

# MEDIA LIFECYCLE AND CONTENT ANALYSIS IN SOCIAL MEDIA COMMUNITIES

*Lexing Xie*

Australian National University  
lexing.xie@anu.edu.au

*Hari Sundaram*

Arizona State University  
hari.sundaram@asu.edu

## ABSTRACT

This paper examines the role of content analysis in media-rich online communities. We highlight changes in the multimedia generation and consumption process that has occurred the past decade, and discuss several new angles this has brought to multimedia analysis research. We first examine the content production, dissemination and consumption patterns in the recent social media studies literature. We then propose an updated conceptual summary of media lifecycle from a previous research column [6] by Chang. We present an update list of impact criteria and challenge areas for multimedia content analysis. Among the three criteria, two are existing but with new problems and solutions, one is new as a results of the community-driven content lifecycle. We present three case studies that addresses the impact criteria, and conclude with an outlook for emerging problems.

*Index Terms*— Multimedia, social networks, content analysis, media lifecycle.

## 1. INTRODUCTION

Multimedia content analysis is a young research area that underwent significant development in the past decade. In the early 2000s, it “refers to the computerized understanding of the semantic meanings of a multimedia document, such as a video sequence with an accompanying audio track” [32]. Is this definition still accurate, given that the ways multimedia documents are produced, disseminated, and consumed are constantly shaped by new tools and improved infrastructure? Moreover, what is the “Holy Grail for content-based media analysis in practical applications”? The answer was [6], “less about a specific algorithm or service than a rigorous methodology to formulate and evaluate content-based analysis research”. Does this methodology still apply, what about the answers it produced?

The goal of this paper is to answer the two meta-questions above, in the current context, ten years after the original definition was in place and the analyzes were carried out.

Inspired by the insights from Chang [6] and Wang et al [32], we set out by examining patterns in the multimedia content lifecycle, taking into account the prevalence of social media platforms and rich user interactions around the audio-visual content. We summarize changes in the multimedia content chain, showing that multiple feedback loops has formed from a linearly-shaped content chain ten years ago [6]. We then propose three impact criteria for content analysis research, two of which existed in prior analyzes while each has a new angle, and one is new due to the content co-creation lifecycle. We further illustrate the impact criteria with three case studies, and conclude with a brief outlook for emerging problems. The main messages of this position paper include:

- There is a strong element of co-creation in content production and consumption patterns. This has brought changes to the

traditional broadcast and personal content space, while creating a novel pattern for community-driven media generation (Fig. 1 (a-b)).

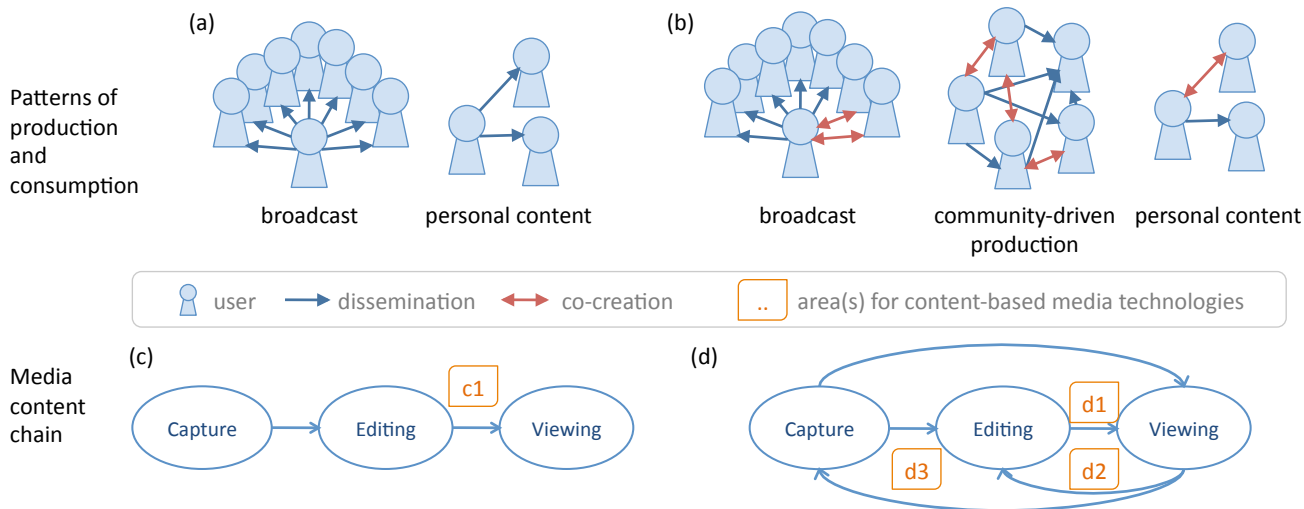
- The multimedia content chain has evolved from a linear progression to one with multiple closed loops between capturing, editing and consumption (Fig. 1 (c-d)). This opens new opportunities for content-analysis in community-driven production.
- The impact criteria for generating metadata and on high-volume low-value content still hold. Although their interpretation and application domains has changed.
- A new impact criteria should be added for incorporating joint analyses by human(s) and automated algorithm.

In our opinion, such a discourse is timely. Although there has been many work on community-centric content analysis for the past few years, no articulation is yet available about how our problems and solutions has evolved from the web 1.0 era. On the other hand, we are aware of several apparent limitations of this paper. First, this paper is not intended as a literature survey. We refer the reader to a number of excellent surveys in content analysis [27, 8], knowledge discovery from community centric media [14], and social signal processing [30] instead. Second, we risk stating the obvious for experts in the area, and are not able to many significant work within social multimedia analysis within the space limit. Third, all three case studies are our own work, due to space constraint and our in-built knowledge of their limitations. Nonetheless, we sincerely hope that by analyzing the problem space and its causes will lead to lively meta-discussions about media-rich social networks in the multimedia community itself.

## 2. NEW PATTERNS IN CONTENT LIFECYCLE

Considerable change has taken place in how multimedia content was produced and consumed in the past ten years. In this section we start by reviewing the multimedia production and consumption model in the early 2000s. We outline patterns that arise with new media platforms and tools, we then summarize the observations into a new content chain for media generation and consumption, along with a note on the inequality of online social participation.

In 2002 Chang presented a summary [6] of the multimedia content chain, accompanied by important areas of research that tackles problems in various stages of such content chain. A simplified rendering of Chang’s content chain is in Fig.1(c). Here multimedia content (e.g. video and audio) go through three main stages: capturing, editing and viewing, each governed by research principles in “production model”, “media integration”, and “viewer model” [6]. This content chain is driven by two prominent user patterns (Fig.1(a)).



**Fig. 1.** Evolution of multimedia content chain. (a) Content dissemination patterns in traditional broadcast and personal content. (b) Content production patterns in online platforms with media-rich communities. (c) Multimedia content chain by Chang in 2002 [6]. (d) Evolved multimedia content chain in online communities.

One mode is broadcasting, where content is generated by professional producers and consumed by many in the general public (such as films and television). The other mode is personal content sharing, where content is generated by amateurs or average users, and consumed by the user herself or a handful of other people (such as personal photos and videos, surveillance video feeds, and so on).

### 2.1. Participatory culture and vernacular creativity

Such uni-directional media production and consumption model has changed. The main driving forces of such change include three aspects. The first include improved device and infrastructure for capturing, editing, and sharing media. For instance, numerous smartphones and portable cameras made producing media easy; 3G or wi-fi connections enable users to share content almost anywhere, anytime. The second is highly usable content-sharing platforms, such as Flickr, YouTube, Facebook. Such platform hosts not one type, but a wide range of content from diverse user groups, they also provides the social incentive for producing and sharing content with feedback mechanisms for commenting, likes and popularity ranking by a large audience. The third is the discussion around media content in online social networks. This is exemplified by “generation Y” or other avid social network users who are constantly online, constantly sharing, and whose interest extend well beyond mass-media content [3].

These three interacting factors has added a prevalent user pattern and changed the two existing ones, as illustrated by Fig.1(b). In the middle is the community-driven production model, users often play the dual role of content creators and consumers. Media content is not only being produced and consumed, it is also constantly discussed, remixed, and evolved on these social platforms. Such consumer co-creation model is also influencing the broadcast and personal content sharing practice (the left and right part of Fig.1(b)). Major broadcasters has opened up channels for users to share content<sup>1</sup>. And personal content space has also included feature for adding narratives, tagging people, and linking to other content easily.

All three usage patterns coexist on social media platforms, and users co-participate in more than one mode. YouTube, for example, is known as a “high-volume website, a broadcast platform, a media

archive, and a social network, . . . , and illustrates the increasingly complex relations among producers and consumers in the creation of meaning, value and agency” [5].

Such coexistence created a continuum from broadcast to social discussions, from professional producers to casual amateur users. One example of the mix of user expertise in the coverage of ongoing news topics. Youtube is “ruled by the clip and the quote” [5], and that users tend to create curated selections based on what they liked or thought was important ([28], page 270), especially since making original videos is difficult for most video makers who are not at the news site. It is recently observed that video content can be remixed by professional producers from citizen journalists, or vice versa [33]. Another example of such continuum is in the skill progression of users. In an interview of several Flickr users, Burgess [4] noted that these users use the media sharing platform as a place, tool, and incentive for learning and discovering photography. This study has also observed that several users has since become skilled photographers by reading books, taking courses, or self-taught in practice ([4], chapter 4). Such skill progression is also seen in video bloggers (or vlogs, in [5] “Literacy and the Social Network”), who progress from “straight to camera shots” to “more complex overlay with photos, captions and other editing operations”.

### 2.2. Evolution of content chain

From the observations on media participation platforms, we propose a revised abstraction to the multimedia content chain, illustrated in Fig.1(d). Such online content chain is important and increasing relevant, as multimedia takes up half of consumer internet traffic in 2012, and this is poised to grow to over 60% by 2015 [2].

As shown in Fig.1(d), the proposed contain lifecycle is no longer chain-shaped, there are several connections being added among the three basic stages. The skip arrow between capture and viewing is due to the prevalence of short video clips and portable recording devices. It is now commonplace to share an unedited image or short video clip directly from one’s mobile phone, some of which has become important journalistic icon or viral video [33]. The first feedback arrow from viewing to editing is due to social remix, where viewers decide which clips and segments are worth clipping and quoting, and then re-edit and re-assemble new video clips based on

<sup>1</sup> <http://ireport.cnn.com>

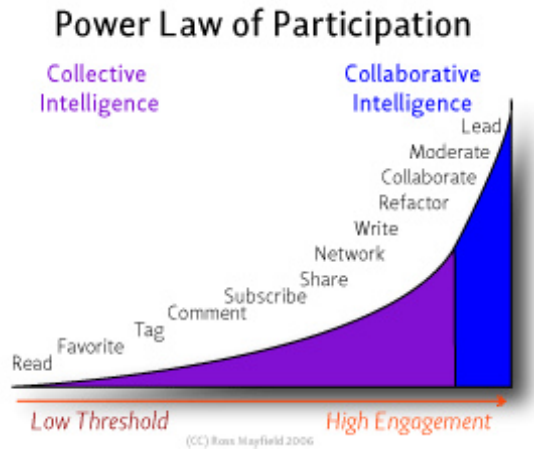


Fig. 2. Powerlaw of participation by Ross Mayfield [21].

the same captured material, sometimes adding a new audio overlay. The second feedback arrow from viewing to capturing often requires community intelligence. This applies to scenarios when viewing community content will affect when, where, or what to capture, such as depicting landmarks [15] or planning travel [11].

### 2.3. The power law of participation

It is difficult to gauge to what extent the community-centric media cycle is replacing or complimenting the traditional content cycle, while it is well known that user activities in a social information network has a highly skewed distribution. This is known as the 90-9-1 rule of participation inequality, famously summarized by Nielsen [22] as: “In most online communities, 90% of users are lurkers who never contribute, 9% of users contribute a little, and 1% of users account for almost all the action”. Specific percentages in different social platforms vary, wikipedia has a ratio of 99.8-0.2-0.003, for example, however the general trend holds for most systems.

Such uneven rate of participation has two seemingly opposite effects on social media systems. On one hand this incurs (unknown) sampling bias – we need to be careful when making conclusions about user preference and behavior from ten percent of the population who like to use the product. On the other hand “the hurdles that users cross as they transition from lurkers to synthesizers to creators are also filters that can eliminate noise from signal” [13].

Fig.2 presents a visualization of the power-law participation pattern, proposed by Ross Mayfield [21]. The horizontal axis of this graph highlights the the different modes of participation. Building algorithms and systems for exploiting such a large variety of engagement is part of the challenge we face. What are the social and preferential implications of low-threshold participation and intense engagement? What are the right incentives? One success example of implicitly using different participation mode is flickr interestingness [13], where a subtle implementation taking into account photo views, comments, favorites have generated apparently good results.

## 3. WHAT DOES CONTENT LIFECYCLE MEAN?

The new trends in multimedia content lifecycle has brought changes in both problems definition and solution domain for multimedia content analysis. There are two main driving forces behind such changes.

The first is that training data has become increasingly cheap and available in large quantities, while the label quality can vary

and requires careful filtering and calibration. One point in case is image labeling. A number of large datasets has emerged in the compute vision and multimedia community, such as ImageNet [10], LabelMe [25], NUSWide [7], Social20[18] and others. Crowdsourcing platforms such as Amazon Mechanical Turk has become a popular means for gathering data, and data-quality measures must be in place for the process to work, such as eliminating input from unreliable users and use redundant labels to reduce label noise [23].

The second force is that visual recognition has seen significant progress in the past ten years. Realistic object matching, detection and search in mobile apps are in wide use (such as [1]), thanks to the invention of scale-invariant image keypoints [20] and boosted cascade of visual classifiers [31]. This progress scaffolds the expectations on multimedia indexing and recognition to a higher level. Both users and technologies are demanding more reliable and more versatile set of visual indexes, including visual attributes, affect, face identities, and so on.

### 3.1. Impact criteria and problem areas

The new content trends and the extent of user participation need to be taken into account for problem definitions and solutions in a number of areas. This shifts existing notions of impact, and has brought new angles.

- A *Providing metadata that are not available from production, and humans are not good at generating* [6]. While this is still true, the application domains and types of metadata needed are constantly changing. Take broadcast content, for instance, although more and more production-time metadata are available from online sources such as imdb.com for films or tv.com for television episodes, users are demanding new level of details from visual indexing. For films, video google allows people to search for particular objects and people in different scenes [26]. In sports [29], user-created summaries of a particular athlete, or snippets of either highlight or controversial moments are part of active discussions of the games.
- B *Focusing on content with large volume and low individual value* [6]. This remains true, while its solution strategy changes from content analysis for each multimedia document, to the joint analysis of a collection of such documents in social context. Take personal content, such as photos and home videos, for example, each media item has low creation effort and low individual value. Many approaches aggregate group or community content to make sense of individual media item, making the sum of many parts greater than the whole. Examples include landmark recognition [15], social event recognition [24], social media travel guides [11], and the analysis of group relations in Sec. 4.1.
- C *Focusing on joint analyses with both human input and content analysis as components*. This is a new problem area that result from community-driven content lifecycle, which itself can be seen from two angles. On one hand we can pose a human-driven problem with automatic analysis as a component. Take the traditional domains of surveillance and satellite imagery, for example, there are abundant content, scarce annotation, and crowd-sourced analysis is naturally suitable for high-value targets. The collaborative search for Jim Gray [12] has shown that an effective large-scale collaboration can be put in place within a short time, it also opened up many questions for composition of human-in-the-loop processes, quality verification, and scaling out to many media modalities. On the other hand we can pose a media

understanding problem with a human-understanding component. The case study in Sec. 4.2 makes use of large-scale video remix to infer content interestingness and user influence. Another case study in Sec. 4.3 uses the conversation and comments around a YouTube video to characterize conversational themes, participants, and the quality of the conversations.

#### 4. CASE STUDIES

In this section we present three case studies for multimedia content analysis in social network communities. All three cases highlights the value of collective intelligence on social media – it is the uploading, tagging, commenting and remixing behaviors that led to new insight on the content and people who are interacting around them. Specifically, the Flickr group study in Sec. 4.1 shows the collective value of large amounts data from the long-tail, addressing impact criteria B; the visual memes study in Sec. 4.2 shows the value of community action, i.e. remixing, as a tool for ranking users in terms of influence and ranking content in terms of virality, addressing criteria C; the YouTube conversations study shows that user roles and content interestingness can be estimated from activities around a content, but not the content itself, again addressing criteria C.

##### 4.1. Multiple modalities in Flickr group interactions

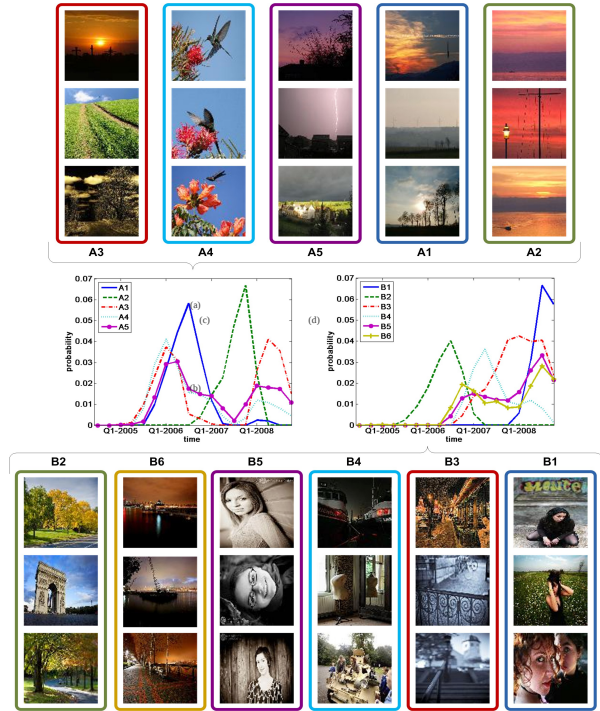
We extract relational structure from Flickr group photo pools<sup>2</sup>. We define a *group photo stream* (or group, for short) to be a collection that includes: photos posted in a *shared space*, all users who posted the photos, and tags associated with the photos. In this work, the shared space specifically refers to Flickr group pools.

There are two key ideas in this analysis: extraction of relations from social media streams and extraction of relational clusters. Relations connect different aspects of the photo stream data. Specifically, this work models visual content, associated tags, photo owners, and post times. We find soft relation clusters with non-negative joint matrix factorization techniques. Details can be found in a separate publication [19].

Fig. 3 shows the results from two Flickr groups: “sky’s the limit” (denote as group A, with 2,278 photos), and “Full Frame Sensor group” (denote as group B, with 3,961 photos). We obtaining five thematic clusters in group A and six in group B. For each thematic cluster we show the top three photos based on the data likelihood  $p_{ik}$  for a photo  $i$  to belong to theme  $k$ . We can determine through other coefficient matrices, the most likely users who post photos belonging to the theme and the most likely tags associated with the theme photos. The middle part of Fig. 3 shows aggregated cluster strength over time for group A and group B. We can see that the theme strengths vary over time; some themes, such as A2 and B2, only appear at certain time periods and then diminish. Some others—A3 and A5 are examples — appear, then fall and then re-appear. We have observed that these themes emerge due to dedicated users (e.g. the “bird” images in A4 are taken by the same user), tag co-occurrences (e.g. “sunset” in A2, “water” in B6, etc.), as well as similar visual content (e.g. A2, A4, A5, B2, B5, B6, etc.). These empirical results suggest that our analysis captures the dynamics of group patterns and gives meaningful summary of group photo streams.

##### 4.2. Visual memes in real-world events

This case study develops a method for tracking “visual quotes” (i.e., memes) in social media, and uses the outcome of meme tracking to



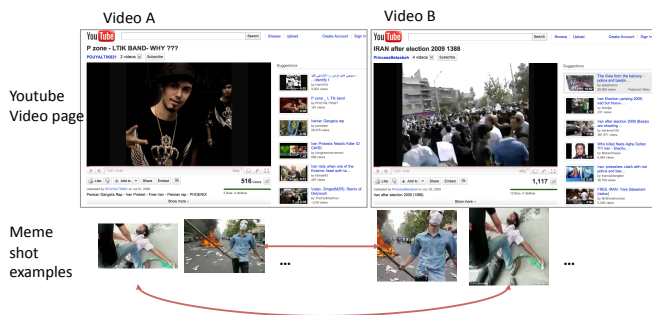
**Fig. 3.** Theme discovery for Flickr group A “sky’s the limit” (top figure) and group B, “Full Frame Sensor group” (bottom figure). Theme evolutions are shown in two middle plots, with the mid-left plot from “sky’s the limit” and the mid-right plot from “Full Frame Sensor group”. The results show that group patterns emerge due to dedicated users, tag co-occurrences, as well as similar visual content.

infer social influence in news events. Remixing and reposting are prevalent on video sharing platforms like YouTube. A number of studies has covered quoting and remixing on text-based social platforms, especially Twitter. activity indicators that reflect information propagation on text-based media such as blogs, news [17] and microblogs [16]. No quoting equivalent has so far been available for videos. *Visual meme* is a short segment of video that is frequently remixed and reposted by more than one user. Visual memes are not visible in current video-sharing platforms, as shown in Fig.4. Visual memes are detected using scalable algorithms on millions of video shots with high accuracy (>96% precision and 80% recall) [33].

We can estimate the influence of content and of authors using visual memes. Visual memes can be viewed as *links* between creators videos that share the same visual segment. We derive a link-based measure—*diffusion influence index*—to depict the influence of a meme and its author. We take data set Iran3, containing more than 20,000 videos over a three-month-period In Fig. 5 (right), we examine author influence index (a measure that gives higher score for earlier posters of a popular meme) versus the number of videos produced by each author. We can see a few distinct types of contributors. We call one type “maven” (marked in red), who post only a few videos that end up being massively remixed and reposted—this particular maven was among the first to post the murder of Neda Soltan<sup>3</sup> and one other instance of a student murder on the street. The

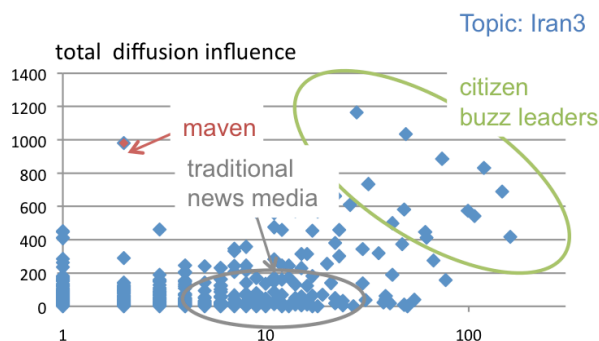
<sup>2</sup><http://www.flickr.com/groups/>.

<sup>3</sup>[http://en.wikipedia.org/wiki/Death\\_of\\_Neda\\_Agha-Soltan](http://en.wikipedia.org/wiki/Death_of_Neda_Agha-Soltan)



**Fig. 4.** Visual meme examples. Two YouTube videos that share a few different memes. Note that it is impossible to tell from metadata or the YouTube video page that they shared content, and that the appearance of the remixed shots (bottom row) has large variations.

former become the icon of the entire event and the face of Iranian struggle during this turbulent period. A second group can be dubbed “citizen buzz leaders” (circled in green), who tend to produce a large number of videos with high total diffusion factor, yet relatively low influence per video. They aggregate notable content and come relatively late in the timeline, which is penalized by the influence factor. We examined the YouTube channel pages for a few authors in this group, and they seem to be voluntary activists with screen names like “iranlover100”. Some of their videos are slide shows of iconic images and provide good summaries. Note that traditional news media, such as AljezeeraEnglish, AssociatedPress and so on (circled in gray) are ranked rather low for this topic, partially because the Iran government severely limited international media participation in the event, and most of the event buzz was driven by citizens.



**Fig. 5.** Total diffusion influence vs. the number of videos produced by each author on Iran3 dataset.

### 4.3. Analyzing conversations on YouTube

In this case study, we examine a simple question: does the presence of social interaction — facilitated by a social media site — around a media object, add to the value of the media object? We conjecture that people participate in conversations when they find the conversation theme interesting, see comments by people whom they are familiar with, or observe an engaging dialogue between two or more people (absorbing back and forth exchange of comments). Importantly, a conversation that is interesting must be consequential - i.e. it must impact the social network itself.

Our framework has three parts: characterizing themes, characterizing participants for determining interestingness and measures of consequences of a conversation deemed to be interesting. First, we

detect conversational themes using a mixture model approach. Second, we determine interestingness of participants and interestingness of conversations based on a random walk model. Third, we measure the consequence of a conversation by measuring how interestingness affects the following three variables - participation in related themes, participant cohesiveness and theme diffusion. We have conducted extensive experiments using dataset from the popular video sharing site, YouTube. Our results show that our method of interestingness maximizes the mutual information, and is significantly better (twice as large) than three other baseline methods (number of comments, number of new participants and PageRank based assessment).

We can see from Fig. 6, that the themes exhibit high strength, shown in high color intensity, during the times of well known and thematically related news events. The correlation between intensity and external events suggest that our textual theme models accurately capture latent themes including theme dynamics. We observe that the degree of participation is low (thin ring)—at some of the times—when the theme strength is high (ring color intensity). The high theme strength is explained by high communication activity amongst participants interested in the particular news event. Detailed description and results are in a separate writeup [9].

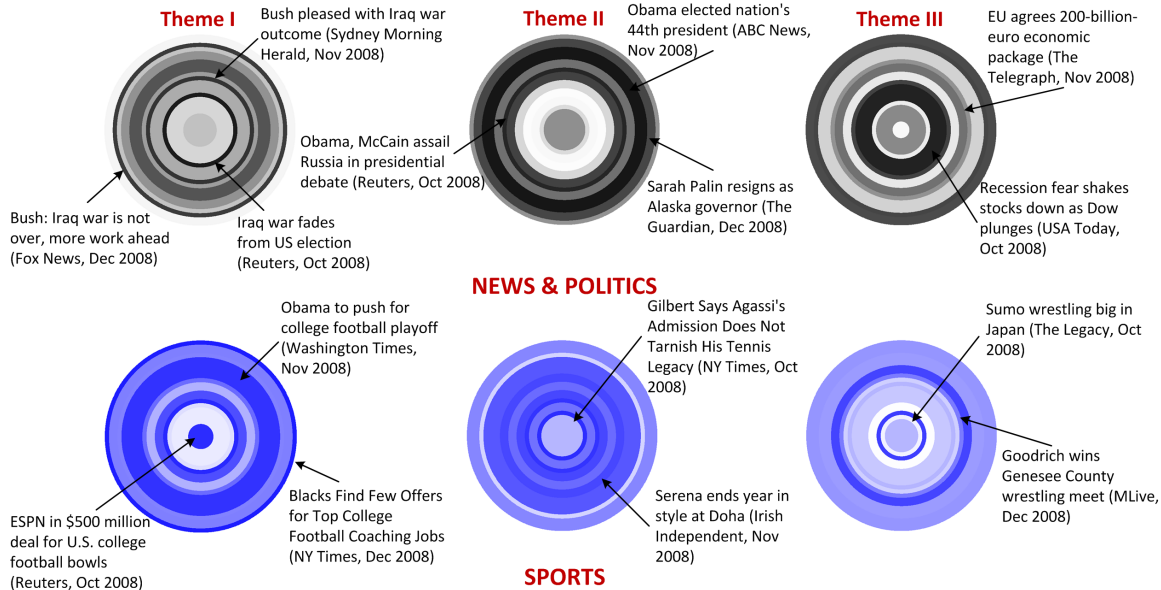
## 5. CONCLUSION AND OUTLOOK

This position paper argues that considerable change has taken place for the definition of multimedia content analysis, as well as its problems and solutions. We examine the current multimedia production and consumption patterns and contrast with those ten years ago [6]. We summarize the media content chain including social co-creation loops, we discuss the resulting problem areas and impact criteria. We showed that two of the well-known criteria can be adapted to the new patterns and one new criterion need to be taken into account.

The focus areas that are may soon receive much attention include crowd-sourced video production, community-related metrics for evaluation, quantify the evolving social incentives for human-in-the-loop analysis.

## 6. REFERENCES

- [1] Google goggle. <http://www.google.com/mobile/goggles/>, retrieved December 2011.
- [2] Cisco visual networking index: Forecast and methodology, 2010-2015, 6 2011.
- [3] C. Anderson. *The Long Tail: Why the Future of Business is Selling Less of More*. Number 10. Hyperion, 2006.
- [4] J. Burgess. *Vernacular creativity and new media*. PhD thesis, 2007.
- [5] J. Burgess and J. Green. *YouTube: Online Video and Participatory Culture*. Polity, 2009.
- [6] S.-F. Chang. The holy grail of content-based media analysis. *IEEE MultiMedia*, 9(2):6–10, 2002.
- [7] T.-s. Chua, J. Tang, R. Hong, H. Li, Z. Luo, and Y. Zheng. NUS-WIDE: A real-world web image database from national university of singapore. In *ACM International Conference on Image and Video Retrieval*, pages 0–8. ACM Press, 2009.
- [8] R. Datta, D. Joshi, J. Li, and J. Z. Wang. Image retrieval. *ACM Computing Surveys*, 40(2):1–60, 2008.
- [9] M. De Choudhury, H. Sundaram, A. John, and D. D. Seligmann. What makes conversations interesting?: themes, participants and consequences of conversations in online social media. In *Proc. World wide web*, pages 331–340. ACM, 2009.



**Fig. 6.** Ripple visualization of conversational themes over time—15 week period between Oct.–Dec. 2008. There are three themes in two categories: politics and sports. The intensity of the ring color represents theme strength and ring thickness represents degree of participation.

- [10] J. Deng, W. Dong, R. Socher, L.-j. Li, K. Li, and L. Fei-fei. ImageNet: A large-scale hierarchical image database. In *IEEE CVPR*, pages 248–255, 2009.
- [11] Y. Gao, J. Tang, R. Hong, Q. Dai, T.-S. Chua, and R. Jain. W2go: a travel guidance system by automatic landmark ranking. In *Proc. ACM Multimedia*, pages 123–132, 2010.
- [12] J. M. Hellerstein and D. L. Tennenhouse. Searching for Jim Gray : A technical overview. *Comm. ACM*, 54:77–87, 2010.
- [13] B. Horowitz. Creators, synthesizers, and consumers. *Eatable Bradley Horowitz*. (25th April 2007), 2006.
- [14] W. Hsu, T. Mei, and R. Yan. Knowledge discovery over community-sharing media: from signal to intelligence. *IEEE ICME'09*, pages 1448–1451, 2009.
- [15] L. S. Kennedy and M. Naaman. Generating diverse and representative image search results for landmarks, 2008.
- [16] H. Kwak, C. Lee, H. Park, , and S. Moon. What is Twitter, a Social Network or a News Media? . In *Proc. WWW*, 2010.
- [17] J. Leskovec, L. Backstrom, and J. Kleinberg. Meme-tracking and the dynamics of the news cycle. In *Proc. SIGKDD*, page 497. ACM Press, 2009.
- [18] X. Li, C. G. M. Snoek, and M. Worring. Learning social tag relevance by neighbor voting. *IEEE Trans. Multimedia*, 11(7):1310–1322, 2009.
- [19] Y.-R. Lin, H. Sundaram, M. De Choudhury, and A. Kelliher. Discovering multirelational structure in social media streams. *ACM TOMCCAP*, 8(1):4:1–4:28, Feb. 2012.
- [20] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *Intl. J. Computer Vision*, 60(2):91–110, 2004.
- [21] R. Mayfield. Power law of participation. Creative-commons licensed, <http://www.flickr.com/photos/ross/135959002> .
- [22] J. Nielsen. Participation inequality: Encouraging more users to contribute, 2006.
- [23] G. Paolacci, J. Chandler, and P. G. Ipeirotis. Running experiments on amazon mechanical turk. *Judgment and Decision Making*, 5(5):411–419, 2010.
- [24] S. Papadopoulos, C. Zgkolis, Y. Kompatsiaris, and A. Vakali. Cluster-based landmark and event detection on tagged photo collections. *Multimedia, IEEE*, (99):1–1, 2010.
- [25] B. C. Russell, A. Torralba, K. P. Murphy, and W. T. Freeman. Labelme: A database and web-based tool for image annotation. *Intl. J. Computer Vision*, 77(1-3):157–173, 2007.
- [26] J. Sivic and A. Zisserman. Video google: a text retrieval approach to object matching in videos. In *Proc. ICCV*, volume 2, pages 1470–1477. Ieee, 2003.
- [27] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain. Content-based image retrieval at the end of the early years. *IEEE Trans. PAMI*, 22(12):1349–1380, 2000.
- [28] P. Snickars and P. Vonderau. *The YouTube Reader*, volume 12. National Library of Sweden, 2009.
- [29] M. Stauff. Sports on YouTube. In *The YouTube Reader*. National Library of Sweden, 2009.
- [30] A. Vinciarelli, M. Pantic, and H. Bourlard. Social signal processing: Survey of an emerging domain. *Image and Vision Computing*, 27(12):1743–1759, 2009.
- [31] P. Viola and M. J. Jones. Robust real-time face detection. *International Journal of Computer Vision*, 57(2):137–154, 2004.
- [32] Y. Wang, Z. Liu, and J. Huang. Multimedia content analysis—using both audio and visual clues. *Signal Processing Magazine, IEEE*, 17(6):12–36, 2000.
- [33] L. Xie, A. Natsev, J. R. Kender, M. Hill, and J. R. Smith. Visual memes in social media: tracking real-world news in youtube videos. In *ACM Multimedia*, pages 53–62, 2011.