TWEETBUBBLE: A TWITTER EXTENSION STIMULATES EXPLORATORY BROWSING

A Thesis

by

AJIT JAIN

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	Andruid Kerne
Committee Members,	James Caverlee
	Steven M. Smith
Head of Department,	Dilma Da Silva

December 2014

Major Subject: Computer Science

ABSTRACT

We extend the Twitter interface to stimulate exploratory browsing of social media and develop a method to establish its efficacy. In *exploratory browsing*, users seek and traverse diverse and novel information as they investigate a conceptual space. Social media has become a predominant source of information. Social media is characterized by rapidly evolving content and relationships. We need interface mechanisms to effectively and iteratively browse, search, and filter, i.e., *explore* vast social media networks.

We develop the TweetBubble browser extension, extending Twitter to enable expansion of social media associations—`@usernames` and `#hashtags`—in-context, without overwriting initial content. We build on a prior metadata type system, developing new presentation semantics, which enable a look and feel consistent with Twitter.

We identify exploratory browsing as a creative ideation process. We use prior ideation metrics as a basis for new ideation metrics of exploratory browsing. We conducted a study, with data from 54 participants, amidst the 2014 Academy Awards. Quantitative and qualitative findings validate the technique of in-context exploratory browsing interfaces for social media. Their consistency supports the validity of ideation metrics of exploratory browsing as an evaluation methodology for interactive systems designed to promote creative engagement.

This research impacts the design and evaluation of interfaces that stimulate intrapersonal creativity, and thereby mutual understanding, by supporting exploratory browsing of connected perspectives in a shared, structured, conceptual space.

ACKNOWLEDGEMENTS

I express my immense gratitude to Dr. Andruid Kerne, for his continuous guidance, support, and motivation. He wants his students to achieve only the very best. I admire in him, his enthusiasm about each and every thing he does, including research, discussion, and programming with his students. I have learned immensely in last three years, under his supervision, and I wish to match him up in skills, knowledge, and dedication.

I am extremely grateful to Dr. James Caverlee and Dr. Steven M. Smith, for the wonderful guidance and teaching I received from them. It always amazes me how they transform teaching of science with such engaging methodology. I have really been lucky to have studied under and been guided by such an elite panel.

A special thanks to my lab mates at the Interface Ecology Lab: Yin Qu, Nic Lupfer, Rhema Linder, and Andrew Webb. I really enjoyed working with them and received an enormous amount of help from each. I would also like to thank Bill Hamilton and Shenfeng Fei who have always been there to help, in whichever way possible.

I would like to thank Kade Keith, Cameron Hill, and the University Writing Center for their great help in reviewing different pieces of this content. I also want to thank Shannon Kerne who helped me by lending her voice to the TweetBubble video.

I would like to thank Dr. Zach Toups, under whose supervision I started my research work at the Interface Ecology Lab, for helping me quickly get an understanding of the lab's research. Once again, I want to thank all the Interface Ecology Lab members for always making me feel at home while I set out on this ardent trek

of research.

I thank my parents, for their relentless love and faith, which always made me see things in a positive light. I thank all my family and friends for their continuous love and support. Finally, I thank Shirdi Sai for blessing me always.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

TABLE Page

1. INTRODUCTION

We extend the Twitter interface to stimulate exploratory browsing of social media and develop a method to establish its efficacy. According to OED, to browse is "to feed on the leaves and shoots of trees and bushes; to crop the shoots or tender parts of rough plants for food: said of goats, deer, cattle" (Figure 1.1). We define *exploratory browsing* as a process through which users seek and traverse diverse and novel information as they use the web to investigate a topic or conceptual space. Browsing involves reading, surveying, sampling, and responding to interesting phenomena [61].

The Twitter microblogging service is a vital medium for news, politics, scholarship, and other social discourse [37]. According to Boden, exploratory creativity arises through traversing a structured conceptual space [5], such as the hypermedia

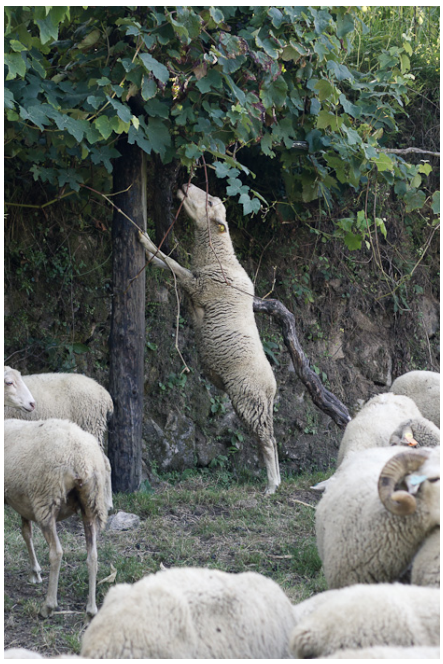

Figure 1.1: Browsing in practice (source: flickr [3]).

of Twitter. This sense of exploration as a creative process motivates our derivation of a new method for evaluating exploratory browsing interfaces for social media. Exploration enables users to learn about events, discover trending topics, and connect with others across the social network.

Twitter provides two primary mechanisms for making *social media associations* within a tweet. `@username` references enable user-to-user exchanges [26]. Hashtags in Twitter are short strings prepend with a hash (#) symbol. Tags in general are used to collectively organize and categorize social media content on topics, spontaneously creating *folksonomy* [44]. Fallon and Timberlake demonstrate how widely recognized hashtagging has become in popular culture [19].

Ideation is the creative generation and development of new perspectives and ideas. Creativity occurs on various scales. Mini-c creativity simply involves personally novel and significant interpretation of experiences, actions, and events [34]. It is inherent in learning. Little-c creativity is interpersonal, involves contribution, and arises through social dialogue and collaboration. We observe that Twitter users develop multiple perspectives through exploratory browsing of `@username` and `#hashtag` social media associations. This provides a basis for mini-c ideation, while little-c arises through following, favoriting, and retweeting.

This paper contributes quantitative methods to measure the mini-c ideation that arises through people's meaningful experiences of browsing social media. Work by Kerne et al. on information-based ideation investigated creative products that people assemble by gathering and annotating content [35]. The present research recontextualizes their metrics, previously used to evaluate creative products authored in ideation activities, to derive *ideation metrics of exploratory browsing* for Twitter. We use these metrics to compare our Twitter interface extension, *TweetBubble*, to the prior Twitter interface.

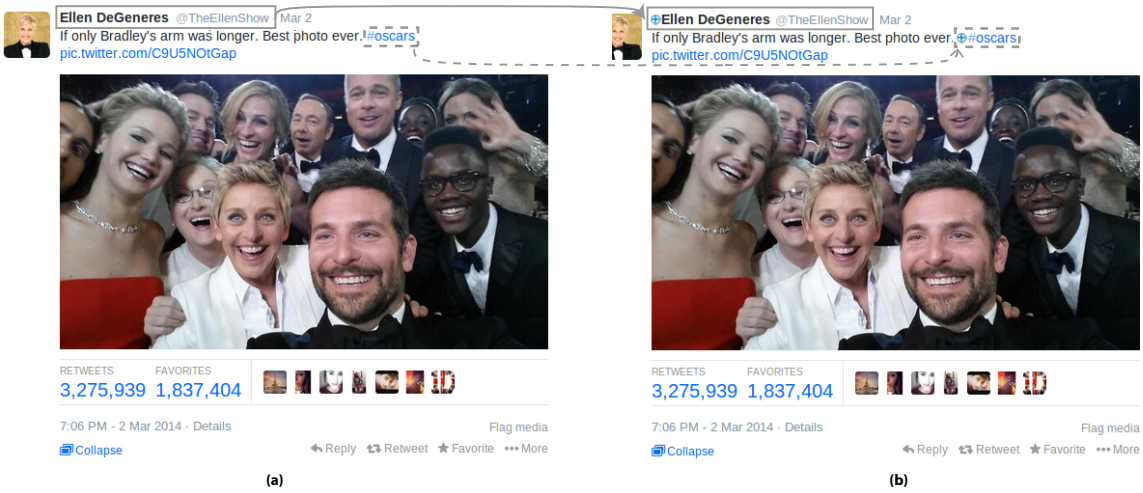

Figure 1.2: Social media association interfaces for Twitter. (a) Prior Twitter interface. Clicking `#hashtag` and `@username` overwrites content in the current window. The user experiences a loss of context. (b) TweetBubble interface prepends expansion affordances (+ sign in circle) to `#hashtag` and `@username` associations. Clicking results in retrieval and presentation in the current context, without overwriting initially rendered content.

When users click social media association links using the prior Twitter interface, initially rendered content is overwritten (Figure 1.2a). They lose context due to disorientation and digression. *Disorientation* is the problem of getting lost in a network of linked information [11]. Disorientation leads to *digression*, the problem of losing focus and going off track [21]. Whether browsing in a single view or with multiple windows and tabs, the paradigm of replacing the current view when following a link challenges human cognition. When users engage in exploratory browsing, they follow multiple, extended chains of hyperlinks. Disorientation and digression grow severe. As working memory is limited to ∼4 items at a time [12], cognitive load often exceeds the breaking point.

Dynamic exploratory browsing interfaces [48] enable traversal of hyperlinked associations while maintaining context. Multiple web pages are simultaneously presented as connected summaries. Qu et al. studied an exploratory browsing interface's effect

Figure 1.3: In-context expansion of social media associations with the TweetBubble Chrome Extension. `@username` and `#hashtag` feeds are now rendered without replacement of initial content.

on students seeking prior work for a paper [48].

The present research alternatively investigates the impact of dynamic exploratory browsing interfaces in everyday engagement with social media. The TweetBubble dynamic exploratory browsing interface subtly transforms the prior Twitter interface, expanding social media associations in context (Figure 1.2b). Clicking a hyperlink

results in contextualized retrieval and presentation of associated social media, via expansion instead of replacement (Figure 1.3).

In sections 2 and 3, we discuss exploratory browsing interfaces and everyday use of social media, respectively. In section 4, we discuss information-based ideation and its applicability in everyday use of social media. In section 5, we show how the TweetBubble exploratory browsing interface system is developed by extending the metadata type system of Qu et al. [48]. In section 6, we present a mixed methods study of the TweetBubble's exploratory interface, with 54 users, on the evening of the 2014 Academy Awards. We combine ideation metrics of curation with qualitative data to show how the exploratory browsing interface stimulates creative exploration of social media. We finish with a discussion on the findings and implications for design, in section 7.

2. EXPLORATORY BROWSING SYSTEMS

In this section, we are going to discuss how external cognition, through means of information visualization, helps mitigate issues of disorientation and digression in exploratory browsing. We will discuss dynamic exploratory browsing interfaces which enable users to see connected web pages in perspective.

2.1 Disorientation and Digression

In section 1, we defined exploratory browsing as a process through which users seek and traverse diverse and novel information as they use the web to investigate a topic. When people engage in exploratory browsing, they follow several trails of hyperlinked information, which causes them to lose track of their exploration. They lose context due to the issues of disorientation and digression.

Disorientation and digression, along with cognitive load, are long standing issues in hypertext systems [11, 21]. Also, the requirement of showing the underlying network structure of the hyperlinked entities is not addressed by current browsing systems. These issues cause an impoverished cognition of a topic as a whole and reduce the extent of learning and exploration that is possible through browsing of associations.

In exploratory browsing of Twitter, people follow chains of hyperlinked @username and #hashtag social media associations to discover connected perspectives. However, as we also discussed in the previous section, initially rendered content gets overwritten on clicking these associations. Disorientation and digression grow severe with time, making it difficult for users to follow the topic and understand the conceptual space as a whole. The limits that exist on human cognitive resources only worsen the issue. Therefore, it becomes vital to employ methods and techniques that

ameliorate the process of exploratory browsing.

2.2 Working Memory and External Cognition

Thinking about many things at one time is not possible for humans, due to the limits that exist on our working memory. The number of information chunks that our working memory can hold at a given time is ∼4 [12]. If elements are related in some manner, then these can be held together as a single chunk [53]. This, in effect, makes it possible to work on a greater number of elements.

External cognition is often employed to overcome limits of working memory. Graphical representations facilitate cognitive processing, learning, and problem solving. Scaife and Rogers [50] investigated relationships between graphical representations and internal representations and the computational processes that mediate between the two, which inform design.

Among graphical representations, *information visualization* is the "use of computer-supported, interactive, visual representations of abstract data to amplify cognition" [6]. Information visualization can increase the available memory and processing resources, reduce search for information, enhance detection of patterns, and enable perceptual inference, among several possible cognitive amplifications.

Tufte demonstrates that effective visual representations can be designed, even for most complex of the information sets [57]. Clutter and confusion are failures of design, not attributes of information. The information design principles that Tufte developed include consistency, foreground/background relationships, stratification, limited use of saturated color and heavy weight, and other particulars of layering and separation.

2.3 Dynamic Exploratory Browsing Interfaces

Wilson et al. found that in knowledge discovery tasks, interfaces that visually present connections help users assess how individual elements fit into a domain [62].

Dynamic exploratory browsing interfaces [48] present metadata summaries of multiple web pages and their connections simultaneously. In a study that involved students seeking papers for an Information Retrieval course project, it was learnt that users prefer to discover associations through the proof-of-concept Metadata In-Context Expander interface [30], as compared to using regular browser mechanisms such as tabs or windows. Users also reported that the interface helped mitigate disorientation, overcome digression, and facilitate comparison.

Fluid Links [65] enable users to see additional information from hyperlinked pages in the source context. Glosses, the brief explanations of the connected material, are positioned below the anchor, through interline expansions. These can be explicitly authored or dynamically derived. For the dynamically derived glosses, the most important information, such as title, first line, first paragraph, is displayed–as it is not always possible to describe the exact relationship between the source and destination. Nested glosses and multi-way links are supported.

Web Summaries [15] also allow extraction and presentation of metadata summaries from the hyperlinked web pages. The extraction patterns for these hyperlinked pages initially get defined through interactively clipping and tagging pieces of webpages. The extracted content is presented in a separate Web Summaries window. In the qualitative feedback [16], many users reported that they liked the feature of viewing linked page the most, but also, that they spent too much time switching between the browser window and the summary window.

TweetDeck [58], an alternative Twitter interface, shows multiple feeds in separate

columns. The user customizes the interface to add, remove, and organize @username and #hashtag feeds. TweetDeck enables concurrent watching of multiple feeds. It provides limited flexibility of organization, as each feed occupies the entire vertical space. Relationships through Tweets are not visually presented across feeds.

2.3.1 Evaluation of Exploratory Browsing Systems

Measuring efficacy of exploratory browsing systems is an interesting research challenge. Elsweiler et al. note the need to establish suitable metrics for evaluating casual-leisure information-seeking environments [17]. Kerne et al's approach to evaluate products of information-based ideation—which exploratory browsing process form a part of—uses ideation metrics of curation. White and Roth recommend measures for evaluating exploratory search systems [61]. In section 4, we further discuss the approaches and derive an evaluation methodology.

3. SOCIAL MEDIA USE

In this section, we discuss social psychology of creativity as the basis of social media participation, everyday use of social media in a variety of contexts, and interface and mining techniques that have been developed for social media.

3.1 Social Psychology of Creativity

Amabile [1] discusses how cognitive abilities, personality characteristics, and social factors contribute to different stages of creative process. A componential framework is presented for conceptualizing creativity which includes domain-relevant skills, creativity-relevant skills, and task motivation. The work discusses that extrinsic motivation can have detrimental effects on the process of creativity.

Deci and Ryan's [13] Self-determination theory (SDT) discusses that human motivation is tied to innate psychological needs for competence, autonomy, and relatedness. It states, "Social contexts and individual differences that support satisfaction of basic needs facilitate natural growth processes including intrinsically motivated behavior [. . .]."

Social media witness participation by a large number of people as trending issues or topics directly or indirectly affect them. Opinions are shaped with online interactions of participants, driving the topic in different directions, and greatly impact the process and the outcome [55].

3.2 Everyday Use of Social Media

Social media has proliferated into every nook and corner of our lives. The infographic in Figure 3.1 (source: [38]) shows the wide usage and popularity of websites such as Facebook, Twitter, Pinterest etc. From news to politics to sports to every

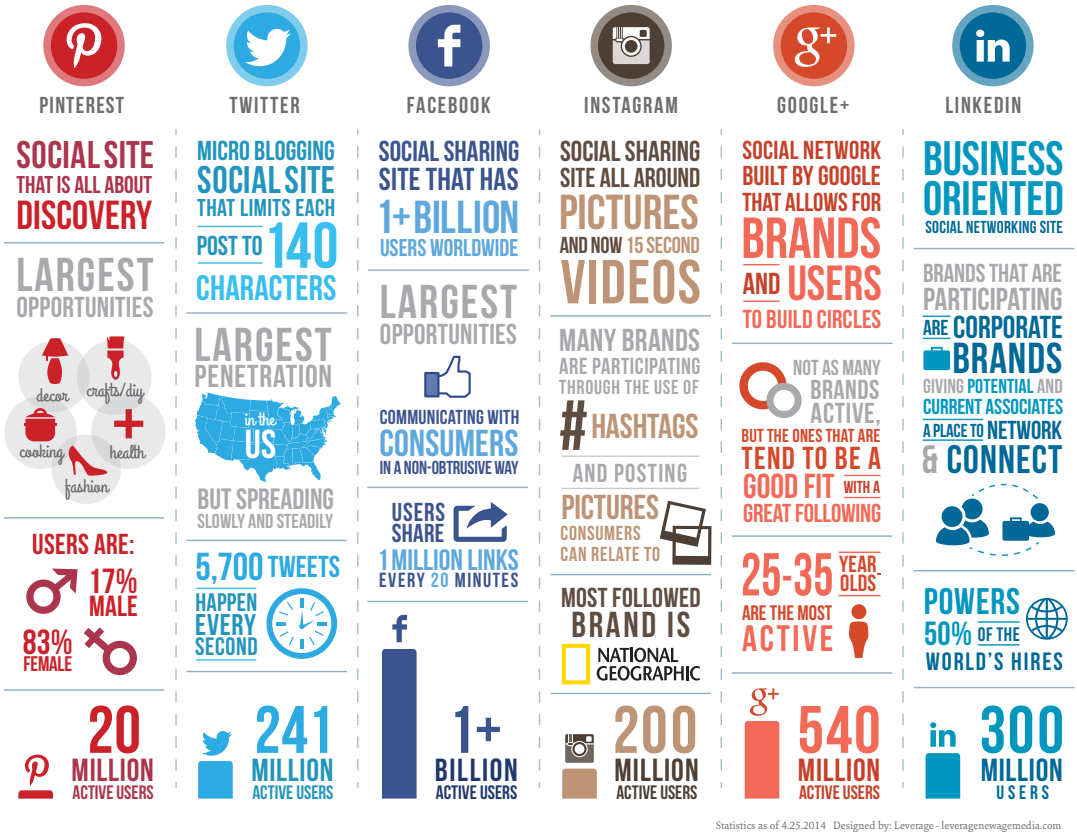

Figure 3.1: The infographic shows engagement of users with different social media platforms. These platforms attract diverse demographics based on the variety of contexts supported with them.
(source: https://leveragenewagemedia.com/blog/social-media-infographic/)

important event, social media has become the de facto platform for participation, communication, and awareness [52, 37]. Deriving from Linder et al's work on *everyday ideation* [39], we call such use of social media as *everyday*. In Lindley et al's [40] categorization of people's usage of web, engagement with social media websites comes under *opportunistic use*, among the 5 categories. It is noted that users are always able to find something engaging to pursue with social media websites. Their use typically involves satisfying some curiosity.

As Figure 3.1 also shows, different platforms have different reach and are suitable

for different purposes, which makes it possible for people to use social media for an array of personal, educational, and business contexts. With social media, people get the ability to express their views to a built-in audience. They are able to share their feelings, or put across a subtle point, on an otherwise serious or delicate issue [41, 7]. Businesses are able to drive gains by utilization of social media [33]. Higher education faculty find value in using social media for purpose of teaching, collaborative learning, and sharing [46].

Exploratory Browsing forms an inherent part of users' interaction, as users seek and encounter related content on the social media websites. It leads them from one link to another, eventually causing them to lose track of their exploration. We observe that `@username` and `#hashtags` links are a lightweight mechanism for providing what Chi et al. call 'distal information scent' [10] in tweets, based on emergent trending topics and social networks. It is important to develop means that support users' creative exploration through browsing of these associations.

3.3 Social Media Interfaces and Mining

With an enormous amount of user-generated content on social media websites, both information interfaces and mining are being seen in a new light.

Recent social media interfaces research investigated how to make it easier for users to understand entities and their relationships. TwitInfo [43], a first of its kind focus+context exploration interface for the social streams, identifies and summarizes events, user sentiments, and frequently occurring terms, tweets, and links. Users were able to construct meaningful summaries of events in a small amount of time. Vizster's [25] node-link layout allows users to see and interact with complex relationships that underlie a large social network. It supported visual search and analysis and stimulated playful exploration of network.

More solutions exist that uncover different aspects of the large information network. For e.g., Link Different [22] deals with a structure awareness problem by relating the theory of social translucence to the social network structure. It lets Twitter users know how many of their followers have already seen a link from someone else. Methods like Facebook Graph Search [18] have been introduced to provide easy mechanisms to lookup information otherwise not readily available.

Sophisticated algorithms have been developed to mine useful information from social media. Starbird and Palen [55] studied how tweet characteristics could have been used to identify participant groups (on the ground vs remote) during the Arab Spring uprising, thus facilitating emergency response. Recent work by Caverlee et al. [8] mines, models, and analyzes the geo-spatial footprints in the social media. In [32], they develop an approach for spatio-temporal meme prediction, building on a global footprint of 755 million geo-tagged Twitter hashtags.

Other characteristics of social media usage have also been studied. Suh et al. found that tweets containing `#hashtags` are more likely to be retweeted [56]. We interpret this to mean `#hashtag` content is significant to users. Participants engage in significant discourse through social media, arranging themselves in various social network patterns, in which attributes of the underlying information, such as events and political positions, play important roles [54].

4. INFORMATION-BASED IDEATION

Information-based ideation (IBI) is a paradigm for investigating open-ended tasks and activities, in which people generate and develop new ideas while browsing, searching, and collecting information [35]. Users perform exploratory browsing and search as part of engagement in IBI.

4.1 Measuring Information-Based Ideation

For evaluating exploratory search systems, White and Roth [61] suggested measuring information novelty, encountering sufficient information, and extent of topic space covered. They did not provide methods for these measures.

Building on Guilford's factors of creativity [24], engineering design researchers measured ideation in solutions to design problems [51]. Extending this, Kerne et al. derived a quantitative methodology for evaluating IBI support tools through a battery of *ideation metrics* [35]. Among these are elemental ideation metrics:

> *Fluency* is the number of ideas. According to Darwinian theories of ideation, the more ideas a person considers, through survival of the fittest, the more likely it is that one idea will survive and grow to achieve creativity.
>
> *Flexibility / Variety* is the number of categories of ideas. It addresses exploration of alternative interpretations. Flexibility measures the span of the solution space explored during ideation. Flexibility in thinking describes the cognitive process of trying out a Variety of different ways of looking at a problem. Variety provides opportunities for more remote associations, and more remote analogies, both of which are likely to lead to creativity.

> *Novelty* is the rareness of an idea. It can be measured with statistical infrequency, which requires an appropriate norm for the space of possible ideas. While potentially difficult to assess globally, Novelty is straightforward to measure in the context of a laboratory experiment by building a master list and inverted index of all ideas generated by all participants. Then, count the number of participants that presented each idea. The lower the count, the higher the Novelty. The Novelty metric is analogous to information retrieval's inverse document frequency (IDF) measure [49].

Kerne et al. applied ideation metrics to measure creativity in the curation products that people author through engagement in IBI activities, by collecting, organizing, and annotating. The present research applies these metrics instead to processes of exploratory browsing.

4.2 Ideation Metrics of Exploratory Browsing

We derive new ideation metrics of exploratory browsing by recontextualizing Kerne et al's metrics of curation [35]. We shift application of the metrics from the products of information-based ideation to its processes of exploratory browsing. Elemental, but not holistic ideation metrics of curation are appropriate for measuring creativity in exploratory browsing, because it consists of a series of actions, without deliberate assemblage of a whole. We rework prior methods for computing elemental ideation metrics—Fluency, Flexibility, Novelty—to derive new ideation metrics for creative engagement in exploratory browsing of Twitter social media.

Define the the set of types $M = \{username, hashtag\}$ of Twitter social media associations. Using type m from this set, $o_{m,i}$ refers to each feed a user browses, of each type.

Then, we express the set of all social media `@username` and `#hashtag` associations

traversed in a user's exploratory browsing process, which corresponds to Kerne et al.'s set of all found digital media objects that a user curated [35]:

$$c = \{o_{m,0}, o_{m,1}, ...\}; \forall m \in M \tag{4.1}$$

$Fluency_m(c)$ will be the total number of feeds the user browsed, of type m

$$Fluency_m(c) = ||o_{m,i} \in c||; m \in M \tag{4.2}$$

Note that recurrent browsing of any particular feed is aggregated by this Fluency measure.

Flexibility will be the total number of distinct `@usernames`, `#hashtags` browsed. This means that any recurrences of browsing a particular feed do not contribute to the Flexibility measure. To compute Flexibility, form the set of all distinct feeds of type m, $DistinctFeeds_m(c)$, from collection c of all the `@usernames`, `#hashtags` a user browsed.

$$DistinctFeeds_m(c) = \bigcup_{o \in c} o_{m,i}; m \in M \tag{4.3}$$

$$Flexibility_m(c) = ||DistinctFeeds_m(c)|| \tag{4.4}$$

We also categorized user feeds into types. We compute $Flexibility_{user\_type}$, drawing on Kerne et al's site type granularity. For this purpose, we extend the categories of Twitter users established by Wu et al.: *celebrities, media, organizations, and blogs* [64], adding: *fashion, professional*. We separate *organizations* into categories: *community, activist*. We performed manual classification to assign types to the 179 `@username` feeds browsed by study participants. The resulting mapping is defined by

a *UserType* operator.

$$Flexibility_{user\_type}(c) = || \bigcup_{o \in c, m = username} UserType(o_{m,i}) || \tag{4.5}$$

To compute Novelty, build an inverted index that shows for each feed f, of type m, the number of users who browsed it.

$$Occurrences_m(f, C) = \{c | c \in C \wedge f \in DistinctFeeds_m(c)\} \tag{4.6}$$

$$Novelty_m(f, C) = \frac{1}{||Occurences_m(f, C)||} \tag{4.7}$$

Then, Novelty of feeds browsed by each user, of type m, can be computed as aggregated mean of Novelty scores for `@username` or `#hashtags` feeds that the user browsed.

$$Novelty_m(c, C) = \frac{\sum_{f \in DistinctFeeds_m(c)} Novelty_m(f, C)}{||DistinctFeeds_m(c)||} \tag{4.8}$$

5. EXPLORATORY BROWSING SYSTEM DESIGN

In this section, we present TweetBubble's system design. A dynamic exploratory browsing interface for Twitter social media associations requires multiple feeds to be rendered within a single web page. To facilitate retrieval and presentation from connected `@username` and `#hashtag` feeds, we extend the *Meta-Metadata* language for metadata types, a component of the open-source *BigSemantics* framework [28]. Meta-Metadata integrally describes data models, extraction rules, and presentation semantics, for semantic representation of linked web pages [48], such as Twitter feeds. We used the types to build a look and feel consistent with Twitter. *Presentation semantics* enable detailed customization of the visual interface for particular metadata types.

We developed TweetBubble as a *Chrome extension* to investigate how exploratory browsing interfaces can help users understand, work with, and think about complex information. The TweetBubble runtime is constituted by a content script that transforms the Twitter interface. More details of using Chrome extension technology can be found in Appendix A.

We present an interface design, followed by a mini-scenario. Next, we discuss the metadata type system and its role in dynamic exploratory browsing interfaces, including the Metadata In-Context Expander. This is followed by details of new types and presentation semantics for Twitter.

5.1 Interface Design

TweetBubble interjects affordances for in-context expansion of `@usernames` and `#hashtags` into Twitter (Figure 1.2b). Pressing a hyperlink causes download of the corresponding feed, extraction of metadata, and rendering the extracted metadata

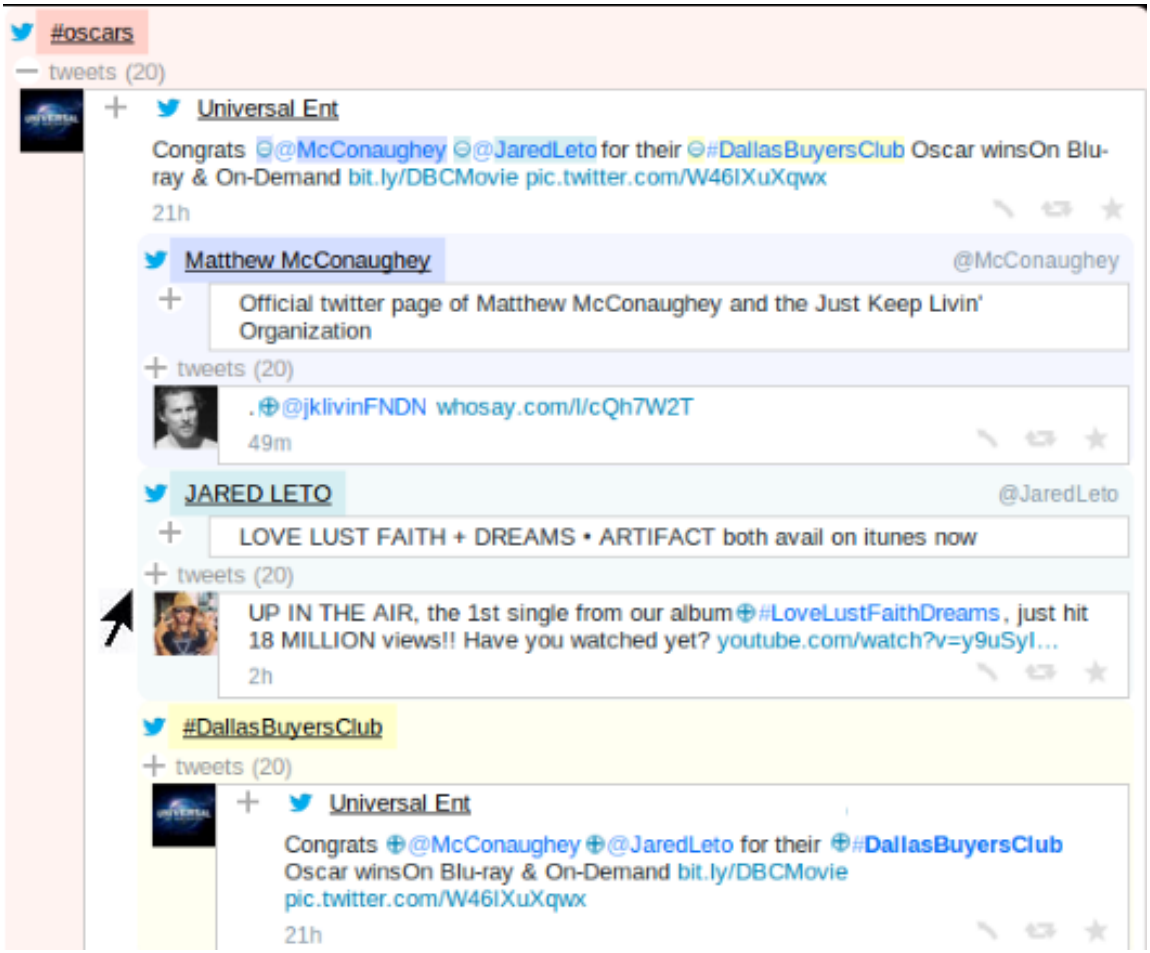

Figure 5.1: Recursive in-context expansion of `@usernames`, `#hashtags` facilitates understanding of associations and seeing the big picture.

amidst the current webpage (Figure 1.3). `@usernames` and `#hashtags` appearing in the rendered metadata are again expandable, creating recursive branches of metadata expansion (Figure 5.1). This provides the ability to simultaneously view multiple feeds, form associations, and make comparisons. A palette of related colors is used to connect a social media association and its expansion as metadata.

Repetitive information becomes a problem as Twitter feeds and tweets are recursively expanded. More feeds are presented at once than in default Twitter. Tweet-

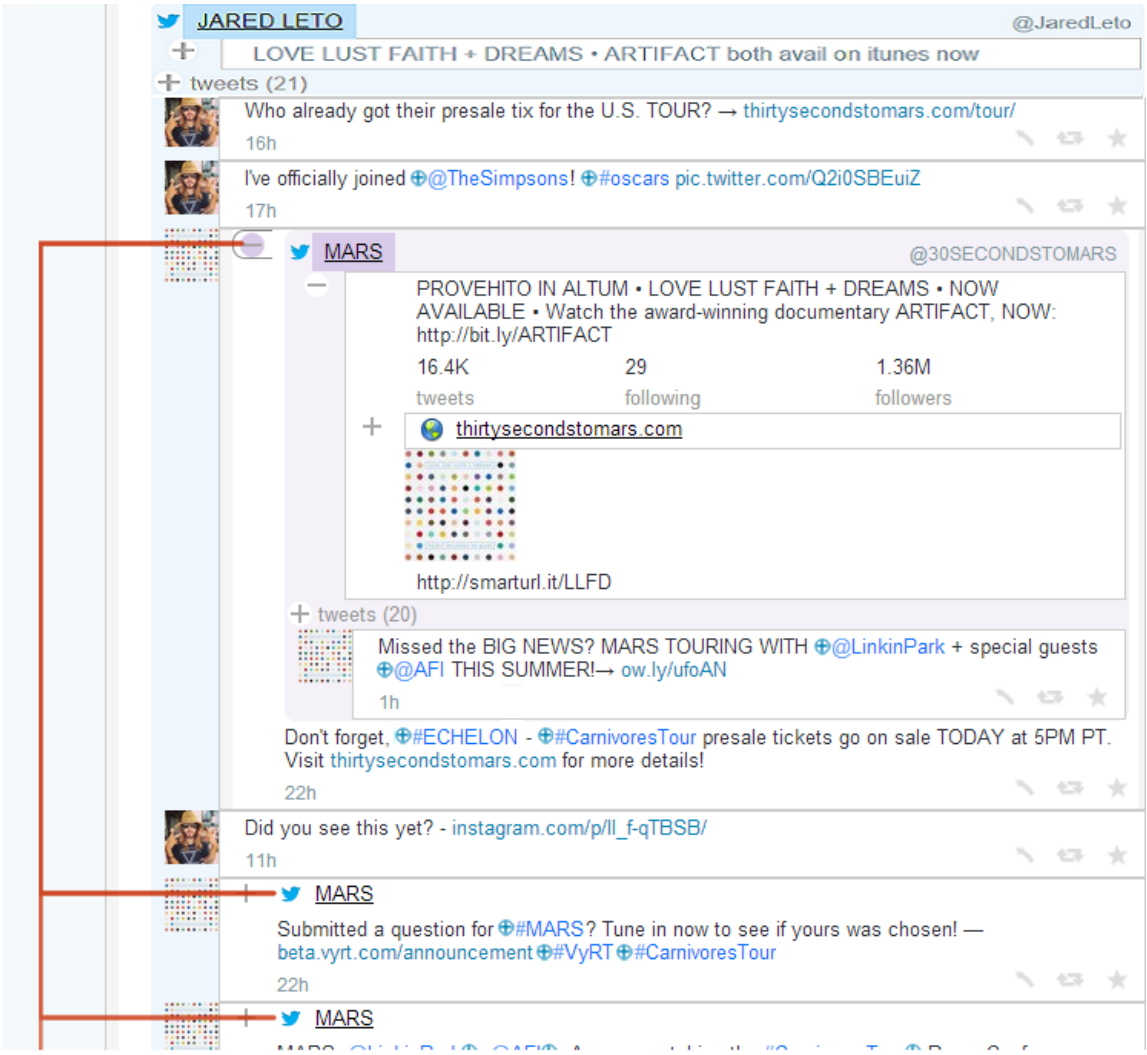

Figure 5.2: Presentation semantics suppress redundant tweeter name and handle information within tweets rendered for an expanded username social media association (See the `tweet` metadata type in Figure 5.3, and Tables 5.1 and 5.2). For retweets, connection lines highlight a recurring expandable username association, to aid users in not repeatedly expanding the same one.

Bubble uses presentation semantics to suppress redundant tweeter name and handle metadata. Further, to help orient the user, when s/he activates an `@username` field in a retweet, the interface draws red lines to visualize recurrences in other tweets (Figure 5.2).

TweetBubble provides interaction consistent with Twitter. In Twitter, after expanding a tweet to see prior social discourse actions—*replies, retweets*, and *favorites*—pressing the background collapses. TweetBubble extends the mapping of press background to collapse all expanded information, streamlining navigation. Consistent interaction also includes affording social discourse actions with each tweet. We form tweet action URLs, using Twitter's unique ∼18-digit tweet identifier, extracted via metadata types.

5.2 Mini-Scenario

We present a mini-scenario to demonstrate how user experience is facilitated through the type system's integration of presentation semantics. Janet clicks on the expansion affordance of `@JaredLeto`'s tweet collection (Figure 5.1). This collection is rendered using presentation semantics in the `twitter_microblog` type (Figure 5.3). The `tweets` collection uses `child_use_values_as_label`, `child_show_expanded_always`, and `label_at` directives (Tables 5.1 and 5.2). It also applies the `is_child_metadata` attribute to the `twitter_microblog` reference within each tweet.

`child_use_values_as_label` is a variant of the `use_values_as_label` semantic; here, it specifies re-presenting members of a collection. Applying this directive to the `tweets` collection renders the label of each tweet composite using the contained `twitter_microblog`'s `photo` field (See result, Figure 5.2). In `@JaredLeto`'s feed, we see the `is_child_metadata` directive instruct the renderer to compare each `tweet`'s `twitter_microblog` reference to the feed it is a part of, and when they are the same, to suppress redundant title and handle information from those tweets. Elimination of redundant information reduces cognitive load [63].

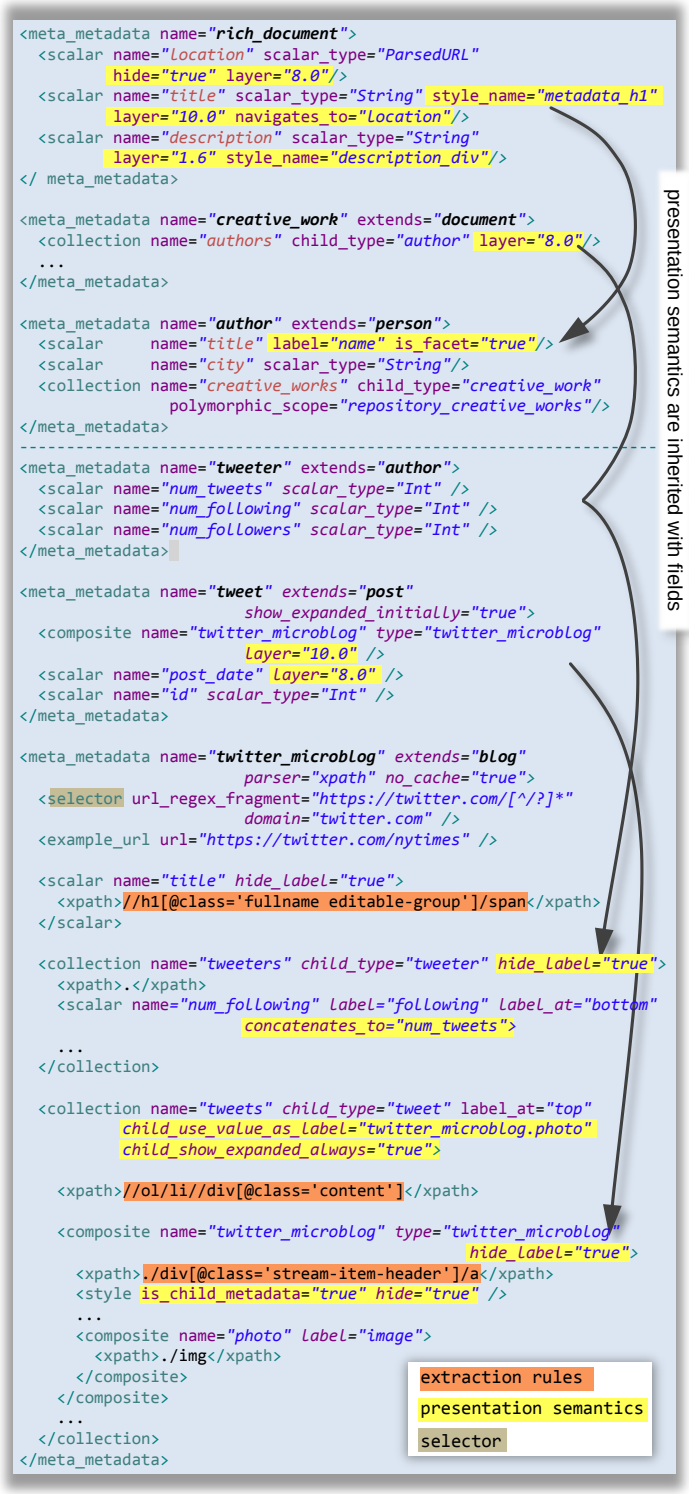

Figure 5.3: Metadata types involved in Twitter. Inheritance in Meta-Metadata promotes reuse of data models and presentation semantics (See Tables 5.1 and 5.2) across different information sources.

Semantics	Description
`layer`	order of the metadata field; higher value means higher in order
`hide / always_show`	visibility of the metadata field; `always_show` is used to override the `hide` attribute
`navigates_to`	hyperlink metadata field to the specified destination
`style_name`	css class name to be applied to the metadata field
`shadows`	show this metadata field instead of the specified field
`is_facet`	metadata field can be used as a facet

Table 5.1: Presentation semantics previously defined in the Meta-Metadata language for metadata types.

Semantics	Description
label_at	positioning of label w.r.t. value of metadata field
concatenates_to	concatenate the metadata field to another
use_value_as_label	use value of another metadata field as label for this field
hide_label	visibility of the label of metadata field
show_expanded_initially	show the composite / collection field expanded initially
show_expanded_always	show composite / collection field expanded without affordance for collapse
style	attributes is_child_metadata, is_same_metadata, is_only_element, is_top_level and presentation semantics to be applied e.g. hide

Table 5.2: New presentation semantics in Meta-Metadata enable seamlessly integration of exploratory browsing of @usernames and #hashtags into the standard Twitter interface.

5.3 Type System for Wrappers

The Meta-Metadata type system for metadata supports inheritance and polymorphism. Unlike in the triples of The Semantic Web [2] data models within metadata types represent underlying information structures as abstract data types, each comprised of typed fields. Together with extraction rules and presentation semantics, data models are used to dynamically derive and present metadata summaries of web pages. Reuse of type elements is facilitated by inheritance of data model fields, and also of attribute values, used to specify extraction rules and presentation semantics.

The raison d'être of presentation semantics is to enable structuring how metadata will look to users to improve legibility, and so comprehension and usability. Qu et al. developed an initial set of simple, yet effective presentation semantics in the Meta-Metadata language. They enabling hiding, ordering, and hyperlinking of fields (See Table 5.1) [48]), which are summarized in Table 5.1.

The Meta-Metadata language also includes specification of selectors as part of types. A *selector* specifies a URL pattern to be used at runtime to match any web page with the appropriate type.

An important base type provided by Meta-Metadata is `rich_document`, which models a general webpage. It contains `title`, `location` (URL), and `description` fields. Common code in the runtime operates polymorphically on `rich_document` subtypes, extracting and rendering metadata from heterogeneous sources. The Meta-Metadata wrapper repository provides types for diverse information sources, including products, movies, scholarly articles, and news.

5.4 Metadata In-Context Expander

Metadata In-Context Expander (MICE) [30] is a metadata renderer that visualizes extracted metadata fields as a tree. It provides a dynamic exploratory browsing

interface with metadata summaries. MICE operates on the `rich_document` subtype in order to render the extracted metadata with the aid of presentation semantics. It joins together the metadata and presentation semantics to create an optimal presentation of the source. MICE is similar to the Tabulator, which operates on linked RDF data [4]. The present research uses MICE to expand social associations within Twitter.

5.5 Metadata Types for Twitter

We extended prior types—`rich_document, author, creative_work, blog,` and `post`—to derive new metadata types for Twitter (Figure 5.3). We developed types `twitter_microblog` and `twitter_search_results` for `@username` and `#hashtag` feeds. We developed supporting metadata types `tweet` and `tweeter` by extending `post` and `author`, respectively. We developed presentation semantics and extraction rules to match the new data models.

5.5.1 Presentation Semantics

Figure 5.4 shows how TweetBubble is able to mimic `@username` feeds using the presentation semantics from Table 5.1 and 5.2, in the derived `twitter_microblog` type (Figure 5.3). Its underlying base type, `rich_document`, uses the `navigates_to` directive, so that the title affords pressing to browse the original web page in a new tab. The `hide` directive eases the user's visual load by specifying to not directly render the URL field.

Our goal for TweetBuble was to uniformly integrate into Twitter by mimicking look and feel. This required us to provide more fine-grained control in Meta-Metadata. We developed new presentation semantics: positioning, expanding, collapsing, concatenating, and changing labels of the fields (Table 5.2).

Figure 5.4 presents a further example. In the `tweeter` composite contained in the

Figure 5.4: Presentation semantics enable highly detailed metadata expansion. The rendering of Ellen's `@username` feed is customized with semantics `concatenates_to`, `navigates_to`, `hide_label`, and `label_at` (See Tables 5.1 and 5.2). The `twitter_microblog` metadata type plays a central role (See Figure 5.3).

`twitter_microblog` type, we position labels below values, Twitter style, using the `label_at` presentation semantic in conjunction with the `num_following`, `num_followers`, and `num_tweets` fields. Meanwhile, `concatenates_to` renders these fields in the same horizontal span.

5.5.2 Extraction Rules

Extraction rules in metadata types [48] are specified by using XPath [60] to select one or more nodes in a web page's Document Object Model (DOM). An XPath is like a filesystem path into the tree of the DOM with optional predicates and wildcards. In Meta-Metadata, XPaths are specified in conjunction with the declarations of particular fields in the data model for a type. Figure 5.3 shows examples, such as in the declaration for the `title` field in `twitter_microblog`.

5.6 Contextualized Extraction

The TweetBubble Chrome extension extends the experiences of users who are logged in, in the contexts of their Twitter environments. We developed in-browser Javascript metadata extraction, which enables sidestepping authentication issues with social media websites. TweetBubble uses the browser's JavaScript `XMLHttpRequest` to download web pages for their contextualized social media associations, in the background. `XMLHttpRequest` enables obtaining the DOM object of the web page, to which we are directly able to apply the XPath extraction rules specified in appropriate metadata types. Details of extraction can be found in Appendix B.

6. EVALUATION: ACADEMY AWARDS

We conducted a study at the time of the 2014 Academy Awards. More than 5 million people sent about 19.1 million Tweets about the Oscars and more than 37 million people viewed those Tweets [20]. The high level of Twitter participation during the event was conducive to recruiting engaged study participants via Amazon Mechanical Turk.

6.1 Study Design

We conducted a between-subjects study with the Twitter interface as the independent variable. The TweetBubble extension condition offered affordances to expand `@username` and `#hashtag` social media associations. The control condition was the default Twitter interface, in which clicking a social association link overwrites initial content. Across conditions, the apparatus logged how users explored encountered information.

The study asked users to browse one of two trending Twitter topics, `#ERedcarpet`, or `#oscars2014`, for 15 minutes. The Red Carpet event attracts a fashion-oriented audience. Celebrities walk up to the Oscars, wearing outfits from prominent fashion designers. We first published 12 Mechanical Turk HITs with `#ERedCarpet`, around 3pm PST, when the TV broadcast of Red Carpet started. This was followed by 38 more HITs for `#ERedCarpet`, a couple of hours later. As the main Academy Awards event began, around 7:30pm PST, we published 50 HITs with the topic `#oscars2014`. Overall, there were 59 male and 41 female participants in the 3 sets. Participants were given pre- and post-questionnaires.

The HITs specified that participants should be regular Twitter users (having used Twitter at least once in the last week). We didn't strictly validate this, instead we

relied on participants' self-reports, in conjunction with sanity and attention checks [45, 36]. The sanity check questions we asked were: (1) Who of the following is not nominated for Best Actress or Best Supporting Actress? Response choices were (a) Meryl Streep (b) Jennifer Lawrence (c) Miley Cyrus (d) Sandra Bullock; and (2) Which of the following is not known as a dress designer? Choices were (a) Vera Wang (b) Estee Lauder (c) Alexander McQueen (d) Christian Dior. The attention check question was like that of Marshall and Shipman [45]. The complete pre- and post-questionnaires can be found in Appendix C and Appendix D respectively.

Based on sanity and attention check responses, we discarded 18 and 15 responses respectively from the TweetBubble extension and control conditions. We also eliminated 13 participants post hoc because of logging problems. This leaves 25 participants for the TweetBubble condition and 29 participants for the control. We developed and applied a quantitative method for measuring the creativity of exploratory browsing engagement with Twitter. We collected and analyzed qualitative data about users' social media experiences.

6.2 Results: Ideation Metrics of Exploratory Browsing

We used this method to compute ideation metrics of exploratory browsing. Results across conditions were compared with t-tests, assuming unequal variances. We found that all ideation metrics of exploratory browsing were significantly higher for `@username` social media associations in the TweetBubble extension condition (Table 6.1). For `#hashtag` associations, Fluency and Flexibility significantly increased in the TweetBubble condition (Table 6.2). Figure 6.1 graphs the results.

6.2.1 Continuing Use

Users continue using TweetBubble after the study. In an evaluation 2 months after the study and releasing the extension, it is found that 11 study users continue to

Metric	TweetBubble μ	SE	Control μ	SE	p <
Fluency	8.76	2.007	1.137	0.431	.0000186
Flexibility	6.84	1.614	1.103	0.421	.0000746
Flexibility (User Type)	3.6	0.881	0.759	0.267	.00000352
Novelty	0.599	0.118	0.288	0.096	.001744

Table 6.1: Ideation metrics of exploratory browsing for `@username` associations. All were significantly higher in the TweetBubble condition.

Metric	TweetBubble μ	SE	Control μ	SE	p <
Fluency	5.6	1.473	1.517	0.755	.00397
Flexibility	3.32	0.888	1.068	0.509	.008859
Novelty	0.269	0.065	0.16	0.064	.084316

Table 6.2: Ideation metrics of exploratory browsing for `#hashtag` associations. All except Novelty were significantly higher in the TweetBubble condition.

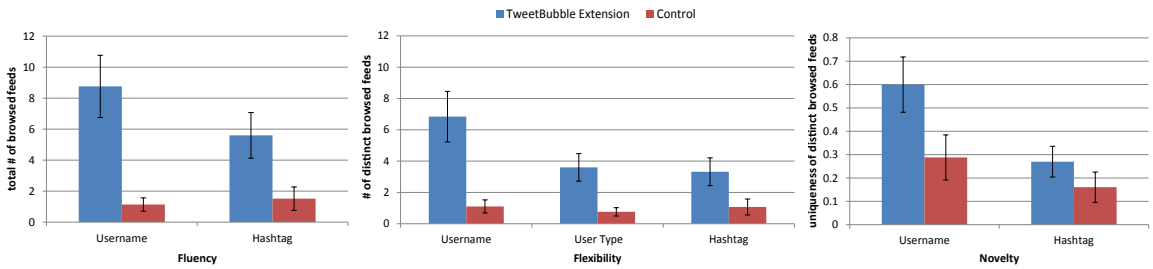

Figure 6.1: Mean Fluency, Flexibility, Novelty of `@username` and `#hashtag` feeds that users browsed. The TweetBubble Extension interface condition (in blue) resulted in significantly more creative exploratory browsing than the Control condition (in red) for 6 out of 7 ideation metrics. Error bars show standard errors of the mean.

use TweetBubble. 12 new users who installed the extension in month of March and 15 new users installed the extension in month of April also continue using TweetBubble.

6.3 Qualitative Results

TweetBubble condition participants were asked, via questionnaire, how clicking on related content, seeing it on the same page, and thus simultaneously browsing multiple feeds affected their experiences. Drawing on grounded theory methods [9], we performed a qualitative analysis of resulting data. We first performed open coding of a set of participant responses. This resulted in 25 codes. We followed with focused coding of remaining responses. We aggegated codes into 5 categories: 'Reduces Disorientation and Digression', 'Helps Explore Related Content', 'Connecting Perspectives to Gain Broader Understanding', 'Fluid and Effective', 'Too Much Information?'.

6.3.1 Reduces Disorientation and Digression

Participants reported that TweetBubble helped them focus. They liked that the original context of reading is maintained when browsing related content. This helped them stay oriented and on topic. They reported having more information at their fingertips.

P71: *"...It was great to be able to browse through all the #hashtags and @usernames, as I love to do, and yet easily get back to where I was originally."*

P99: *"Being able to see everything on one page kept me much more focused."*

P93: *"I learned more about what I wanted to know about without wasting the time clicking to different pages."*

P3: *"It made it more cohesive and allowed me to keep a train of thought instead of getting so hopelessly sidetracked by having new windows constantly."*

6.3.2 Helps Explore Related Content

Without going back and forth between windows, participants found themselves more engaged. They reported easily and efficiently viewing and discovering more content, exploring more Twitter users' feeds, following online conversations better, and building more in-depth understanding of topics. TweetBubble introduced them to more diversified content, often missed when using the default interface.

P37: *"...the more you have to go back and forth, the less interested I am in staying and learning/reading about topics, because it is too time consuming. So having everything all together made me interact a lot more than I would've before."*

P99: *"It helped me delve into the subjects more. It also kept me more engaged in browsing the Twitter feed."*

P91: *"It allowed me to see more of the conversation and subject."*

P68: *"It made my experience so much quicker, with so much more content. I was amazed by how easy it was to see tweets between two of my favorite celebrities."*

6.3.3 Connecting Perspectives to Gain Broader Understanding

Participants reported new ability to see relationships across topics and perspectives. They more readily saw diverse viewpoints. Exposure to contrasting viewpoints enabled them to connect perspectives, and so gain broader understanding.

P47: *"This allows one to understand how the system connects related topics and to see how and why they are linked together"*

P21: *"You get to see many different points of view, professional and civilian. You could get quick quips or detailed facts. It allows for a much broader reading experience."*

P33: *"It gave me a broader view of the topic all in one area, instead of being segmented across several pages. Made it a lot easier to understand as a whole."*

6.3.4 Fluid and Effective

Users found the experience of the TweetBubble interface to enable more fluid and effective navigation of Twitter.

P53: *"...The way the information is presented in such a fluid way and color coded is ... great..."*

P79: *"It was a lot smoother and more efficient. I'm definitely pleased. I'm leaving the app installed."*

P45: *"I experienced user-friendly navigation options."*

6.3.5 Too Much Information?

A few participants felt overwheled, as the number of expansions increased.

P59: *"The extension made the @usernames and #hashtags open up a tree menu which ended up being confusing and cluttered. I understand how it can link topics and users together, but it is too messy."*

P85: *"It was straightforward but overwhelming at this time."*

7. CONCLUSIONS

We begin with a discussion of the study results and derive implications for design. We discuss intellectual merit and broader impact of the work. We also briefly discuss future work.

7.1 Design Implications

We frame investigation of exploratory browsing by delineating the everyday contextual role of social media in society. We derive implications for design based on the study results, browser extensions as design laboratories, and the role of the polymorphic metadata type system.

According to Marchionini, browsing is an approach to information seeking that is informal, opportunistic, and iterative [42]. One respondent in Lindley et al.'s study addressed this for social media: 'there's always something' engaging to pursue [40]. Exploratory browsing interfaces are designed to reduce cognitive load from disorientation and digression in opportunistic information seeking. The preponderance of `@username` and `#hashtag` references in meaningful tweets (such as those retweeted), highlights their role. Our qualitative and quantitative study data validate the present approach to exploratory browsing interface design and development.

An exploratory browsing interface should provide means to explore related content, while maintaining originating context. It should depict browsed chains of association. In the study, users began from the same `#hashtag` search page, `#ERedCarpet` or `#oscars2014`. With the TweetBubble extension, they discovered significantly more diverse and novel content, as compared to the participants using the default Twitter interface. This shows that the TweetBubble's in-context expansion of social media associations promotes users' exploratory browsing. It provokes discovering associa-

tions, stimulates users to learn about diverse perspectives on a topic, and so promotes flexibility in thinking.

Popular social media, particularly Twitter and Facebook, have attained such a pervasive role in the conduct of society, that, drawing on the work of Linder et al. [39], we call them *everyday*. An indicator is how the medium of television—e.g., The Tonight Show, news, and sports—positions Twitter in leading roles, framing discourse. Similarly, personal lives and business relationships are extensively built and expressed through Twitter and Facebook. Everyday use refers to people's engagement with these platforms for a panoply of meaningful events in their lives. Thus, how people engage in exploratory browsing of social media is very significant, on a societal level.

Exploratory browsing becomes an inherent part of people's everyday social media lives, as they seek and encounter trends, events, and experiences, in personal, educational, and business contexts. In doing so, they draw and generate new value for themselves and peers.

Exploratory browsing of `@username` and `#hashtag` social media associations is a form of mini-c creativity, because people encounter and learn about new perspectives, themselves. Little-c creativity arises through social discourse tweet actions of following, favoriting, and retweeting. People combine and connect tweets, transforming their contexts, and impacting social networks. We, therefore, took a creative cognition approach and reworked ideation metrics of curation [35], to compare interfaces designed for exploratory browsing of social media. The diversity and novelty of personal and social forms of engagement constitute measures of the extent of their creativity.

We investigated a particular everyday social media context, the 2014 Academy Awards, in order to study how interfaces support people in creative exploratory

browsing. To engage typical users with TweetBubble dynamic exploratory browsing interface, we directly embedded it into Twitter. Users of the TweetBubble interface performed significantly higher on 6 out of 7 ideation metrics of exploratory browsing (the 7th was close). Qualitative data mirrors the quantitative results. This grids the claims of validity both for the exploratory browsing interface, and for the method of using ideation metrics of exploratory browsing as the basis for interfaces. The metrics match the user experiences.

How to evaluate exploratory browsing interfaces is an interesting research challenge. For evaluation of systems for casual information seeking, Elsweiler et al. demand new metrics [17]. We found that ideation metrics [35] provide a means to measure efficacy of an interface in stimulating creative exploration. To measure creative exploration in browsing social media, we recontextualized the prior ideation metrics, in terms of the `@username` and `#hashtag` associations that Twitter users explore. In general, the methodology can be adopted for interfaces designed to promote exploratory browsing and exploratory search, to compare interface conditions based on diversity and novelty in experiences of creative exploration.

Browser extensions serve as *interface design laboratories*. In conjunction with recruiting micro-task workers, validated by reliability measures [45, 36], the browser extension approach facilitates elicitation of feedback on interactive system design, utility, and performance of an interactive system's design. Both quantitative and qualitative data can fruitfully gathered in browser extension ideation laboratories. Users become involved in the design of new technologies, while researchers gain valuable insights from their feedback. Browser extensions have the capacity to serve as technology probes [27], when their use is open-ended and extended. Researchers can study new interaction techniques via browser extensions to enable data collection over extended time periods.

The polymorphic metadata type system of the Meta-Metadata language provides a strong foundation for dynamic exploratory browsing systems, by integrating data models, extraction rules, and presentation semantics [48]. The type system enables creation of dynamic exploratory browsing interfaces through polymorphic extraction and presentation of information from heterogeneous sources. Presentation semantics enable customizing exploratory browsing interfaces to match the look and feel of diverse websites. The present research extended these semantics to mimic Twitter social media. Implementation of the polymorphic metadata type system in a browser extension enables direct capture of the user's social media context, avoiding the need to address authentication. Further, re-use of data models and presentation semantics facilitates construction of interfaces for diverse social media. This technique is extensible to other social media, such as Facebook, Reddit, Pinterest, and Google+.

7.2 Summation

TweetBubble stimulated exploratory browsing of the Twitter social media associations, during the Academy Awards study. Ideation metrics of exploratory browsing, which we derived from existing research, enable quantitative comparison of the interfaces, in how they support users' creative exploration. We found that the TweetBubble dynamic exploratory browsing interface users performed significantly higher on 6 out of 7 metrics of Fluency, Flexibility, and Novelty. Users reported that with the TweetBubble interface, the gaining of perspectives, introduction to new people and topics, and following of conversations becomes more easy, engaging and efficient.

What is the significance of these findings? Social media, particularly Twitter, have attained an international transformative impact, playing a catalytic role in social movements, such as Arab Spring, [55] and politics, such as U.S. presidential debates [52]. They transcend individuals, communities, and organizations and re-

define participation, communication, and awareness. Social media content is vast and fast moving. With increased exploratory browsing of Twitter, not only will users make more connections across `@usernames` and `#hashtags`, tweeters will also get better reach for their social media posts. Both the producers and consumers of information gain as diverse and novel perspectives are shared, growing mutual understanding.

Using Boden's notion of exploratory creativity, and Kaufman et al.'s mini-c creativity, we identified exploratory browsing as a form of creative experience, and so of information-based ideation. Creativity is vital to personal well-being [5, 34], as well as innovation that is critical to commercial and national success [47]. Thus, techniques for building interfaces that promote creative engagement in exploration of social media, and methods for evaluating these interfaces, have great potential to significantly impact how social media transforms the world. We see exploratory browsing interfaces playing a key role in human engagement in information-based ideation activities, which, in turn, are the heart of humanity's potential to transform the raw material of our digital society and super-abundant information resources into connectedness, well-being, and success.

7.3 Future Work

Users found the experience with the interface fluid and engaging. New presentation semantics were introduced that provided more fine-grained control of the source presentation. However, users also reported that the representation grows in clutter and becomes messy after a certain point. Further research should investigate other visualization and presentation techniques, providing means to present a lot of information at the same time, while managing cognitive load [59]. It should investigate new techniques to promote flexibility and novelty of exploration and engagement,

which is key to nearly all the facets of creativity [14].

REFERENCES

[1] Amabile, T. M. The social psychology of creativity: A componential conceptualization. *Journal of personality and social psychology 45*, 2 (1983), 357.

[2] Antoniou, G., and van Harmelen, F. *A Semantic Web Primer*. The MIT Press, 2004.

[3] Bernando, A. Flickr photo sharing - ovelhas da raa bordaleira de entre douro e minho (obtained permission to use). `https://www.flickr.com/photos/noussnouss/6096177001/in/photostream/`. Last accessed: 10/19/2014.

[4] Berners-Lee, T., Hollenbach, J., Lu, K., Presbrey, J., and schraefel, m. Tabulator redux: Browsing and writing linked data. Tech. rep., 2008.

[5] Boden, M. A. *The creative mind: Myths and mechanisms*. Psychology Press, 2004.

[6] Card, S. K., Mackinlay, J. D., and Schneiderman, B. *Readings in information visualization: using vision to think*. Morgan Kaufmann, 1999.

[7] Carr, D. View of #Ferguson Thrust Michael Brown Shooting to National Attention - NYTimes.com. `http://nyti.ms/1vsV1WT`. Last accessed: 09/24/2014.

[8] Caverlee, J., Cheng, Z., Sui, D. Z., and Kamath, K. Y. Towards geosocial intelligence: Mining, analyzing, and leveraging geospatial footprints in social media. *IEEE Data Eng. Bull. 36*, 3 (2013), 33–41.

[9] Charmaz, K. *Constructing grounded theory: A practical guide through qualitative analysis*. Pine Forge Press, 2006.

[10] Chi, E. H., Pirolli, P., Chen, K., and Pitkow, J. Using information scent to model user information needs and actions and the web. In *Proc CHI* (2001), pp. 490–497.

[11] Conklin, J. Hypertext: An introduction and SurvevJ. *IEEE Computer 20*, 9 (1987), 17–41.

[12] Cowan, N. The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and brain sciences 24*, 1 (2001), 87–114.

[13] Deci, E. L., and Ryan, R. M. The" what" and" why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological inquiry 11*, 4 (2000), 227–268.

[14] DeYoung, C. G. Higher-order factors of the big five in a multi-informant sample. *Journal of personality and social psychology 91*, 6 (2006), 1138.

[15] Dontcheva, M., Drucker, S. M., Wade, G., Salesin, D., and Cohen, M. F. Summarizing personal web browsing sessions. In *Proc UIST* (2006).

[16] Dontcheva, M., Lin, S., Drucker, S. M., Salesin, D., and Cohen, M. F. Experiences with content extraction from the web. In *Proc UIST* (2008).

[17] Elsweiler, D., Wilson, M. L., and Lunn, B. K. Understanding casual-leisure information behaviour. *Library and Information Science 1* (2011), 211–241.

[18] Facebook. Introducing Graph Search. `https://www.facebook.com/about/graphsearch`. Last accessed: 04/17/2014.

[19] Fallon, J., and Timberlake, J. "#hashtag" with Jimmy Fallon & Justin Timberlake. https://www.youtube.com/watch?v=57dzaMaouXA. Last accessed: 07/30/2014.

[20] Fleischman, M. The reach and impact of oscars 2014 tweets. https://blog.twitter.com/2014/the-reach-and-impact-of-oscars-2014-tweets. Last accessed: 05/19/2014.

[21] Foss, C. L. Detecting lost users: Empirical studies on browsing hypertext. Tech. rep., 1989.

[22] Gilbert, E. Designing social translucence over social networks. In *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems* (2012), ACM, pp. 2731–2740.

[23] Google. What are Extensions? - Google Chrome. https://developer.chrome.com/extensions. Last accessed: 10/08/2014.

[24] Guilford, J. Creativity. *American Psychologist 5* (1950), 444–454.

[25] Heer, J., and Boyd, D. Vizster: Visualizing online social networks. In *Information Visualization, 2005. INFOVIS 2005. IEEE Symposium on* (2005), pp. 32–39.

[26] Honey, C., and Herring, S. C. Beyond microblogging: Conversation and collaboration via twitter. In *Proc HICSS* (2009), IEEE, pp. 1–10.

[27] Hutchinson, H., Mackay, W., Westerlund, B., Bederson, B. B., Druin, A., Plaisant, C., Beaudouin-Lafon, M., Conversy, S., Evans, H., and Hansen, H. Technology probes: inspiring design for and with families. In *Proc CHI* (2003), pp. 17–24.

[28] Interface Ecology Lab. BigSemantics Framework. https://github.com/ecologylab/BigSemantics/wiki. Last accessed: 04/15/2014.

[29] Interface Ecology Lab. BigSemanticsService. https://github.com/ecologylab/BigSemanticsService/wiki. Last accessed: 04/15/2014.

[30] Interface Ecology Lab. Metadata In-Context Expander. http://ecologylab.net/mice. Last accessed: 04/17/2014.

[31] Interface Ecology Lab. TweetBubble project on GitHub - BigSemanticsJavaScript repository. https://github.com/ecologylab/BigSemanticsJavaScript/tree/master/TweetBubble/Plugin/chrome. Last accessed: 10/08/2014.

[32] Kamath, K. Y., and Caverlee, J. Spatio-temporal meme prediction: learning what hashtags will be popular where. In *Proceedings of the 22nd ACM international conference on Conference on information & knowledge management* (2013), pp. 1341–1350.

[33] Kaplan, A. M., and Haenlein, M. Users of the world, unite! the challenges and opportunities of social media. *Business horizons 53*, 1 (2010), 59–68.

[34] Kaufman, J. C., and Beghetto, R. A. Beyond big and little: The four c model of creativity. *Review of General Psychology 13*, 1 (2009), 1.

[35] Kerne, A., Webb, A. M., Smith, S. M., Linder, R., Lupfer, N., Qu, Y., Moeller, J., and Damaraju, S. Using metrics of curation to evaluate information-based ideation. *ACM ToCHI 21*, 3 (2014), 48.

[36] Kittur, A., Chi, E. H., and Suh, B. Crowdsourcing user studies with mechanical turk. In *Proc CHI* (2008), pp. 453–456.

[37] Kwak, H., Lee, C., Park, H., and Moon, S. What is twitter, a social network or a news media? In *Proc WWW* (2010), pp. 591–600.

[38] Leverage New Age Media. Social Media Stats 2014 INFOGRAPHIC (permission to use). https://leveragenewagemedia.com/blog/social-media-infographic/. Last accessed: 04/17/2014.

[39] Linder, R., Snodgrass, C., and Kerne, A. Everyday ideation: all of my ideas are on pinterest. In *Proc CHI* (2014), pp. 2411–2420.

[40] Lindley, S. E., Meek, S., Sellen, A., and Harper, R. It's simply integral to what i do: enquiries into how the web is weaved into everyday life. In *Proc WWW* (2012), pp. 1067–1076.

[41] Mahler, J. On Twitter, Mourning is Collective - NYTimes.com. `http://nyti.ms/Y3qfcc`. Last accessed: 09/01/2014.

[42] Marchionini, G. *Information seeking in electronic environments*. No. 9. Cambridge University Press, 1997.

[43] Marcus, A., Bernstein, M. S., Badar, O., Karger, D. R., Madden, S., and Miller, R. C. Twitinfo: aggregating and visualizing microblogs for event exploration. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (2011), ACM, pp. 227–236.

[44] Marlow, C., Naaman, M., Boyd, D., and Davis, M. HT06, tagging paper, taxonomy, flickr, academic article, to read. In *Proc HYPERTEXT* (2006), pp. 31–40.

[45] Marshall, C., and Shipman, F. Experiences surveying the crowd: Reflections on methods, participation, and reliability. In *Proc WebSci* (2013), pp. 234–243.

[46] Moran, M., Seaman, J., and Tinti-Kane, H. Teaching, learning, and sharing: How today's higher education faculty use social media. *Babson Survey Research Group* (2011).

[47] National Academy of Engineering. *Rising Above the Gathering Storm, Revisited: Rapidly Approaching Category 5*. The National Academies Press,

2010.

[48] Qu, Y., Kerne, A., Lupfer, N., Linder, R., and Jain, A. Metadata type system: Integrate presentation, data models and extraction to enable exploratory browsing interfaces. In *Proc EICS* (2014), ACM.

[49] Salton, G., and Buckley, C. Term-weighting approaches in automatic text retrieval. *Information Processing and Management 24*, 5 (1988).

[50] Scaife, M., and Rogers, Y. External cognition: how do graphical representations work? *International journal of human-computer studies 45*, 2 (1996), 185–213.

[51] Shah, J. J., Smith, S. M., and Vargas-Hernandez, N. Metrics for measuring ideation effectiveness. *Design studies 24*, 2 (2003), 111–134.

[52] Shamma, D. A., Kennedy, L., and Churchill, E. F. Tweet the debates: understanding community annotation of uncollected sources. In *Proc SIGMM* (2009), pp. 3–10.

[53] Simon, H. A. How big is a chunk. *Science 183*, 4124 (1974), 482–488.

[54] Smith, M. A., Rainie, L., Shneiderman, B., and Himelboim, I. Mapping twitter topic networks: From polarized crowds to community clusters | pew research center's internet & american life project. `http://pewrsr.ch/1oWq6Am`. Last accessed: 04/21/2014.

[55] Starbird, K., and Palen, L. (how) will the revolution be retweeted?: information diffusion and the 2011 egyptian uprising. In *Proc CSCW* (2012), pp. 7–16.

[56] Suh, B., Hong, L., Pirolli, P., and Chi, E. H. Want to be retweeted? large scale analytics on factors impacting retweet in twitter network. In *So-*

cial Computing (SocialCom), 2010 IEEE Second International Conference on (2010), pp. 177–184.

[57] Tufte, E. R. Envisioning information. *Optometry & Vision Science 68*, 4 (1991), 322–324.

[58] Twitter. About TweetDeck. `https://about.twitter.com/products/tweetdeck`. Last accessed: 04/11/2014.

[59] Von Landesberger, T., Kuijper, A., Schreck, T., Kohlhammer, J., van Wijk, J. J., Fekete, J.-D., and Fellner, D. W. Visual analysis of large graphs: State-of-the-art and future research challenges. In *Computer graphics forum* (2011), vol. 30, Wiley Online Library, pp. 1719–1749.

[60] W3C. XML Path Language (XPath) 2.0. Tech. rep., 2010.

[61] White, R. W., and Roth, R. A. Exploratory search: Beyond the query-response paradigm. *Synthesis Lectures on Information Concepts, Retrieval, and Services 1*, 1 (Jan. 2009), 1–98.

[62] Wilson, M. L., Kules, B., schraefel, m., and Shneiderman, B. From keyword search to exploration: Designing future search interfaces for the web. *Foundations and Trends in Web Science 2*, 1 (2010), 1–97.

[63] Wilson, M. L., and schraefel, m. Improving exploratory search interfaces: Adding value or information overload? In *Proc HCIR* (2008).

[64] Wu, S., Hofman, J. M., Mason, W. A., and Watts, D. J. Who says what to whom on twitter. In *Proc WWW* (2011), pp. 705–714.

[65] Zellweger, P. T., Chang, B.-W., and Mackinlay, J. D. Fluid links for informed and incremental link transitions. In *Proc HYPERTEXT* (1998), ACM, pp. 50–57.

APPENDIX A

CHROME EXTENSION

Chrome extension technology enables broader distribution and deployment of a new feature or interaction technique by making it available to users through the Chrome Web Store. Users are able to review the permissions required by the application when they choose to deploy it. A manifest file describes the permission that the app requires and the resources it is going to use. Additionally, it describes the content scripts that will get injected into the web pages whose urls match the patterns included in the manifest file.

The manifest file specifies background pages that allow managing tasks and states through a singleton instance that is common across all the web pages into which the content scripts have been injected. Background pages are also particularly enabling in that they allow cross-origin requests. This overcomes any permission issues that developers are faced with when using APIs such as `XMLHttpRequest`. Several Chrome extension features such as storage, access to cookies, tabs, and windows are only available to the background pages. Content scripts can communicate with background pages via message passing.

A small part of the TweetBubble extension manifest file is included below (Figure A.1). A complete version can be seen at the GitHub project location [31]. The extension is available for free on the Web Store and the TweetBubble project is open-source. Further details on Chrome extension technology can be obtained from the Chrome Developer site [23].

```
{
 "manifest_version": 2,

 "name": "TweetBubble",
 "description": "Expand to explore @usernames, #hashtags on Twitter. See related
content without switching back and forth between tabs and windows",
 "version": "0.9.51",
 "permissions": [
  "https://twitter.com/*",
  "http://ecology-service.cs.tamu.edu/",
  "tabs“,
  ….
 ],
 "content_scripts": [
  {
   "matches": [
   "https://twitter.com/search?q=%23*",
  "https://twitter.com/*"
   ],
   "run_at": "document_end",
   "css": ["content_script/css/mice.css",
    "content_script/css/infosheet.css",
    "content_script/css/twitterMice.css"
   ],
   "js": [
    ….
    "content_script/js/debi.js",
    "content_script/js/twitterDebi.js",
    "content_script/js/mice.js",
    "content_script/js/twitterMice.js",
    "content_script/js/twitterICE.js",
    "content_script/js/ICE.js"
   ]
  }
 ],
 "background": {
  "page": "background/background.html"
 },
 "web_accessible_resources": [
  ….
 ]
}
```

Figure A.1: Manifest file describes the permissions and resources required by the Chrome extension. It also contains the version number and a description which appear on the Chrome Web Store and chrome://extensions page.

APPENDIX B

METADATA EXTRACTION

On dynamically encountering a social media association —@username or #hashtag—the extraction and presentation are derived using corresponding type specification in the Meta-Metadata language. To obtain the respective Meta-Metadata description, we used BigSemanticsService [29] a RESTful web service interface to the Meta-Metadata semantics, which is available as a part of the open-source BigSemantics framework [28].

The service request *params* include request type (metadata / meta-metadata), response format (xml or json), and the source url. At the server end, a *selector* logic facilitates identification of the appropriate Metadata type based on the url pattern and mime type of the encountered source. The Meta-Metadata description is returned in a serialized form from the post-inheritance version of metadata types. In the post-inheritance version, the references contained within the types are resolved, making each source description a self-contained unit and describing the corresponding data model, extraction rules, and presentation semantics.

The extraction rules are then applied to the Document Object Model (DOM), which is obtained through `XMLHttpRequest` in the extension's background page. As discussed, this also helps sidestep any authentication issues and capture the user's social media context. Extraction rules consist of XPaths, direct binding of names to elements (in case of XML or JSON documents), regular expressions, and field parsers, which make it possible to extract the exact information that is needed. The details of using these constructs can be found in Qu et al's work [48].

APPENDIX C

PRE-STUDY QUESTIONNAIRE

Questions marked with an asterisk * were mandatory.

1. What is your gender?*

   (a) Male (b) Female (c) Other / Prefer not to say

2. What is your age? Please enter a numerical age.*

3. How many hours do you usually spend on twitter each week?*

   (a) 0 (b) 1 (c) 1-5 (d) 5-10 (e) >10

4. How many hours do you spend on Facebook per week?

   (a) 0 (b) 1 (c) 1-5 (d) 5-10 (e) >10

5. Pinterest?

   (a) 0 (b) 1 (c) 1-5 (d) 5-10 (e) >10

6. How often do you usually tweet?*

   (a) Almost Never (b) Once a month (c) Once a week

   (d) Once a day (e) Almost every hour (f) Multiple times per hour

7. What category of topics are you mostly interested in? [Multiple Choice]*

   (a) sports (b) entertainment (c) arts (d) science (e) business

   (f) marketing (g) world news (h) US news (i) others: please specify

8. Who do you interact most with? [Multiple Choice]*

   (a) friends (b) celebs (c) trends (d) others: please specify

9. What event or trending topic do you remember last noticing on twitter?*

10. How did you explore the event / topic on twitter?*

11. I have decided not to explore an event or topic I noticed on twitter because it was too much work*

    (a) Never (b) Rarely (c) Sometimes (d) Often

12. How much do you browse across Twitter feeds, such through #hashtags and @users?*

    (a) Almost Never (b) Once a month (c) Once a week

    (d) Once a day (e) Almost every hour (f) Multiple times per hour

13. How easy do you find it to browse across Twitter feeds, such through #hashtags and @users?*

    (a) Very difficult (b) Moderately difficult (c) Neither easy nor difficult

    (d) Moderately easy (e) Very easy

14. If you encounter, any difficulty in browsing across Twitter feeds, please describe.

15. How much do you follow links from Twitter to other websites?*

    (a) Almost Never (b) Once a month (c) Once a week

    (d) Once a day (e) Almost every hour (f) Multiple times per hour

16. How easy do you find it to follow links from Twitter to other websites?*

    (a) Very difficult (b) Moderately difficult (c) Neither easy nor difficult

    (d) Moderately easy (e) Very easy

17. If you encounter, any difficulty in following links from Twitter to other websites, please describe.

(Sanity Check Questions)

18. Who of the following is not nominated for Best Actress or Best Supporting Actress:*

    (a) Meryl Streep (b) Jennifer Lawrence (c) Miley Cyrus (d) Sandra Bullock

19. Which of the following is not a famous dress designer:*

    (a) Vera Wang (b) Estee Lauder (c) Alexander McQueen (d) Christian Dior

APPENDIX D

POST-STUDY QUESTIONNAIRE

Questions marked with an asterisk * were mandatory.

1. Did you face any issue in installing the extension, or any other issue that prevented using it?*

   (a) Yes (b) No

2. If yes, please specify.

3. I browsed across twitter feeds, such as through @usernames and #hashtags*

   (a) Never (b) Rarely (c) Sometimes (d) Often

4. How did clicking related content e.g. @usernames, #hashtags affect your Twitter experience?*

5. I sometimes did not click on the related content as I did not want to lose the context*

   (a) Strongly disagree (b) Disagree (c) Neither disagree nor agree

   (d) Agree (e) Strongly agree

6. I could relate a lot of content on a single page and didn't require to switch back and forth much between open tabs / windows*

   (a) Strongly disagree (b) Disagree (c) Neither disagree nor agree

   (d) Agree (e) Strongly agree

7. How did what you could relate on same page affect your experience?*

(Attention Check Question - #8)

8. How many times in the last 5 minutes have you had a heart attack while riding a unicycle in the Alps?

   (a) 0 (b) 1-3 (c) 4-6 (d) >6

9. I could develop a good understanding of the topic browsing across a large number of related feeds*

   (a) Strongly disagree (b) Disagree (c) Neither disagree nor agree

   (d) Agree (e) Strongly agree

10. How did browsing across a large number of related feeds affect developing an understanding of the topic?*

11. Briefly describe what and how you browsed on Twitter. Did you learn anything new? How were the different #hashtags and @usernames related?

    Please mention a few #hashtags, @usernames, or related content that you browsed. Tell a story.*

12. Do you think the extension would be useful for regular twitter browsing as well?*

    (a) Strongly disagree (b) Disagree (c) Neither disagree nor agree

    (d) Agree (e) Strongly agree

13. Any feature you liked or disliked. Any feature you would like to see added.

14. I plan to continue using TweetBubble.*

    (a) No (b) Perhaps (c) Undecided (d) Probably (e) Definitely

15. I plan to refer friends and people in my network to use TweetBubble.*

    (a) No (b) Perhaps (c) Undecided (d) Probably (e) Definitely