13th Annual International Symposium on Online Journalism

Day 2, April 21, 2012: Morning Session - 11:15-12:45 p.m.
New Approaches in Engaging with the News Community

Chair: Gabriela Warkentin, Professor, Universidad Iberoamericana, Mexico City

Panelists:

- **Angela M. Lee,** University of Texas at Austin, and **Seth C. Lewis,** University of Minnesota-Twin Cities: *Audience preference and editorial judgment: a study of time-lagged influence in online news*
- **Alfred Hermida,** University of British Columbia (Canada) and co-authors, **Seth C. Lewis**, University of Minnesota-Twin Cities, and **Rodrigo Zamith**, University of Minnesota-Twin Cities: *Sourcing the Arab Spring: A case study of Andy Carvin's sources during the Tunisian and Egyptian revolutions*
- **Avery Holton, Mark Coddington** and **Homero Gil de Zuniga,** University of Texas at Austin: *Who knows best? Attitudes and perceptions of citizen journalism and the news through the lens of creators and consumers*
- **Emily T. Metzgar** and **Hans P. Ibold,** Indiana University: *Asserting "truth" in political debates: A study of partisan Twitter users*
- **Katarina Stanoevska-Slabeva**, Vittoria Sacco, **and Marco Giardina, University of Neuchâtel (Switzerland):** *Content Curation: a new form of gatewatching for social media?*

Q & A: Gabriela Warkentin and the Panelists

Gabriela Warkentin: Well, hello, and good morning still. It's not noon yet. And welcome to a very, very interesting panel we're going to have right now. It's a research panel. And first of all, I want to thank Rosental and Amy for having me moderate this panel, because that made me read the whole papers. [laughter] And I think that that's something I really want to encourage you all to do. I mean, these are…. We're going to see, I'm sure, very good presentations, but I really encourage you to read the papers, because it's always better to have them.

Just to introduce myself, I'm Gabriela. I come from Mexico City. And I work both in university and in media. So, I am really glad to be here. I have been attending some and following almost all of these conferences in the last years. And I think that this is a very unique opportunity to have people from universities—scholars, researchers—and people from the media and just

trying to talk to each other, which is not always easy. I mean, I've been in academia long enough to know that in communication it's something [that] is really different to talk to the professional sector and visa versa. So, I'm really glad about the opportunity of being here.

And I must say we've had some fascinating panels. And I was skeptical about something. It was interesting to hear about the last panel when Angelica Peralta said that she encounters skepticism when they talk about data and so on. I must say I was quite skeptical about the panel before, and it was absolutely fascinating. It's so beautiful to see that data can be beautiful. And I'm really glad. And the panel before that, the topic before that with Bob Metcalfe, and then of course the day yesterday.

So today, you're going to hear about audiences, about engaging audiences, which is something that has been in the air here this day-and-a-half. And I have more questions when I went through the papers and actually certainties, which is good, because I think that the five papers we're going to hear today, they address very new phenomena. And we really have to think a lot about what you're seeing.

So, we're going to talk about and hear about tracking audiences and something that's interesting, the path of influence. Who influences whom? Which is interesting about editors and audiences and sort of like the rating of readership, which is interesting. We're going to hear about the use of Twitter in Arab Spring, of course, but also in different areas, and what that means for building a story and how to use it. We heard about that yesterday also with the tornado issue and so on in Memphis. But I think that this is an important issue. And of course, then again about participatory journalism and one issue which I think is basic which is the issue of trust. And then something that is a very nice word or a very nice new concept with this, content curation, which I think it sometimes sounds a bit more elegant to be a content curator than maybe an editor. I don't know. We have to talk about that also. So, that's what we're going to hear today.

And I'm going to just invite our first speaker. And she is Angelica — Angela, sorry, Angela. No Angelica. Angela Lee. And so please, take the floor.

Angela Lee: Thank you. So, good morning.

Audience: Good morning.

Angela Lee: The paper I'd like to talk to about today looks at time-lagging facts of audience preferences on editorial judgments. But before I go into the paper, I would like to first contextualize the media landscape a little bit. So, as we all know, with the rise of cable television and the Internet today, we have a lot more supply for news than there are demands. So, we're in a high choice media environment. So with this, that means that audiences have the power to choose what they're looking at and what they're consuming. And in

addition to that, Bockowski and other people have found that there is this taste disparity between what audiences like to see. Audiences prefer non-public affairs news, and journalists believe that we should be offer public affairs news, so there is a taste disparity. And on top of that, there is also this power struggle where journalists believe that we know the best. We know what people should know. But at the same time, audiences are the ones that are driving the traffic online, which also translates into profit.

And on top of that, how many of you in here are familiar with Charpy? Yes? So, this is basically an audience metrics service that allows newsrooms to look at any given moment, what people are clicking on, what stories are getting the clicks, and what has the potential of becoming more viral. And the bottom line for this is that audience data are here to stay. As time goes on, we'll only know more and more about audiences because of these services.

So, the questions that were asked in this paper is essentially so we know that there is a taste disparity between journalists and audiences. So, who's influencing whom in this disparity? And just real quickly about the methods. We did a fixed effects regression [model] using SEM [structural equation modeling]. And we collected the data at four time points. So Time 1 will be 9:00 in the morning, Time 2 will be 12:00 p.m., Time 3 will be 3:00 p.m., and Time 4 will be 6:00 p.m. And this is a secondary analysis data, so the data [is from] New York Times, New York Post, and the Daily Post. And so, we're basically looking at whether editorial judgment [is] affecting audiences or visa versa.

And so, just a visual presentation of the model that we actually estimated. The red arrows that you see are the ones that were presented in this paper. There are the effects, the time-lagged effects we're looking at. All the other lines that you see on there are a little scary or confusing. They're just the controls that we included in the analysis to make sure that whatever we're presenting to you are the pure effects.

So, for example, when I'm looking at whether editorial judgment affect time — I'm sorry — whether editorial judgments affect audience preferences three hours later, it's possible that earlier editorial stories affect later editorial and the same thing for audiences. So, the thing is that it has a really solid control of spurious factors.

And just real quickly, basically, editorial judgment is the same thing as news placements. And so, we're looking at how prominent the story is on the homepage. And the way that we look at this is when you're on the top of the screen, that's suggests that a story is more important to the editor. So, it's from top to the bottom, left to the right, and also the headlines are from larger headlines to smaller headlines. And for audience preferences, it's pretty much just audience clicks and we use most viewed clicks from the three newspapers.

So now, going into the results. Are there time-lagging facts at the three-hour time-lag? And it turns out that audiences preferences actually does affect how other editors place the stories on the homepage three hours later, but we do not see the same effect in the other way around. And I'm going to talk a little bit more about the effects in the next few slides.

And so, kind of to sum up the finding is that audience preferences do affect editorial decisions three hours later across the three newspapers. And more interesting is that the effect actually intensifies during the course of the day. It makes sense in the sense that if we think about how data works or audience data, as time goes on, we're only going to have more and more data, because more people are coming to our websites to click on the stories. So, we know what people like to read and what they are actually reading, and so that makes sense in that way. But on the other hand, it is a little surprising that we do not see time-lagged effects of editorial on audiences.

And now turning to a different question. As you may be wondering, New York Times, New York Post, and the Daily News are really different papers. They are different in terms of their content, their goals, their organizational norms, and their subscription models. So, isn't it likely that the three papers may be different in how they use the audience data and so forth? And that's pretty much what we asked in the post hoc analysis.

And so, if we look at the Daily News, the data suggests that in the morning when a story gets more popular on the most viewed rank, it actually gets moved down on the homepage. And this is from 9:00 to 12:00 and 12:00 to 3:00. But when we look at the editorial effect on audience three hours later, there is no effect. And the only positive effect for New York Post would be in the morning from editorial, which suggests that when the story gets more popular on the homepage, it also gets moved up a little bit on the most viewed rank.

And the even more interesting thing about the New York Times is that we actually see a symbiotic relationship between audiences and editorial, in the sense that just as the more popular a story gets on the most viewed rank, it gets moved up on the homepage three hours later. We see the same thing when we reversed analysis. So, the more popular a story gets on the homepage, it also gets moved up on the most viewed rank. And we thought that was really interesting.

And so, the implication of our findings suggest that journalists actually pay more attention [and] were more responsive to audiences than the other way around. So, this leads to the question of, does this suggest a new kind of agenda-setting functions that editors play in the age of data? And does this suggest there's a rise of audience power because of the findings that we found in the study?

So, this study is based on three papers in New York. And this is also the first qualitative study that looks at the time-lagged effect. So, we would like to know if this is actually idiosyncratic or if it's universal. So, if anyone in the audience or in your organizations are interested in understanding the future of journalism from this relationship between audiences and editorial judgments, please send me an email or just send me a tweet, and I'll be real interested in collaboration.

So, going back to the question of the overall finding that was the negative coefficient, why is [it] that for some newspapers and for some stories when they get more popular on the audience rank, it actually gets pushed down on the homepage? We don't really know the answer, and we really invite people to maybe give us your thoughts on this, but one possibility is that online speed is important. It's all about immediacy. So as the day goes on, no matter how important a story is, we have to make room for fresher content to come into the page and so naturally it's likely that stories get pushed down. But at the same time, that doesn't explain why for the New York Times, for example, we see the positive symbiotic relationships. And this is where we believe that qualitative studies can come in to really help us understand it. Or, if any journalist in the audience would like to talk to me afterwards and explain to me why you think this may be or if you see the same kind of findings in your newspapers, I think that'll be really interesting.

And that's the presentation. Thank you.

[Applause.]

Alfred Hermida: So today, I'm not going to actually talk about my book. I'm going to talk about a different project, because I think I'll just leave the talking to other people about that stuff. This is a project done with Seth Lewis and also Rodrigo Zamith, who is in the audience today. And we wanted to find out what was happening on Twitter in terms of how Andy Carvin was reporting the uprisings in Tunisia and Egypt. You know, we know that Carvin emerged as this key broker of information during this period.

And the aim of our research was to find out how sourcing is evolving in this sort of network distributed, very communicated and connected social sphere, and to see whether…. We talk about social media allowing perhaps more opportunities for individuals to be heard, to take part in the filtering, in the distribution and the interpretation of news. To what extent was this happening here? And we looked at sourcing, because sourcing matters.

Normally when journalists quote people, this is who they quote. They quote people—lovely, shiny, happy politicians, because they have institutional power. And sourcing is important, because who we talk to as journalists affects not just what we report, but also the meaning we ascribe to those events. You know, they shape public perception and understanding. We know from the literature, we rely on elite sources who have institutional power.

Then when it comes looking at how protests are reported, these tend to be the people we rely on as our sources—the ones with institutional authority. We will go to the police to see what is happening. Has there been any violence? Who's involved? How many arrests? That kind of reporting. When journalists cite non-elite sources, alternative voices, they tend to be not just in the minority, but often particularly when it comes to protest reporting, we treat them as deviant, as the other. So, we see that the powerful and the privileged dominate sourcing, dominate our perception of the world.

And we looked at Carvin to see, well, he was doing a very different type of reporting during this period. You know, he was sending out messages on Twitter, retweeting mentions for 16-18 hours a day, seven days a week. Hundreds of tweets were flowing through his stream. And the Columbia Journalism Review at the time described him as this 'must-read newswire.' So, we wanted to find out, well, if he is a must-read newswire, who are we reading? For this, we took a dataset of all his tweets. We then reduced it a couple of key periods: the week leading up to the resignation of President Zine El Abidine Ben Ali, Tunisia, who also was the president who expelled me from there in 1994, [laughter], and after lots of data wrangling—and we're doing a separate paper on our methods, because it was very complicated—we ended up with 162 sources. We did the same thing for Egypt during the week leading up to the resignation of Mubarak, who didn't expel me when I was a correspondent in Egypt, [laughter], and we had 185 sources.

We then coded them into these sort of categories, and this is based on a study by Lotan et al of the Arab Spring, and we used very much their methodology. Full details in the paper as well. But essentially, we wanted to break it down into various groups: media, institutional elites, and then what we might call alternative voices, either people involved in the protest, people who are identified as activists, people connected with organizations. That's broadly that.

So, bear with me. The results are kind of complicated. There's a lot more in the paper, and this is simplifying it to the basic bones. But when you look at Tunisia in terms of the source type, who are the types of sources in Andy Carvin's feed during this key period in the uprising? Well, what you see is that alternative voices are 23%. There are lot of institutional elites, but that figure, what that figure hides is that a lot of that institutional elites are what we classed as digeraties—people involved in social media and digital media. And what was happening at this time is Carvin was having a lot of discussions with people involved in social media about what was happening in Tunisia and his style of reporting.

But this is just by source type, and it doesn't quite give you the full picture. Because then we did the same thing, but by frequency. Who did Andy Carvin site and how often did he cite them? And when we look at frequency, we get a very different picture. Again, we see then that alternative voices make up

30%. We still have the institutional elites, because he spent a lot of time talking to other digerati there. But we start seeing a trend here. We start seeing that he's privileging these alternative voices. That they're more prominent than the percentage they make in terms of sources themselves.

We did the same thing for Egypt. And here, when you look at source type, this surprised us initially. We find that mainstream media is a large chunk of the sources that Andy Carvin cited in this key period of the Egyptian uprising. What's more remarkable, out of that 39%, 30% of that are journalists. And we can sort of understand that Egypt was pretty well covered [and] came after Tunisia. There were a lot of journalists on the ground. So, in terms of source type, journalists seem to have a great deal of influence. But again, when we look at frequency, we get a very different picture.

And this is what happens when you look and see who was cited the most times by Andy Carvin. A very different picture emerges. What we then see is that these alternative voices enjoyed an outside influence in this feed. In terms of source type, they made about a quarter, but in terms of how often he quoted them, they make about half of his sample. Institutional elites, far lower. Political elites within that, almost nonexistent. So, what we see here is that Andy Carvin was clearly privileging alternative voices and they had an outside influence in his feed.

So, in trying to make sense of this, I think one thing we see is sort of a reversal of the traditional patterns of sourcing; particularly, when it comes to sourcing protests. And instead of privileging institutional elites, he's privileging the alternative voices. And, you know, rather than sort of discrediting them, he's amplifying those alternative voices. I think what we potentially see here is the ability through a social media like Twitter, which is an open form of media connected and distributed, that you have an ability to bring in a broader range of sources. And they're sort of real-time reporting where you're trying to tap into what is happening at that time on the ground. You can bring in a broader range of voices. And this is what seems to have happened in Andy Carvin's. He was turning to people on the street, people who were documenting what was happening at Tahrir Square, who became activists by taking part in those events and emerged as prominent sources in the network.

The question it raises, in terms of when we think about journalism norms and ethics, is, you know, journalism norms and ethics, balance, subjectivity, fairness, in this case, outside these alternative voices that are normally outside of mainstream journalism had an out-sized influence. So, balance may be an issue here. What I think is also interesting when you look at Carvin is the ability of the information to go beyond his actual feed. Given the amount of journalists who followed what he was doing, how other people followed his feed to get a sense of what was happening, we have this idea of potentially him influencing other people's reporting by citing these alternative

voices. And, you know, we didn't analyze the content of his messages, but essentially he was promoting rebel voices.

And I think then that raises questions in terms of, how far will this reshape the way we see the news? Because we know that sourcing is a vital component of journalism; that who we talk to shapes what we report and how we report it. So, it raised for us a research question in terms of, how far does this reshape the narrative coming out of the streets of Tunisia, coming out of Tahrir Square? At what point did Mubarak switch from being a trusted ally of the U.S. to a dictator that we must overthrow? It raises those kind of questions.

To conclude, when journalists adopt new forms of technology, communication tools, they tend to go through a process of normalization, fitting them within existing norms and practices. Carvin didn't do that. He actually did quite the opposite. He used Twitter in a very new way when it came to sourcing. He overturned the sourcing paradigm that dominates most of journalism. And I think it speaks to this role that we've hinted at—the role of the journalist as curator, who can synthesize events in real time, pulling the best of the back channel in social media. This style of real-time collaborative reporting involving a whole range of voices, where the journalist plays this central node in a distributed network. That's a node that is trusted to authenticate, to interpret, to contextualize the information on social awareness streams. And he does this by drawing on the distributed knowledge in that network.

So, thank you for your time. There's an awful lot more here. And I look forward to your questions.

[Applause.]

Mark Coddington: My name is Mark Coddington. I'm a grad student here at the University of Texas. This is a paper I did with Avery Holton and Homero Gil de Zuniga, both here at the University of Texas. Avery is actually the lead author and put together this PowerPoint, but wasn't able to be here today. So, I'm pinch hitting for him. And what we wanted to look at was people who create content and people who consume it. There's been a lot of study and talk both in academia and in the newsroom about news creators. And we've known them by a variety of different names. We've known them as user-generated content. That term was big for a while. Citizen journalist, participatory journalist, interactive journalist—all these different sort of terms for what's essentially the same basic concept, which is people participating in the creation of content; particularly, news.

But we also know that there are relative few compared to the people who are consuming news. Research has shown that it's a kind of niche group that's actually participating in the creation of the news. And we heard more about this yesterday as well from people in the newsroom. We also know that they

are largely reactionary. They are responding to content that traditional professional news organizations have already made, rather than truly creating content themselves originally with some exceptions. We know that they are also, partly for that reason, co-opted by news organizations often. When they work with that newsroom process, news organizations often say, "If you want to create content, you need to do it within our standards, our routines, our rules." And then it ends up being, in some views, sort of a free labor sort of process for the newsroom. That's often how it works. We also know that they are valued increasingly just for the data that they provide, the personal data about themselves that we can package and sell to advertisers. Think about Facebook and its walled garden. It's incredibly valuable with its initial public offering coming soon because of all the data that we're providing it through our creation of content.

But at the same time, we also know that they are influential. We've seen examples all through the last day-and-a-half of people coming from outside professional journalism and working with people in professional journalism to really help reshape how we think about how the news is created. As far as how their content is actually perceived, studies have shown that their content is generally perceived as less important, less trustworthy than professional content. But at the same time, when we isolate news content specifically, we've seen in a few studies that they are seen — citizen journalism is seen as valuable; though, maybe not quite seen as valuable as professional content, but people are still seeing some value there.

So given all this, we wanted to — we had a couple hypotheses and wanted to ask one question. The hypotheses were pretty straightforward. We figured that people who were creating content and creating news-oriented content would have more positive attitudes towards citizen journalism, and we figured that people who were consuming this citizen journalism content would have more positive attitudes towards citizen journalism. Makes sense. If you're creating it or consuming it, you probably like it. The question that we wanted to know was, what sort of association would there be between news consumers as a whole—people who are just generally consuming news whether online or off—and their attitudes about citizen journalism? The other set of questions we wanted to ask related to the values of professional journalism. This idea of what is good journalism? And what separates that from just, you know, people writing about stuff or tweeting about stuff?

So historically, professional journalists have said that what they see as good journalism is accuracy, autonomy—this would be the divide between news and advertising—objectivity, and sort of a watchdog role for the public. The public's tenets have overlapped somewhat, but they've been a little bit different in the past, especially over the past several years. They viewed it from more of a populous perspective, a sort of public, civic journalism perspective. They say, "Good journalism gives voice to the people. It interprets news." And they also think objectivity is a big deal, but they believe that journalists are failing miserably at it.

So, from that, we had a couple more hypotheses and a question. We believed that people who are creating this news content, your citizen journalist or users generating content, are going to affirm these more — be more likely to affirm these professional tenets of good journalism. We also believe that people who are generally consuming lots of news are also going to affirm these tenets. And we wondered, what about the people who are particularly consuming citizen journalism? Are they absorbing these values of professional journalism?

So, the way we wanted to find this out was through a survey. We did a national representative survey. 312 adults. We used these key variables that we've been talking about, all indexes, multiple items on all of these things. We controlled for all the demographics. We also controlled for political efficacy and trust, because we know that those play a big role and we specifically want to look — we wanted to isolate the key variables on the top that we wanted to look at. So, we controlled for them.

And this is what we found. Yeah, warning — here be tables. [laughter] So, don't worry. I've circled the good parts for you. And the asterisks mean there's something going on there statistically. There's something significant. So, on the left and on the right, what we see are the people who are creating journalism, these citizen journalists, and the people who are consuming journalism, the citizen journalists. And they have a positive…. They are related to the positive attitudes towards citizen journalism, but they're also related to a set that we had termed as the negative attitudes towards citizen journalism, which is citizen journalism is biased, citizen journalism is opinionated, which kind of leads us to believe that maybe that's not necessarily a negative thing in people's minds. That's just the way it is. But they are not necessarily affirming the tenets of good journalism. But on the other hand, the people in the middle are the people who are just general news consumers. We didn't find anything as far as either way, as far as their attitudes towards citizen journalism, but we did find that they were really affirming those tenets of good journalism.

So now this, I believe, is the first hierarchical regression of the conference, so I'm really sorry, but I'll go through this really quickly. [laughter] So, we have positive attitudes towards citizen journalism [on the left], our negative attitudes in the middle, and then the affirmation of the professional tenets of good journalism on the right. And again, if you look at the circled areas here, on the bottom, we have the only people who are positive towards these — positive attitudes towards citizen journalism are the people who are consuming citizen journalism. Not even the people who are creating it. And this is after we've isolated everything else, so these are our most robust numbers. And then as far as the professional tenets of good journalism, the people who are creating news are not identifying with it either. In fact, there's a negative relationship; though, it's not significant. Only the people

who are consuming that news identify with that, and they identify it to a *huge* degree.

So, what does this mean? Well, first of all, we've been thinking along the lines of a distinction between content creation and content consumption. And this indicates that there may be some sort of distinction that's a bit more important that's between news consumers in general and news consumers of citizen journalism. So, the type of consumption may be a little bit more of an intentional act. So, those who consume citizen journalism may have the best attitudes towards it, but they do not necessarily identify with the professional tenets of good journalism. On the other hand, those who consume the news and generally identify with those professional attitudes or those professional values of the journalism that they're consuming, they're sort of absorbing those values, but at the same time, they do not have any affirmation for citizen journalism.

So, where does that leave those who are creating content? Well, we would expect the line between creating content and really identifying with these values to be extremely strong, and it turns out it was way weaker than the one for consuming. So, that could mean that those who are creating content might just create. It may be so embedded into just everyday activity that it's not really seen as journalism to them. It's just, you know, "I pass on a news story. I post a photo of an event I'm at that I find interesting. I comment on a story that interests me." It's not really seen as journalism. That could be a good thing, because it helps take content creation out of that niche that it's been in and makes it more sort of, as Alfred has said in past papers, ambient journalism consumption. We may be seeing sort of an ambient journalism creation. On the other hand, that's not really being connected to the values of journalism, and that might be a disconnect that educators and newsrooms might really seek to make. So, that's something that we could potentially take away from this study.

So, thank you very much. Look forward to your questions.

[Applause.]

Emily Metzgar: My name is Emily Metzgar. I'm on faculty at the School of Journalism at Indiana University. My co-author on this paper is Hans Ibold, who's also on faculty with me at IU. He is unable to join us today, but he's here in spirit and maybe even online. [chuckles] Hey, Han. [laughter]

All right. So, our research project is titled *Asserting "truth" in political debates.* The preceding papers have set the table perfectly for what we're talking about here. We're talking about Twitter. And you cannot go to a journalism conference or a social science related conference of any kind without running into, stumbling over Twitter in one way or another. And we are guilty of that ourselves, but it's actually a very interesting form of communication, right?

And one of the questions that comes up sort of inherent in these discussions is—and it's something that Mark touched on—is it journalism? Is what's being created journalism? Do we consider it journalism? Does it matter if we consider it journalism? What do we know about Twitter generally? One, we know that it's somehow influencing the communication ecosystem. We know that many people who are tweeting are performing some of the same behaviors that traditional journalists had traditionally performed. We know that citizens can now connect on their own and organize. This is ala Clay Shirky—the power to organize without organization. You can be out there freelancing on your own. And we also know that, with deference to Jay Rosen, we know that the former audience is creating and producing. We know that this is disruptive inherently in some way. And so what we want to do is talk about how Twitter fits into this media ecosystem.

So, the big picture. The view from 30,000 feet, with respect to providing context for the research that we're doing here is, what do we know? We know that Twitter is growing in popularity. We know that it's increasingly used for political discourse. And we know that it's increasingly possible to analyze the content that we see on Twitter, whether in a qualitative or a quantitative fashion. We know that it can be analyzed—content analysis, textual analysis, individual tweets. The Andy Carvin study is a classic example of being able to look at particular issues.

What do we want to know? What Hans and I are asking is, how are journalistic behaviors just on their own manifesting on Twitter? And how is political rhetoric being used? Our method for getting there is pairing with colleagues at the School of Informatics at Indiana University. There is a project called The Truthy Project funded in part by the National Science Foundation. And they are essentially collecting massive amounts of data. They've got the big data on Twitter on certain categories of hashtags that they've been following, and we are mining that data for our purposes here. So, what we've done is put together a mixed methods approach, where we're sampling data from their massive database and then doing hand-coded analysis of the content to try to tie it into some of these broader themes with respect to journalism, mass communication, political communication, [and] everything else.

So, the basic question that we're looking at is, if Twitter is becoming a powerful new form for journalism or a powerful new form for storytelling with a purpose, as Kovach and Rosenstiel have described journalism, then what we want to know is, to what extent are Twitter users engaging in this kind of behavior? How are they making use of this new technology that lets the former audience be part of the game?

We're trying to pull together four major threads of literature — five major threads of literature. And it's not an easy task. First, of course, the question of Twitter in context. Where does it fit overall? The new State of the News

Media Report of 2012 has a great analysis of where Twitter fits in the media ecosystem. We're also interested in looking at user-generated content. What kind of user-generated content do we see in the Twitter-sphere? We're interested in the Internet and politics; the intersection of it. I was at a conference recently where someone said, "We really need to stop treating the Internet as something separate and need to start talking about it as 'the mother of all intervening variables.' And that's part of what we're trying to do here.

Also, issues of media credibility. How do we know what we trust? Where does trust come from? How do we get other people to trust us? And finally, issues of media literacy. How do we make sense of the massive amounts of information that are coming our way? And how do people who are using Twitter make use of that information?

So, we first look at types of journalistic behaviors. We borrow from Kovach and Rosenstiel. And I think we should get a finder's fee or a commission of some kind every time we plug this book, because we're big fans of *Blur*. And one of the things that they have done in that book is identify four types of journalism that persist today. We have verification, which is the traditional journalism of objectivity. Assertion: the statement of facts, the journalist sort of as the conduit for information without arguing the point. Journalism of affirmation, where people go to particular sources of information to have their beliefs affirmed, not to be challenged, but to be affirmed. And special interest, which is, as the title suggest, is content designed to look like journalism, but in fact to promote a particular point of view, political orientation. Or for the purposes of our coding, we then came up with the all purpose 'none of the above.'

We also try to incorporate types of political rhetoric as designated by original research by Benoit and later by Wicks and others, who classify political discussion into three different categories: attacks, rebuttals, acclaims. And we added 'none of these' for our coding.

So, in terms of the actual work that we're doing, we expand on work from our colleagues from the School of Informatics at IU, Mike Conover and others, who have published work looking at an analysis of the Twitter-sphere based on division of political partisanship, broken down on particular hashtags, looking, coding different — using machine learning to analyze all the masses of data that are coming their way. We wanted to build off of that research.

So, we are using data generated by Truthy. And it's Truthy.Indiana.edu. I recommend it. Check it out. It's an incredible resource. I'm not on the grant, but I do recommend you checking it out. What we do is we identify two particular hashtags that have already through the previous research been coded as associated with political rhetoric. We take tcot, which is associated with the political right in the United States—True Conservatives on Twitter is

what that stands for—and p2, which is Progressives 2.0. This is a snapshot of the system architecture. I am not a computer scientist. This is where the data comes from, and then a miracle occurs, [laughter], and we get the Excel spreadsheet with all the data in it.

So, this is a snapshot of the two hashtag communities that we are targeting. The left it tcot. The right is p2. Now as you can see, anyone who's into network analysis, social networking theory can help interpret these. What you can see is that they're not identical. With p2, you've got sort of two different groups going on. On tcot, it's much more centralized. These are characteristics that have been described at length in the work by Conover and others.

What we are asking is: 1) To what extent do Twitter users produce content that is consistent with the categories of journalism identified by Kovach and Rosenstiel? 2) What are the characteristics of the tweets that are associated with each function? 3) What are differences, if any, between the way that tweets that are coded left or right, in the way that these behaviors are exercised?

So, what we're presenting here today is the results of a pilot study. Sort of our first attempt to dig in with the data and work with our colleagues. We ended up pulling 250 politically oriented tweets, many of which were categorized for partisanship, and the data was pulled from early 2012.

Very preliminary findings. Tweets consisting of retweets — entirely of retweets are much more likely to be associated with the left. Regardless of political alignment, tweets that are fully retweeted tend to be associated with scandals, highly powered rhetoric. There's a general disregard for verification. And links to the outside, when there are embedded links, because we did analyze the embedded links as well, links to the outside, when offered, do tend to assist with verification, to the extent there is verification at all. So, assertion, the journalism of assertion is what we see most frequently in these tweets that we've looked at so far.

With respect to the embedded links, we've got 'none of the above' as the most common source. [laughter] With respect to political rhetoric, attack, perhaps not surprisingly, is the most common form of political rhetoric.

And with respect to next steps, we are in the process of collating the 2,500 tweets that we are going to be manually coding. And we're heading back to Bloomington to do that. We are refining our coding mechanism and would welcome input on helping refine these and perhaps eliminate some of the 'none of the above.' And we're looking at additional categories for describing the individual tweet user, such as, date of account creation, total number of tweets, total number of followers, total number of retweets, that sort of thing.

In short, this is a pilot study. We are excited about what we've found so far, and we're eager for your feedback. So, thank you.

[Applause.]

Vittoria Sacco: OK. Hi.

Audience: Hi.

Vittoria Sacco: I want to first apologize for my accent and my English. It's a pleasure to be here today. I will present our study on media content curation. So, news coverage by way of media content curation is a recently new way of creating news. It's characteristic, potential, and impact on news curation and media have not been explored in science in sufficient manner yet.

On the last decade, online journalism has created new forms and new genres, but it has also contributed to shaping new phenomena by which the audience gets more involved in the news curation distribution; changing thus the role of media organization. But this new technology and social media presents some limits. The information provided is overwhelming and cannot reach everyone. News contribution vanishes from readers' screen. Several sources talk about the same event, but from different perspective. And social media often lack a clear storyline.

Several researchers have provided idea and concept for a new media ecosystem involving intermediary role of journalists. In this context, Bruns has suggested that gatewatching will replace traditional gatekeeping journalistic rule. On the one hand, gatekeeping refers to the traditional role of journalists, the select and narrate event. It depends on factors such as timeframe, concurrency, and unpredictability. In essence, gatekeeping is the practice of deciding why one story is selected to be reported and the other is not. On the other end, the core characteristics of gatewatching concepts is the participatory citizen journalist is part of the news curation process and journalists watch the gate and point them out to the reader, rather than feature information and sources themselves.

Enabled by the ever evolving Internet, new forms of citizen and participatory journalism have been appearing constantly. One most recent form of news curation is media content curation, which include social media and traditional media content.

From a structural point of view, main component of curated stories are the original contribution curated from social media and/or traditional media. The contents and background information provided by the curators and additional metadata. Key aspect of media content curation based on gatewatching are: 1) all submitted stories are published instantly, 2) where editorial decisions

are made, they are entirely transparent to users; 3) news stories and the entire website system itself are freely redistributable.

In September 2010, the American startup, Storify, launched a platform supporting media content curation, whose aim is to help journalists and the audience to collect and feature information for producing stories. So, this is an example of a Storify story. It can be seen as a composition of several social media and traditional media. Moreover, the curators can also add text to the story and publish it on its own blog or content management system. The meaning of Twitter video relating to this has highlighted the importance of social media on source information, but have also revealed the effort for extracting the best content in real time. Indeed, Storify has been used by media professionals and amateurs to help feature and report news about the media is surprising.

The goal of the empirical research has been to explore to what extent social media curation can be considered as gatewatching journalism. Through analysis of sources and stories featured by media professional and amateurs, the control of the flow of the communication and news reported can be evaluated. Representative does the interrogation of social media and traditional media.

So the following research questions have been examined: 1) Which are the original contributions employed in social media curation? 2) What are the type of digital sources commonly adopting social media curation? 3) Is social media curation affecting the level of gates?

The content analysis, original contribution, digital sources, authorship, and time coverage can be explored, as it is a method that allows investigation of content information; in particular, a news article. A story has been considered as a unit of content. Since media curation is an aggregation of social media and traditional media content, several sources could be found in one particular story. So, if one of the source categories was present, the coder has coded one; otherwise, it is coded zero.

So, at the end, we have analyzed 1,450 stories from the beginning of January to the end of August when this study was performed. And we have searched them with some key words in Google. So, this is only the coding process and the entire coding ability exists for internal validity of the results.

So, the content analysis revealed that 52% of the stories are written by media professional and 48% by amateurs. Media content curation has all the potential allowing the audience to be involved in both the curation and reception of news. For stories written by media professionals, journalists choose what to keep or omit, so that the professional remind the guardian of what content is to be distributed.

42% of the sources used to craft stories were few hours old. This shows that original contribution in social media are curated almost in real time. 28% of the story cover a daily coverage and the remaining 30% cover several days. The amateurs seem to prefer reporting already news coverage giving the most important information. Contrary wise and maybe more traditionally, media professionals choose a multi-day coverage summarizing the facts and giving a deeper insight into the event.

For tomorrow, as it is shown in this graph, the analysis reveal that the major source of information are media organizations, citizens, non-Arab people, Arab people, storyteller themselves, NGO, unknown sources, and official sources. As Bruns has argued with respect to gatewatching, this confirms that also media content curation as a gathering process become more transparent. As the reader, I'm more likely to quote original sources. The verification of sources if left to the reader, of course, require links and a social media account, but it can be also motivated by the comment of the curator.

So, these are some results about the digital sources. In the area of coverage, collection and aggregation of tweet focusing on the ongoing action. Overall 47% of tweets are using the already covered. Instead, in the daily and several day coverage, photo and video are predominant.

As a conclusion, media content curation consists in several sources from social and traditional media, while at the same time posting one's own sources as the primary source information.

So to summarize, in gatekeeping, at the input stage, journalists themselves choose new stories to be reported. At the output stage, editors select from journalist's material; story to be reported and published. And at the response stage, a number of audience resources are selected to be compared in today's paper or in the on-air broadcast.

Instead, based on Bruns' definition and given example, many attributes of gatewatching are disclosure and access to original source information, openness to our user to check input information, and participation of user in all the stages of news production. At the current stage, both amateur and media professional Storify stories present the primary gatewatching characteristics. News sources open to all users. Our repeated findings compared that curated story at gatewatching in the first stage and in the second stage if written by amateurs.

It was not the case at the time of this study, but now media content curation platforms support the discussion and comment open to all the user, so they have evolved to the first stage of gatewatching. So some media curation can be a first attempt to combine aperture of gatewatching, which supplement automatic aggregation with human generated content. In addition to professionals with some media search and filtering of variable sources,

traditional journalistic skills are still necessary in order to glue the curated piece of information to a story. Curators have to be trained in their assessment of a story and the curation of information. Media professional using this tool can benefit from their expertise and organizational research for just making a significant contribution.

So, thank you.

[Applause.]

Question & Answer Session:

Gabriela Warkentin: I know that we have someone here from El Pais, but I was going to ask you, Angela, when I read your paper and when I went through the study and what you presented, there was an interesting case in El Pais recently, where a story that was essentially a quite old story—I don't know, four or five years old or so—ended up in the last couple of months being the most read story, the most viewed story, again, on the homepage. And it led to El Pais even to disclose and say, "OK, this is an old story." And the person from that newspaper, she even wrote a paper on that and said, "This is something very strange that sometimes happens." A very old story that somebody started tweeting it or Facebooking it and then it just comes up again as the most viewed." So, when I was reading your paper and thinking about influence, I think that the path of influence is something that we still have to understand better.

Angela Lee: Yes. And I guess I have a confession to make. I've never been a journalist myself, so a lot of what I understand [is] actually through research. And so maybe this question would be better answered by the audience, but I do believe that there is a circular causal relationship between audiences, the most viewed, and editorial judgment that we don't really look at yet. And that, I believe, has a real future if we spend the time and resources to look at the research. And that's my confession.

[Laughter.]

Gabriela Warkentin: That's a good confession.

Alfred Hermida: I'll just give you an anecdote. When that happened at the BBC, we had a story, Man Marries Goat, and it was several years old, and then it became one of these most read for the best part of a week. And of course, our editors there then go to their weekly meeting with TV and radio colleagues, and they're comparing what's popular on the sites, and for several days he had to say, "Well, the most popular story is the Man Marries Goat from three or four years ago." [laughter]

Gabriela Warkentin: And then also, Alfred, it came out with Vittoria's paper, you're talking about a mix of sources, which is quite interesting. I

mean, you were just like disclosing the mix of sources from Egypt, in this case, and how the alternative voices sort of like are important there. And you're also talking a little bit about the mix of sources, and I think that that's a big issue that we have to consider. Yesterday, while we were talking about Twitter and the tornado or what happened in Memphis, there was something interesting that they were talking about, how do you filter emotion so that those twitters you're using are actually the ones that you have to use? And I think that it really poses a very important question about, how does that mix of sources work? What would be a better or a worse mix of sources? I'm sure that we don't have the answers yet, but I'm sure that you also have those questions.

Alfred Hermida: Well, I think this is one of the reasons we did this study— to see who's voices are being heard. And what we want to move on next is look at framing [and] how the various messages were being framed. But I think one thing that studies looking into tweets coming out of Tunisia and Egypt have found—Cece Papa Therese has done some work on this—looking at the tweets mix of facts, opinion, and emotion to such an extent that it's almost really difficult to separate the way we would separate it in traditional journalistic practices, in terms of this is a sort of factual piece, here's an opinion piece, here's an emotional piece. And I do wonder whether when you have this kind of mix of fact, opinion, and emotion, what kind of perception that has in the audience in terms of their interpretation of an event. And if you have people who are caught up in the event in Tahrir Square sort of reporting on what's happening around them, how that then shapes not just our perception, but our sympathies for what's happening there.

Gabriela Warkentin: Yeah. Because you were also saying the way that they use some tweets that are also rebel voices. I mean, maybe you are just charging it or inclining bands to some part of the story.

Alfred Hermida: And I think that's what happened here. That even though, particularly in Egypt, a quarter were these alternative voices, they made up half of his feed at this very critical time, so they have an outsize influence. I don't say that's a good or bad thing, but it's something we have to be aware of, and it certainly is contrary to the way traditional journalists would do. If I was at the BBC when I was in the Middle East and did a story where half of the time I was just quoting protestors and virtually ignoring any government officials, my editor would have some questions for me.

Gabriela Warkentin: [chuckles] Yeah, sure. We have a question over here.

Woman: Yeah. This is really the right time for this question, too, because it concerns your paper and your use of the word privileged and outside influence—outsized influence. I wonder if the basis for those assessments, that is just your typology of the types of sources, is really the best way to look at that, because it does seem to me—and I'm happy to hear you're going to do a framing analysis—it does seem to me that a better way of

looking at privileging and outsized influence might have to do with the freshness of the content. Carvin has said repeatedly that he had to use sources all around these various countries in the Middle East because the mainstream media were holed up in just a few places. And in order to get the news, in order to get the fresh content, he needed a wide variety of people from all over the country. Moreover, a lot of the process by which he was gathering data involved an elaborate process of confirmation, which means he needed those tweets both to source his content and to confirm.

Alfred Hermida: And I think certainly our data suggests that that may have been the case in Tunisia, where there were very few journalists there and it sort of caught the Western media by surprise. But I think when you come to Egypt, you know, a third of his sources are journalists, so he does actually quote them. But when you look at frequency, when you look at how much the journalists make up his feed, you know, they make up far less. So, I don't necessarily....

Woman: Perhaps they're saying the same thing though over and over again.

Alfred Hermida: Well, perhaps that's also happening with the activists saying the same thing.

Woman: Yeah. My point is that you need a content.

Alfred Hermida: Oh, yeah. We didn't look at the content. In the paper, we say, you know, this raises a lot of questions. But, say, if we take it as a starting point, particularly the discussion we've had around these new forms of journalism, how does it in this particular sense affect sourcing? What kind of questions that then provokes in terms of looking at how an issue is framed [and] what kind of content gets circulated. But for this initial study, we simply wanted to look at the voices that were being heard.

Gabriela Warkentin: Well, thank you. And yes, as I said when we started the panel, I think that the papers we have here actually raise much more questions than they give answers, which I think is fascinating. That's why I really love those papers, because I'm sure that maybe sometime when you go back to what you said today, it's going to be quite different, because there are so many questions. For example, even through your presentation, you insisted on, we have to reconsider the way we are, for example, in universities or the way we are teaching, the way we are forming our students or forming our future journalists. And I think that is something that we have been saying. Those that are in academia, we have been saying that for a long time. And I'm not so sure if we're doing something about it, but I'm sure that we're saying it for a long time. And sometimes the answer is, OK, they need more courses. They need a lot more courses. And I'm not so sure that that is a good answer. I'm sure that maybe we have to rethink a lot of

stuff and you insisted on that. So, really would like you to have a little bit more in-depth on that.

Mark Coddington: Sure. And I think this is something that might come outside of journalism schools, in that the content curators that we're talking about are not really journalism students and they're not people who were journalism majors. They're just people, you know, who were sharing whatever they thought was interesting or commenting on it. And I think the way to kind of connect them with journalism values, not that they need to self-identify as journalists or anything like that, the label really doesn't matter, but more of like a media literacy sort of perspective. And I've seen several calls over the past several years to like make some sort of media literacy a required course or something like that. And many people have done that. Kovach and Rosenstiel and *Blur*. You know, that's a big part of just telling truth from BS, and those sorts of things. And so, if those values are going to be connected within education, I think that's probably the way they do it. Not necessarily in training our journalists better, but recognizing that everybody's going to be doing this. And so, if we all have an understanding of how to consume and create media, sort of with our minds critically engaged, that's a better thing for all of us.

Gabriela Warkentin: I don't know if somebody has a question. I have another one for Vittoria and because, I mean, what Mark just said, I mean, it's interesting to understand the content curation. We talked about that. It has, as you just showed, it doesn't have to do with the traditional values of news, of journalism. Are they different? What is your reading on that? I mean, on journalism values and the way we are producing content.

Vittoria Sacco: I mean, for what's about the media professional story, they can still use, you know, norms and principles of editing. And what is interesting is even in stories written by, you know, amateurs, they also quote media organizations, so it seems like they search for a curator and trustable information. So, they put this kind of emotion [in] using as witnesses tweets or other social media and then they quote the journalist just to give credit to the right sources.

Gabriela Warkentin: That brings me then to Emily, which I think was interesting. I would love to see the next step of the paper when you have much more data around it. But you were saying about something that is obvious when you see it, but it's important to say it, that a lot of the partisan tweetership and so on is various attack, it's emotion, and so on. But I also think that we have to consider, how do you read this kind of stuff in a general timeline? I mean, if I follow a lot of people or whatever, how can I understand what is, like, say, coming from these deliberate attacks? What is really partisanship? What is just an opinion by someone? I think the timeline, when we take the tweets out, then it's interesting to see them, but they're not within the context of a general timeline. So, how does that connect?

Emily Metzgar: You're right that a timeline is important. And in fact, for the data that we're getting ready to code, [we're] looking specifically at the first three months of 2012. And if you go to the Truthy site, you can see that there is actually a timeline available showing patterns of usage in both the p2 and the tcot. And it's possible to work backwards from the spikes, for example, to see what was happening in the news to provide context for the tweets that are emerging. That said, I think we are at this point focusing less on the affects of it than simply trying to describe what it looks like.

Alfred Hermida: If I can just ask you a question, because I wonder also how far your findings are affected by the type of topics that you're following. I'm peripherally involved in a study in Canada where the results will be out in the next few weeks, and they looked at various political issues; some very legislative political issues and then the other one was occupy. And I found a real difference in the discourse on Twitter when it came to sort of government policy versus a sort of street protest, I mean, that emerged from the streets. So, I wonder how far the nature of political discourse affects the type of voices and also what they say about it.

Emily Metzgar: I think that's a really good questions, and I think Mark touched on that a little bit in what you were presenting. The fact that different people are — that you may have people who are just creating content, who aren't necessarily participating in any other aspect of what's going on, and that puts a giant fly in the ointment of trying to assess where the tweets fit in the broader overall ecosystem.

Gabriela Warkentin: Well, thank you very much. This was really interesting. And thank you all for being here. Thanks.

[Applause.]