table 4 for results from these models.

In general, we find that users who give matching labels and categories are more likely to reblog each other. We also find that users who give labels that indicate shared values around issues such as conceptions of gender and sexuality, and shared interests such as around visual content, are more likely to reblog each others’ content. From the effects we see on content propagation, identity presentations on Tumblr seem to come across to establish solidarity and common ground. Specific findings are discussed below.

7.1 Category vs. Label Alignment Effects

The effects of category features are inconsistent across categories (Table 4). Five out of the 11 categories do not show significant improvement over the content baseline, while 6 categories do improve: interests, age, pronouns, location, sexual orientation, and zodiac. For some categories, this is likely due to the presence of the category itself indicating an alignment of shared values or interests. Providing pronouns can indicate shared conceptualizations

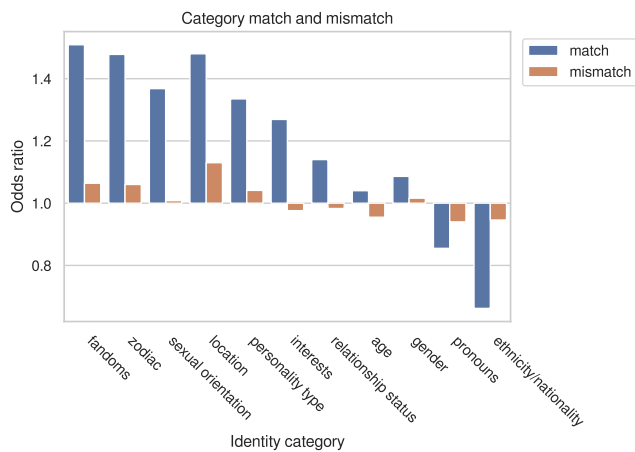


Figure 4: Odds ratios for the Category Match and Category Mismatch features from logistic regression models trained separately across identity categories. Categories are sorted by the difference between match and mismatch.

of gender, providing any location an interest in local (often visual) content, and any zodiac label indicates an interest in astrology.

For other categories, labels are skewed such that providing the category acts as a proxy for providing popular labels that indicate shared values or experiences. For example, 54% of users in our training set who provide an age present an age from 18 to 22, and 30% of users who provide interests list visual interests such as ‘art’, ‘draw’, or ‘photos’. For sexual orientation, only about 10% of users provide labels that indicate being straight, so providing any sexual orientation indicates likely belonging to the LGBTQ community.

Label alignment features significantly improve performance over content features (Table 4), suggesting that matches and mismatches of identity labels are associated with content propagation. Though category alignment features for the ethnicity/nationality, fandoms, gender, and relationship status categories did not significantly improve over the post content baseline, the use of label alignment features in these categories do lead to significant improvements. For these categories, we hypothesize that distinctions between specific label alignments are necessary to indicate shared values or experiences, rather than simply framing one’s participation in a kind of interaction by listing any value in the category. For example, presenting any common gender label, such as ‘male’ or ‘female’, does not express an ideological position, whereas giving pronouns can express an ideology on gender issues.

7.2 Category Alignment Interpretation

For all categories except for pronouns and ethnicity/nationality, models trained with only baseline and category alignment features learned positive weights on the **Category Match** feature (Figure 4). This indicates that users are more likely to reblog content from other users who present the same category. Listing one’s sexual orientation or interests—regardless of the labels used in these categories—signals that these categories are important to users’ self-presentation.

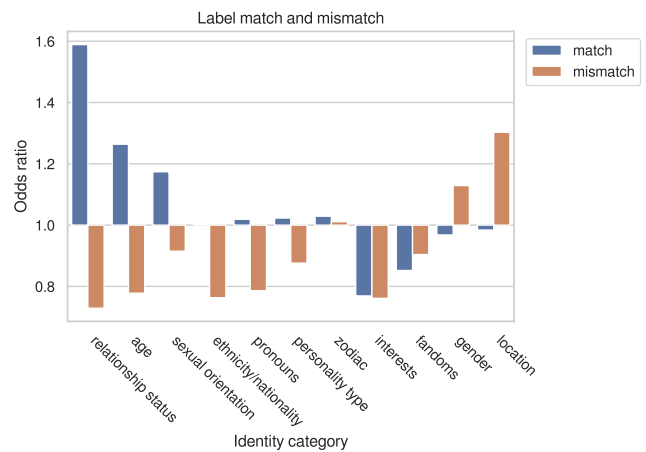


Figure 5: Odds ratios for the Label Match and Label Mismatch features from logistic regression models trained separately across identity categories.

The model trained on pronouns placed a negative weight on category match: both users giving pronouns was associated with a slight decrease in the likelihood of reblogging (odds ratio = 0.855). However, if the user deciding whether to reblog presents pronouns while the user they follow does not, the follower is less likely to reblog this post (OR = 0.781). There have been calls from transgender activists for cisgender people to share pronouns to normalize the practice of not assuming gender.⁵ Not giving pronouns may signal that a user does not share a view of gender that encourages listing pronouns, and we see an association with content propagation.

Similar trends are observed for ethnicity/nationality, where followers who present any ethnic or national label are less likely to reblog users who do not present such labels (OR = 0.609), a stronger association than the negative weight learned on reblogging between users who both present the category (OR = 0.662).

For the zodiac category, the model placed most positive weight on the directional category alignment mismatch in which the user choosing to reblog presented zodiac but the user providing the post did not (OR = 2.395). This suggests that users who list zodiac are more open to propagating content from blogs that do not list zodiac.

In almost all of these categories that showed significant improvement, negative weight was placed on the directional mismatch in which the user making a decision to reblog does not provide a category, while the user they follow does. If a user does not place value in presenting an identity category, they appear less likely to propagate content from a user who does.

7.3 Label Alignment Interpretation

Models using only baseline and label alignment features learned positive weight on the **Label Match** feature for most categories (Figure 5). Most models learned negative weight on the **Label Mismatch** feature. This provides evidence that matches in identity labels generally increase, and mismatches decrease, the likelihood of reblogging.

⁵<https://www.glaad.org/transgender/allies>, accessed 19 February 2020.

However, weights learned on specific **Label Interaction** features were often higher in absolute value, and thus more informative, than matches or mismatches. Looking more closely at these interactions, users who present labels that indicate similar interests in content often are more likely to reblog each other. For fandoms, one user listing ‘star wars’ and the other listing ‘reylo’, a popular character pairing in Star Wars, increased the likelihood of reblogging. For the interests category, users who presented ‘anime’ were more likely to reblog those who presented ‘design’. Users who presented ‘gaming’ as an interest were more likely to reblog those who presented ‘manga’. The model placed negative weight on some specific interactions that indicated differences in tone, such as users who listed ‘memes’ as interests reblogging those who list ‘history’.

Other times specific interactions likely indicative of shared experiences were more informative. For example, in the age category, users presenting similar ages (e.g. users who present an age of 20 following users who present 21), were more likely to reblog each other’s content. Positive weight was learned on interaction features of trans-identified users reblogging users who give no gender terms (OR = 1.23) and those who give other trans labels (OR = 1.06). The same holds for users giving ‘non-binary’ as a label: they were slightly more likely to reblog those who give no gender label (OR = 1.13). This may point to a preference for content from users who do not specify terms that are explicitly on the gender binary, which forms the bulk of the extracted gender labels. Similarly, those listing ‘queer’ were less likely to reblog others who present no term for sexual orientation and who may be more often straight without giving a label (OR = 0.85). Users presenting ‘straight’ were more likely to reblog others presenting ‘straight’ (OR = 1.25). Explicitly cis-identified users were slightly more likely to reblog content from other cis-identified blogs (OR = 1.02). Note that this feature only applied to users who explicitly identified as cisgender, not users who simply did not give any transgender labels. Identifying as ‘cis’, a relatively new and rarer term for gender, likely shows a knowledge of the discourse around transgender issues.

8 LIMITATIONS AND FUTURE WORK

In this study, we only observe user behavior (posting profile images, writing blog descriptions, and sharing posts) and use statistical models to understand this behavior at scale. User interviews or surveys may provide more insight on motivations for these behaviors. For example, what do users themselves say they are trying to signal by using certain categories or labels in blog descriptions?

We use a regression-based analysis on records of naturally occurring interaction to find associations between identity alignment and content propagation. While this approach allows us to make correlational conclusions, we cannot draw any conclusions regarding causality. Though we sample scenarios where it is likely that users will see both posts in our comparisons, Tumblr’s own ranking of posts in users’ dashboards has an effect that is difficult to measure. Other factors such as what users are known as authorities on certain issues may also have an effect on reblogging behavior. A larger sample of users may provide a more comprehensive picture of the relationship between identity and content propagation on Tumblr. Narrowing in on particular communities within Tumblr could reveal community-specific responses to self-presentation.

Since our datasets contained up to five data points for each reblogged post, each time paired with a different non-reblogged post, there is a concern that the model could somehow be memorizing reblogged posts. We split our dataset randomly, and so this would especially be a concern if a reblog appears in both training and test sets paired with different non-reblogged posts. However, we checked the results on a smaller sample without duplicate reblogged posts and the results and feature weights were very similar.

In future work, profile image alignment could be interpreted based on the 1,000 ILSVRC-2012 categories [32]. We could also explore image features relevant to Tumblr profile images, such as drawings and cartoons versus human faces, or facial expressions and emotions [4]. Complementary effects from visual and textual self-presentation could then be investigated. Identity alignment features could also be extracted from blog names.

Nonlinear feature combinations used by the MLP could be investigated. For example, do certain post hashtags increase the likelihood of reblogging in combination with profile images or blog description features? Tumblr has a particular emphasis on identity, so it remains to be seen what effects hold in other social media contexts.

9 ETHICS AND PRIVACY

Tumblr profile images and blog descriptions contain sensitive personal information, so care was taken to protect users’ privacy and remove the possibility of identifying any blogs in this study or amplifying any content from these blogs [14]. We only included public Tumblr blogs accessible without a Tumblr log-in. All examples of blog descriptions, except the staff Fandom account in Figure 1, have been fabricated so as to not easily trace back to any individual blog [6]. Though researchers did view blog descriptions and profile images, these descriptions were not matched with blog names or URLs, which were also never used in analysis.

Regular expression patterns for identity label annotation were constructed from aggregate blog description n -grams, and are not shared. Only labels occurring in more than one blog description were extracted to protect any users who used unique labels. Similarly, hashtags were only considered in aggregate, and we removed any hashtags used by only one blog for feature extraction.

We did not construct classifiers that predict user identity attributes from text or images. However, our classifiers did include self-presentation features in predicting content propagation. This approach could possibly be used for targeted marketing, though this is not the intended purpose of this work, and no classifiers, data, or feature extractors have been made publicly available.

10 CONCLUSION

To explore the effect of identity-based features on content propagation on Tumblr, we constructed a machine learning task predicting which posts users propagate among posts they would have likely seen. We found that features from profile images and text descriptions were informative for predicting content propagation between users. Visual and textual self-presentation information provided complementary signals. Investigating the nature of this effect for text features, we found that users who presented similar identity categories and labels were generally more likely to reblog each other’s content. Specific interactions between labels were also an

informative part of this signal; users who presented labels that indicated shared interests in content or shared values around gender and sexuality, for example, were more likely to reblog each other.

These results suggest that homophily may support content propagation, though we caution that the pattern we have identified is only correlational. This raises the question of what the most effective balance between homophily and diversity would be to support content propagation while encouraging communication across subcommunities.

ACKNOWLEDGMENTS

Thanks to Amanda Brennan, Nicola Barbieri, Beita Li, Sherlan Moriah, and other employees at Tumblr for inspiration and conversations. This work was supported by NSF Grant IIS 154393 and a K&L Gates Presidential Fellowship.

REFERENCES

- [1] Nora Alrajebah, Leslie Carr, Markus Luczak-roesch, and Thanassis Tiropanis. 2017. Deconstructing Diffusion on Tumblr: Structural and Temporal Aspects. In *Proceedings of the 9th ACM Conference on Web Science*. 319–328.
- [2] Rose Attu and Melissa Terras. 2017. What People Study When They Study Tumblr: Classifying Tumblr-related Academic Research. *Journal of Documentation* 73, 3 (2017), 528–554.
- [3] Saeideh Bakhshi, David A Shamma, Lyndon Kennedy, Yale Song, Paloma De Juan, and Joseph 'Jofish' Kaye. 2016. Fast, cheap, and good: Why animated GIFs engage us. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 575–586.
- [4] Tadas Baltrusaitis, Amir Zadeh, Yao Chong Lim, and Louis-Philippe Morency. 2018. Openface 2.0: Facial behavior analysis toolkit. In *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*. IEEE, 59–66.
- [5] Liad Bareket-Bojmel, Simone Moran, and Golan Shahar. 2016. Strategic Self-presentation on Facebook: Personal Motives and Audience Response to Online Behavior. *Computers in Human Behavior* 55 (2016), 788–795.
- [6] Amy Bruckman. 2002. Studying the Amateur Artist: A Perspective on Disguising Data Collected in Human Subjects Research on the Internet. *Ethics and Information Technology* 4, 3 (2002), 207–231.
- [7] Mary Bucholtz and Kira Hall. 2005. Identity and Interaction: A Sociocultural Linguistic Approach. *Discourse Studies* 7, 4-5 (2005), 585–614.
- [8] Liam Bullingham and Ana C. Vasconcelos. 2013. 'The Presentation of Self in the Online World': Goffman and Study of Online Identities. *Journal of Information Science* 39, 1 (2013), 101–112.
- [9] Judith Butler. 1990. *Gender Trouble: Feminism and the Subversion of Identity*. Routledge.
- [10] Yi Chang, Lei Tang, Yoshiyuki Inagaki, and Yan Liu. 2014. What is Tumblr: A Statistical Overview and Comparison. *ACM SIGKDD Explorations* 26, 1 (2014), 21–29.
- [11] Xin Chen, Yu Wang, Eugene Agichtein, and Fusheng Wang. 2015. A Comparative Study of Demographic Attribute Inference in Twitter. In *Proceedings of the Ninth International AAAI Conference on Web and Social Media*. 590–593.
- [12] Justin Cheng, Lada A. Adamic, P. Alex Dow, Jon Kleinberg, and Jure Leskovec. 2014. Can cascades be predicted? *Proceedings of the 23rd International Conference on World Wide Web* (2014), 925–935.
- [13] Michael A. Devito, Ashley Marie Walker, and Jeremy Birnholtz. 2018. "Too Gay for Facebook": Presenting LGBTQ+ Identity Throughout the Personal Social Media Ecosystem. In *Proceedings of the ACM on Human-Computer Interaction, Vol. 2 - CSCW*.
- [14] Casey Fiesler and Nicholas Proferes. 2018. "Participant" Perceptions of Twitter Research Ethics. *Social Media + Society* 4, 1 (2018), 1–14.
- [15] Marty Fink and Quinn Miller. 2014. Trans Media Moments. *Television & New Media* 15, 7 (2014), 611–626.
- [16] James Paul Gee. 2011. *An Introduction to Discourse Analysis: Theory and Method*. Routledge.
- [17] Erving Goffman. 1959. *The Presentation of Self in Everyday Life*. Doubleday.
- [18] Neil Zhenqiang Gong, Ameet Talwalkar, Lester Mackey, Ling Huang, Eui Chul Richard Shin, Emil Stefanov, Elaine Runtong Shi, and Dawn Song. 2014. Joint Link Prediction and Attribute Inference using a Social-Attribute Network. *ACM Transactions on Intelligent Systems and Technology (TIIST)* 5, 2 (2014), 27.
- [19] Mihajlo Grbovic, Vladan Radosavljevic, Nemanja Djuric, Narayan Bhamidipati, and Ananth Nagarajan. 2015. Gender and interest targeting for sponsored post advertising at tumblr. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 1819–1828.
- [20] Oliver L. Haimson and Gillian R. Hayes. 2017. Changes in Social Media Affect, Disclosure, and Sociality for a Sample of Transgender Americans in 2016's Political Climate. In *Proceedings of the Eleventh International AAAI Conference on Web and Social Media*. 72–81.
- [21] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 770–778.
- [22] Bernie Hogan. 2010. The Presentation of Self in the Age of Social Media: Distinguishing Performances and Exhibitions Online. *Bulletin of Science, Technology & Society* 30, 6 (2010), 377–386.
- [23] Thorsten Joachims. 2002. Optimizing Search Engines using Clickthrough Data. In *Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 133–142.
- [24] Barbara Johnstone. 2010. Locating Language in Identity. In *Language and Identities*. Edinburgh University Press.
- [25] Been Kim, Martin Wattenberg, Justin Gilmer, Carrie Cai, James Wexler, Fernanda Viegas, and Rory Sayres. 2017. Interpretability beyond feature attribution: Quantitative testing with concept activation vectors (tcav). In *International Conference on Machine Learning*.
- [26] Diederik P Kingma and Jimmy Ba. 2015. Adam: A method for stochastic optimization. In *Proceedings of the 3rd International Conference for Learning Representations*.
- [27] Zornitsa Kozareva and Makoto Yamada. 2016. Which Tumblr Post Should I Read Next?. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*. 332–336.
- [28] Siddharth Krishnan, Patrick Butler, Ravi Tandon, Jure Leskovec, and Naren Ramakrishnan. 2016. Seeing the forest for the trees: New approaches to forecasting cascades. *Proceedings of the 8th ACM Conference on Web Science* (2016), 249–258.
- [29] Nasir Naveed, Thomas Gotttron, Jérôme Kunegis, and Arifah Che Alhadi. 2011. Bad News Travel Fast: A Content-based Analysis of Interestingness on Twitter. In *Proceedings of the 3rd International Web Science Conference*.
- [30] Abigail Oakley. 2016. Disturbing Hegemonic Discourse: Nonbinary Gender and Sexual Orientation Labeling on Tumblr. *Social Media + Society* 2, 3 (2016), 1–12.
- [31] Mario Piergallini, A Seza Doğruöz, Phani Gadde, David Adamson, and Carolyn Rose. 2014. Modeling the Use of Graffiti Style Features to Signal Social Relations within a Multi-Domain Learning Paradigm. In *Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics*. 107–115.
- [32] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Sathesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg, and Li Fei-Fei. 2015. ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision* 115, 3 (2015), 211–252.
- [33] Ramprasaath R Selvaraju, Michael Cogswell, Abhishek Das, Ramakrishna Vedantam, Devi Parikh, and Dhruv Batra. 2017. Grad-cam: Visual explanations from deep networks via gradient-based localization. In *Proceedings of the IEEE International Conference on Computer Vision*. 618–626.
- [34] Katrin Tiidenberg. 2014. Bringing Sexy Back: Reclaiming the Body Aesthetic via Self-shooting. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace* 8, 1 (2014).
- [35] Balazs Vedres and Orsolya Vasarhelyi. 2019. Gendered behavior as a disadvantage in open source software development. *EPJ Data Science* 8, 1 (2019).
- [36] Sorous Vosoughi, Deb Roy, and Sinan Aral. 2018. The Spread of True and False News Online. *Science* 359, 6380 (2018), 1146–1151.
- [37] Zijian Wang and David Jurgens. 2018. It's Going to be Okay: Measuring Access to Support in Online Communities. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*. 33–45.
- [38] Daniel Xie, Jiejun Xu, and Tsai-Ching Lu. 2017. What's Trending Tomorrow, Today: Using Early Adopters to Discover Popular Posts on Tumblr. In *Proceedings of the 2017 IEEE International Conference on Big Data*. 2168–2176.
- [39] Jiejun Xu, Ryan Compton, Tsai-Ching Lu, and David Allen. 2014. Rolling through Tumblr: Characterizing Behavioral Patterns of the Microblogging Platform. In *Proceedings of the 6th ACM Conference on Web Science*. 13–22.
- [40] Jiejun Xu and Tsai-Ching Lu. 2015. Inferring User Interests on Tumblr. In *Proceedings of the International Conference on Social Computing, Behavioral-Cultural Modeling, and Prediction*. 458–463.
- [41] Robert J. Zeglin and Julie Mitchell. 2014. Using Social Media to Assess Conceptualizations of Sexuality. *American Journal of Sexuality Education* 9, 3 (2014), 276–291.
- [42] Lumin Zhang, Jian Pei, Yan Jia, Bin Zhou, and Xiang Wang. 2014. Do Neighbor Buddies Make a Difference in Reblog Likelihood?: An Analysis on SINA Weibo Data. In *Proceedings of the 2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*. 208–215.
- [43] Qi Zhang, Yeyun Gong, Jindou Wu, Haoran Huang, and Xuanjing Huang. 2016. Retweet Prediction with Attention-based Deep Neural Network. In *Proceedings of the 25th ACM International Conference on Information and Knowledge Management*. 75–84.