15th Annual International Symposium on Online Journalism

Day 2, April 5, 2014: Afternoon Session — 4:30-6:00 p.m.

Knowing Your Audience:
Readership Analytics & Editorial Strategies for Online News

Chair: Joshua Benton, Director at **Nieman Journalism Lab at Harvard University**

Panelists:
- **Todd Cunningham,** Director of Media Impact Project, the Norman Lear Center at **University of Southern California**
- **Melody Kramer,** Digital Strategist and Associate Editor at **NPR**
- **Higinio O. Maycotte,** CEO and Founder at **Umbel**
- **James G. Robinson,** Director for News Analytics at **The New York Times**
- **Nicholas White,** CEO and Founder at **The Daily Dot**

Todd Cunningham: Thank you. So thrilled to be here. Greetings from USC's Normal Lear Center. So, I'm here today basically to talk to you about some things that we have been working on for the past -- basically, I've been on the job about 126 business days. So, for anyone who counting, it's like that's not much time, but we've done a lot of work, and I can tell you it has not been alone. It's like it takes a village. It takes many of us. None of us can do any of this alone. And the whole premise of the Media Impact Project is really about collaborating and doing things together and learning together and making sure that we're sharing all the kind of things that we're doing, just as Melody is doing and showing her leadership there at NPR.

So, let's jump in. first, I thought I would just start by giving you a definition of the way we think about media impact, the way we define it. We define it as, "the role that media play in changing knowledge, attitudes, and behavior among individuals and communities, large and small, around the world." Pure and simple. Nothing more. No window dressing or anything else on that.

When I talk about our mission statement, that really is the crux of what I'm going to talk about today. Really, the Media Impact Project is here as a global hub for collecting, developing, and distributing the smartest approaches and the best tools for measuring the impact of media. I also oftentimes talk about it as being the place where we're here to basically understand how media motivates people to take social action. Same thing. Same idea.

So, we're going to jump into basically each of these three areas—collecting, developing, distributing—and give you some examples of the things we've been doing and what we have in store as we move forward.

So, first is collecting smart approaches. First off, we did a lit review, that if you want to email me, my email address is at the end of this presentation here. It was done by Phil Napoli, who now is at Rutgers. Our focus was really on getting organizations that are kind of culturally ready to actually do impact measurement. A lot of organizations are not. And in fact, they sometimes find themselves in a place where they've seen anything about measurement as being punitive or being something that is something to be shy about or embarrassed about. Because most of the people, of course, who are doing this work are not necessarily researchers. They don't come from an analyst background. So, we are expecting a lot out of them to be able to actually have them step up and to create a culture where it's going to take more than just that analyst to be able to change the way that the business works. It's really about behavior change.

Another thing we're doing…. And this is the crudest drawing I've ever dealt with before. I will have something much better in the future. We're building a data analytic center. In fact, it's already built. It's ready. We've already been taking in data. It will basically be comprised of two levels. One is like it will accept lots of third-party data, open-data sets that are the kind of data that is not sensitive, doesn't have any kind of restrictions around it. And then also we've built a safe data harbor. So, building this from the ground up with a safe data harbor at the center of it, really, with all the policies, and procedures, and safeguards, and everything to ensure that the people who are using [it], the right people actually get to see the right data.

We also will serve as a third party if there are opportunities where you have lots of different data that you don't necessarily want the groups to actually see, but you need someone in the middle to basically kind of air traffic control all that and do that analysis. And we also have the ability to have virtual machines, so that the data doesn't have to leave the companies [or] organizations where it lives. It will all be ingested and go through a whole process. Carl Kesselman, who runs this division, is much more conversant in this. But it's using a lot of USC's computational excellence and group at the ISI Institute to basically help us develop new tools, new dashboards, and new kind of approaches that will all be open source and available to everyone.

So, developing. We've spent a good deal of time on the development front. One development approach that we're actually working on [is] with Participant Media. Participant Media is an organization that has what they call a double bottom line. They have one bottom line, the typical kind of financial one that we all know about, and then they also have another one that's what they would call [the] 'doing well by doing good' bottom line. So, they produce movies like "An Inconvenient Truth," "Food Inc." Movies like that that are more pro-social in mindset. In fact, they are a great company to work with, in fact, to partner with and collaborate with, because they are in the very business of doing the thing that we're all trying to learn about. They've been for a long time talking publically about how they actually

motivate people to change or to take action from seeing the movies that they produce.

Now, they are ready to quantify it, and so that's what we're doing with them, in fact. Basically, what it's called is the participant index. Don't think index as in research index. It's really about a collection of media performance data to assess the combined results of one's actions taken across a number of different screens and different kinds of media content.

Now, how does that relate to news in journalism organizations? It's like, again, those films are news and journalism oriented. They are…. Some of them are documentaries, but many of them are not. We're looking at narrative films, documentaries, TV, narrative and reality, unscripted shows, as well as short online video. So, we will have a good deal of understanding about the entire video landscape, in particular as it relates -- as we heard earlier today, a lot more organizations are going towards video, so we will have a lot of that learning to share, of course.

Plus, we also are working with the internet archive and using all their closed caption data for all news programming. So, not only your six o'clock and evening news, but also shows like The Daily Show, 20/20, Good Morning America, all those kinds of things as well, to be able to understand what role those play in helping to motivate people to take action, make social change themselves.

Another one that we've actually been working on also is, we found that there are hundreds and hundreds of actions that one is challenged with having to think about, how do I measure? When you think about all the things that -- an action that a person could take after seeing some kind of content or after talking to another person, it could get quite mind numbing. And so, we set out, before I actually joined the project last year, the number…. You actually were at a meeting, a convening that they had that Media Impact Project hosted, and it was really around something that I think they called The Ings, which was inspired from folks at The New York Times and some other organizations, where they all basically contributed, dumped all the, basically, the actions that they were actually measuring. From that, we basically created The Verbs. And it's called The Verbs of Online Action. Basically, it's a way, a taxonomy, an organizing principle to think about, how do I -- if I need to actually measure a certain kind of action, I can actually go to the heading here and like find out.

There's a grid. I just wanted to give you a quick snapshot of what the headline is on this. But I could look, in fact, and see the kind of thing that I would actually be -- how it's described, what the activities are that are associated with it, what the means by which I should research it, what expectations I ought to have in it, have from it, what advantages and disadvantages from that approach would be. Just a really great at-a-glance snapshot for people who are not necessarily well steeped in understanding

how to do this kind of research and helping to provide great solutions for you. So, we are really excited about this product.

More development [of] other kind of smart approaches. So, the way that we are looking at all of this is basically—ooh, wow, I didn't see this happen—so, lots of data. We'll use the verbs. We'll then filter it through different cohorts, which a lot of times doesn't seem to happen. When I talk about cohorts, you can think of it typically as demo data. Really thinking about different segments within your audience. Take actions. We'll put it through the types of infrequency. Then an impact model. But the impact model will be influenced by other observations that really come from qualitative research and anecdotal research that many times gets brought in at the beginning of this process, but it gets buried. But oftentimes, that's the nectar of it all. That's the stuff that really influences and helps to give shape to your kind of solutions and, in fact, how you would actually measure the impact.

I'll go very fast. We're also working on another thing. We can take this for another day. But basically, we're working on -- there's a neighborhood within Twitter called Black Twitter. And it's people who, actually, they are not all African-American, and all African-Americans are not all involved with Black Twitter. But basically, we found that they galvanize and mobilize themselves around specific topics. Sometimes a lot of news and political kinds of issues. Well, it just so happens that they are also viewers of Scandal. We have Scandal viewing parties, and we've been monitoring all the Twitter data for Scandal. Because it's actually a show that is a perfect researchable moment for a news and journalism organization, because it's like it has all the things that you'd want. If you ever wanted to design a study to try to understand how news works with people, it's like you'd want it to be a place where there's lots of heavy audience involvement, where the content is engineered for social media activation. That would be the dream state, of course.

Third, it's like the audience's content is created and unfolds in response to real-life experiences just like the way that we respond when we see news. And finally, instead, not everyone has a clearly defined community that they work with, right? It's like sometimes it's fluctuating and it's kind of amoeba-like. So, it's a great collective of audiences that are kind of organized within Twitter. We've been studying it now since the beginning of this season currently. We have lots of learnings, and we will share those with you. You'll be seeing those in the future.

So, distributing is the last one. We're distributing lots of different tools. You can see across pretty much everyone, because of course media impact measurement and the measurement and understanding of impact is pervasive. We can't really get away from it, and mostly it's because media is pervasive. So, we're here dedicated to focusing all those.

What's next? This is one thing I just want to share with you. Typically, we feel like there is not a lot of understanding about the person who is in

charge, who is like on point for doing the analytics in a news organization. I think that we made lots of assumptions about what they know, what their life is like at work. We find a lot of times that if their supervisor or their senior manager or editor at their organization has not made this kind of work a priority, it's like they've put it to the side. It's on their to-do list. It's not something that they're going to make sure gets done. So, we want to basically shine a light, illuminate what the ark of the experience of the person who's actually doing that job is all about and bring light to that, and also talk to some senior managers also to understand what it is that we can do in terms of outreach for that group. Because we believe that they were not brought up in all this culture, right? They are many times reciting talking points and doing the right thing, but it's like when push comes to shove, we need to make sure that the actual work is getting done the right way.

We are hiring. We're hiring. Thank you very much.

[*Applause.*]

Melody Kramer: I'm Mel, and I'm a Digital Strategist at NPR. I work on a two-person social media team. In addition to running our shop's social accounts and strategy, we find and code tools for the newsroom to become more efficient, so that they have time to actually use social. And we create and find tools that help our reporters and producers prioritize their limited time.

Earlier this year, we started with this idea. We wanted to create a tool for NPR journalists to better understand our digital culture. If you're working at a new media startup, it's easy to say that analytics tools are part of your culture from the ground up, but it's a little harder at shops like my own, where there's already a culture in place.

We do use Chartbeat and Google Analytics, but people don't have time to sift through Google Analytics, and even if they did, Google Analytics is really confusing. Chartbeat doesn't say a whole lot about content that's not in the top traffic getters. That's interesting, but it doesn't really help folks learn. So, we had to think, if we create a tool without creating a culture that embraces that tool, does it matter if we have a tool in the first place?

So, we, and by *we*, I mean Wright Bryan, who's the other member of the NPR Social Team, and Scott Stroud, who's a user experienced designer, and myself set out to, number one, create a product that could help our newsroom, and number two, figure out how to collect, display, and share data that would help the newsroom use, benefit, and embrace the tool.

The most important part of the project was not the tool itself. It was to create what I call a sharing culture within NPR. We started doing this about five months before the Analytics Dashboard Project began. We knew that to have success with the dashboard, people would have to notice something,

take an action of some sort, like tweaking a headline or updating a story or sharing something on a social media platform, and then share it with the rest of our news organization so that we could all learn. So, we needed to start sharing stuff internally as a news organization.

So, I created a listserv. It started out as an internal listserv. All I did was praise people for doing good work on social media and on digital media. I also shared really cool stuff that other media organizations were doing, stuff that NPR could learn from. Everyone at NPR was opt-in except for management. I put all of them on, because I wanted them to know what everyone was up to, and I wanted everyone to know that their managers and bosses valued this stuff. The listserv started to make people really happy. They liked being mentioned and they sharing more stuff with me. In turn, I became smarter in the process, because people were involving me earlier in the reporting and sharing stuff they spotted online that could benefit NPR as a whole. It was a win for our entire organization.

There are now 450 people on the listserv. And last month, we decided to make it public. The complete archives is up on that website, and you can share it with your newsroom if you'd like. It's basically everything we're thinking about and doing and all of the tips and tricks we've learned in the social space. As well, there are Cliff Notes for this presentation and the full text of what I'm reading to you.

The next step was to think about the dashboard itself. I work in public radio and couldn't spend money researching this. So I called up The New York Times, BuzzFeed, The Atlantic, The Huffington Post, USA Today, and The Guardian, and asked if we could come and visit and talk about what they were up to in the analytics world. They all said, "Yes." So, Wright, Scott, and I went on a fieldtrip to New York and had very candid, open conversations where we exchanged what we were thinking and saw how they thought about analytics in their newsroom.

After coming back from the other newsrooms, we shared what we learned in a brown-bag lunch session at NPR, as well as on our internal and external listservs. So if you are curious about anything that we learned, you can look on that Tumblr. We talked about how we'd approach analytics in our newsroom. We also met with every show and desk head and asked individual members of each department for feedback. This kept everyone in the loop and invested in our process.

Then we started to think about what UX people call personas. I don't know if you're familiar with this term, but we came up with several. Some of them were, "I'm a homepage editor." "I'm an editor of a section." "I'm a blogger at NPR." "I'm a radio reporter." "I'm a producer." And then we sketched out what each of them might need from a dashboard, because each of those people might need something different.

It was really important to think about what our audience—which in this case was our own newsroom—might need. Because if you build a tool and the newsroom finds it difficult to understand, difficult to use, difficult to access, difficult because of your company culture, or difficult to share, then what's the point?

And I didn't mention this: when I pitched the project, I was given six weeks of development resources. That's not a lot of time. Most newsrooms have a fulltime person doing this all of the time. So, we had to think, how do we get the most impact for what we can build? What's the most bang for our buck with limited resources, staffing, and time? We don't have time to be everywhere, and to think that our reporters can add more stuff to their plates without having a better sense of what to prioritize doesn't realistically address my newsroom or other smaller newsrooms needs.

So, we set out to write our general assumptions about a tool, because as Scott pointed out, this would help us narrow our focus, because we only had six weeks to work. The first assumption had nothing to do with the tool itself. It was: A change in NPR's culture is equally important as building a useful tool. We want to empower individual journalists by providing them the tools and information they need. We believe our social media strategy should be centered around individual empowerment. We want to empower our journalists to build their own individual presences on social media. We want our beat reporters, producers, and editors to become established leaders within their top areas regardless of the platform they are using. And we want to provide our newsroom with the proper tools and training.

This may seem like common sense at a new media startup, but there are many more challenges at a non-startup and legacy organization. There's an existing culture in place, and there's a framework in place. So change can come, but it comes slowly, and usually involves, number one, getting in trouble, number two, not minding getting in trouble, and number three, meetings. I go to a lot of meetings. I also get in trouble a lot. [laughter] This is just part of the process, and you just have to deal with it.

The other assumptions that we made through the process: The tools we create should help us understand and expand the core audience. We assume that the tool must evolve over time, and we should start with the most impactful solution. And back to sharing. We want people to be able to share insights from the dashboard, and share somewhere that everyone at NPR can see in real time, so that we can take meaningful action with the limited amount of time that we have.

We also did a lot of thinking about ways to best integrate the product with our CMS and people's workflows. We could create a bookmarklet. We could link the product within our CMS. And we could think about the landing page specifically, because what users see on the landing page would determine what they thought about our product.

We found it was really helpful to assign everyone in the newsroom a role. And then we thought about what they would do if they noticed something on the dashboard. For example, let's say I'm a homepage editor, and I learn a piece from our archives is spiking from the dashboard. Then, should I put it back on the homepage? Update the page itself? Let the social team know? And so on. And then we made a chart of these things. If I have Job X and I learn Y, then I should do Z. This is a really good way to think. If I have a dashboard, how do I want my audience, and in this case my newsroom, to engage with it?

Our other thought was about sharing. Is there a way to build a dashboard so that if someone notices something they can share [it] with the rest of the newsroom? My colleague Wright Bryan came up with this idea, and I love it! Ideally, this involves the landing page I mentioned before. It would have images of unusual audience behaviors that people could see as soon as they accessed the tool, so they would always be learning and seeing what their coworkers did.

We're not there yet. That may not be built in the time that we have, because we only have six weeks. That's an ideal state that might be many months or even years away, but we have to be realistic about what we can do right now. What we're thinking about right now because we're in the middle of this project is, is there a way to learn from the tool out loud? Is there a way to build actions into the tool to prompt people in my newsroom into action and then share with their coworkers?

We're priming the newsroom to share what they learn in the tool even if the tool isn't in its ideal state at the end of six weeks. And we plan to share everything we learn with other news organizations we met and to keep the conversation ongoing.

Next, we're thinking back to sharing with the public. Is there a way to make part of this tool public so that the audience is also involved? Is there a dashboard that would make the experience more meaningful for our audience? Again, is there a way to learn from our tool out loud? I think we have to. We've been using agile development throughout this process, but analytics also requires agile thinking and being able to share and change and iterate and then share again.

So, if you work in analytics and you can help out in anyway, please let me know. I'm always up for sharing ideas and hearing yours. I would love to integrate the way you're thinking into what we're sharing and sharing online and sharing with our newsroom, and I'm happy to share with your newsrooms.

I'm mkramer on Twitter. I'm MelodyKramer on Github. All of the code that we write is accessible to everyone. We make everything public. I'm

mkramer@NPR.org. if you're in D.C., I'm a tour guide at NPR. I can give you a tour. [*laughter*] And I'm going to sit down.

[Applause.]

Higinio Maycotte: Well, thank you for that. It was one battlefield, [*laughter*], called The Texas Tribune, where a lot of this, you know, thinking started. I'm an addicted entrepreneur and I love solving really big problems. I didn't quite know what I was getting into when I decided to join the Texas Tribune to help them launch the site. That was simply the limit of my obligations. And I ended up there, I think, about another 18 months. And it was with that that I realized that despite our success and despite being very technologically focused, we literally had no idea who our audience was beyond some simple surveys that, embarrassingly, we would run maybe every year, and some very, you know, sort of questionable quantitative analytics, a lot self-reported Google analytics, Omniture, some services out there like Quantcast or comScore that thought they had an idea of what our traffic was like. But nobody could really tell us who our audience was at scale without, you know, the intervention of questions to interrupt our users or anything else. So, once I knew that Texas Tribune was going to be off to the races, I decided to move on to try to solve this problem.

And I think three years ago, this was somewhat of a new concept, but I think today we're all having a major data awakening. We are all realizing how much data we are all creating. You know, if you could actually see it, we're all evaporating data as we speak. Companies are generating massive amounts of information. And this sounds like a brochure, but I think the real reason they call it *big data* is simply because it's piling up.

And so, what do you do with all of this data? And maybe more importantly, who are the rights holders for this data? Like, who owns this data? Who should benefit from it? And I think in the media context, something went very wrong in the early days of the transition to digital. You know, we often criticize the traditional media companies for what they do, but they actually have storytellers beyond just editorial. They have sales teams who tell stories about their audience. The tools we still, I think, like to criticize, like the Nielson's and the Arbitron's, you know, those techniques may be hundreds of years old, but they tell a story about who the audience is and that is really valuable.

When we moved to digital, I think we outsourced everything to Google. And now there's thousands of companies that are fighting sort of for those incremental CPM improvements. And sadly, they know your audience better than you do, and that doesn't make much sense. You know, the relationship that matters is the one between you and your reader, the one between you and your consumer. And if you do it ethically and you respect the relationship, there's a lot of data there that can help you tell stories about the value of the audience.

So, while I totally respect the idea of journalism and content being king, your ability to articulate the value of that audience to a particular sponsor is also critically important. And I think it's simple to say, "Hey, you should get ahold of your data and figure it out," but how many in this room have data scientists, analysts, or storytellers devoted to telling data-driven stories to brands and agencies to make you more money? You know, it's crazy. We're full of some of the world's best storytellers, but we don't apply those same talents to the data.

But getting to that data -- and I think what the Media Impact Project is doing and others is hard. So, if you think of a modern media stack, it's full of all kinds of complex components, massive investments. You know, you have DMPs and you have ad servers. You have social channels trying to make sense of social ROI. It's really, really difficult.

How hard is it to get access to your developers? Probably pretty hard. How many of you think of yourselves as media technology companies and not just editorial, you know, endeavors? So, making sense of all this data, breaking the technology silos, and more importantly, breaking the organizational silos, is really, really hard.

And so, I think in the end, digital media companies are not taking advantage of this data asset. We're starting to work more and more with some of the world's biggest broadcasters. And, you know, taking advantage of these innovations is difficult. Showing the impact of their investment to a sponsor is very difficult, but it's something that everybody must do.

I would assert that in the next three to five years data will start to move from the intangible side of a balance sheet to the tangible side of a balance sheet. This is going to be a mandate from the top down in these companies to take ahold of this data, figure out what it's worth and how you can make it more valuable. You know, the CMOs today still get by with impressions and conversions, but who cares how many times that ad was seen? I think even today I hear constantly that size equals value, and all of these media companies out here are racing to get more uniques and more traffic. But what if you have the most interesting audience in the world and you can articulate that to a sponsor who's willing to pay money for that? You know, that might mitigate the idea that you have to grow a giant audience and might help focus where you really need to focus.

So, I think Umbel's vision and the opportunity for everyone in this room is to make data the most valuable asset that you own. And I say that right now from a pure media company standpoint, but this really applies to companies, and maybe at the end of the day if we do our job correctly, it applies to individuals. You know, I think we as individuals are in control of most of this data, but we're the furthest from being able to actually realize that control.

So, Umbel hopefully will someday help each and every one of us make more sense and understand, and not be confused about the value of our own data.

So, just in summary, stop focusing on ad networks. Figure out how to get those, you know, premium dollars. Devote some of your storytelling efforts to better monetizing your digital assets. I think eliminate company fragmentation. I love what NPR is doing. Make everyone in your company aware. Democratize the data. Help everyone understand the impact of, you know, the stories they are writing and the advertisers that they are bringing on.

And I think for the data rights subject, respect your users. Be transparent. Give them options. Let them opt out. Don't be embarrassed to ask them for more data, but be willing to let them opt out of it if they want to.

So hopefully. the internet will be a wonderful place for, you know, all of us in the future. And I think media companies have a chance to take back what brands and agencies took from them in this transition to digital. And I think data is sort of the key.

So, thank you.

[Applause.]

James Robinson: So, you guys must really like analytics if you're still here. Thank you for coming, and thank you for staying. On behalf of the whole panel, it's great to see you all. I don't know about you. Usually when I'm talking about analytics at 5:15 in the afternoon or 6:15 New York time, I'm usually doing it over a beer at the bar. So, I'm going to spare you a PowerPoint. I have like my talking points on like my business card, right? I'm going to try not to use jargon. And we'll just sort of sit back and pretend we're at the bar, and maybe we can continue the conversation afterwards at a real bar.

So, my name is James Robinson. I'm Director of News Analytics at The Times. Before that, I was Director of Web Analytics at The Times, for about five years doing analytics on the business side, and then about a year ago, April 1st of last year, I moved down into the newsroom to try to bring analytics into the newsroom of The New York Times for the first time and see how we could apply audience insights to editorial and newsroom decisions.

So, it's been a really interesting experience. I think in previous years, it's something that The Times has actively resisted or at least not been too enthusiastic about. I think in part because of the fear that…. It's just so weird looking up at you guys while I'm talking. And I think it's due in part to the fear that our homepage would be covered by like cat videos if we started looking at audience behavior. As I've sort of tried to show over the past year,

that's kind of a false premise. And in fact, we've actually got quite a deal of enthusiasm now amongst our journalists to figure this thing out.

And I'm going to share with you some things that we found along the way in the first year that hopefully will resonate with you and sort of allowed us to start this conversation with journalism editors about how we think about our audience.

So, the first thing that I just want to make clear is that there is a very important distinction to be made between reporting and analytics. They are two very, very different things. It's the difference between metrics and insights. And whenever I hear the term *metrics,* like, "Let's bring metrics into the newsroom," it like bristles the back of my neck, because it's not the point. We are trying not to bring numbers into the newsroom. We are trying to bring our audience into the newsroom. It's a very important distinction.

Numbers, metrics, reporting, in the analytical sense, is really all about evaluation. It's like keeping score. You know, it's like top ten lists, or how much, or counting big aggregates of things. "Wow, we got 30-million visits." That sort of thing. Whereas, analytics is really more about discovery. It's more about insights about our audience. Metrics can sort of be expressed in numbers. Insights can be expressed in ideas; usually surprising ones. And that's what we've chosen to focus [on] at The Times.

One of the questions that I've gotten most from fellow journalists has been…. What do you think it is? "How did my story do?" You heard that question? [*some laughter*] Right. Okay. And the only tools we have to answer that question are tools that were originally built for marketers and salespeople, which are designed to regurgitate big numbers back at you, right? And so, the only numbers we have at our disposal currently to answer that question are just big numbers. So, we'll give them a big number. And the way it breaks out at The Times is that, like, when you hear a number, it sort of equates to like your possible salary. Right, so, "So, how did my story do?" "Oh, 20,000 visits." And they're like, "20,000. I can't really live on 20,000," right? But if you say like, "150,000," they're like, "Oh, that's pretty good. I can live on that." Or like a million, like I'm a millionaire, right? That's how we think about it, but that's not really that satisfying.

And in fact, the follow-up question that we get often is, "Is that good?" And that's a really interesting question, because it really depends on what you're trying to do. And enunciating what we're trying to do is maybe not necessarily about scale. Maybe it's about something else, especially in the newsroom, because opposed to moving from the business side where everything is about the dollar bill, solving for the dollar bill at the end of the day.

Engagement is another thing that people like to measure. Usually outside the building people say, "How do you measure engagement?" And that's a really

baffling question for me, because that's kind of a reporting question. It's like trying to boil engagement down to one number. And to me, engagement is not something that can really be quantified in a single number, because it's like a state of mind. It's like a state of being. It's like a relationship you have. It's like love, right? It's like trying to figure out how much I love my wife by how many flowers I buy her each week, right? And I don't think you can really do that, right?

Engagement, to me, is the alignment between what you want your readers to do and what they want to do. And if those two things are in line, then you have engagement. And to boil that down to a single number, I think, is kind of impossible.

In the end, the insights are really about understanding, right? Understanding our audience. And that's what we've chosen to focus on at The Times. And we have had the benefit of some senior leaders there who get it and have not asked us for endless regurgitations of numbers and have given us the freedom to go out and seek insights, and that's something that is very special for us.

How do you find insights? Well, you listen for questions. That's what a good analyst does is he listens to questions. They devise questions. They think of questions. And I think there's two types of questions. There's explicit questions and implicit questions. Explicit questions are like that question like, "How did I do?" People fishing for a number, right? "Give me the data on this. Give me a dashboard on that." Those are real explicit questions—requests for like a spreadsheet or a dashboard—but what they are really looking for is insight. And when they take that spreadsheet or dashboard, they will sort of like pour over it and try to figure out the insights in it, but it's really hard to get insight from reporting.

The more interesting questions are the implicit questions. They are the ones that are on everybody's mind, but nobody asks. And to find them, you actually have to immerse yourself in what people do, because those questions are core to what they see as success, and they are not always encapsulated in numbers. You listen for implicit questions when you show up, right? Woody Allen says, "80% of life is showing up." That's what you do.

So, I've spent a lot of time in the newsroom of The New York Times just showing up, getting to understand how the newsroom works, what they value, and what they want to do, because at the end of the day, that's why I'm there is to help them do that better, not change what they do. The values of our journalists are very important to us, and they are very important to our readers. It's not something that we necessarily want to change.

Implicit questions. There's some like telltale signs of an implicit question. Usually it's like an assumption or a hypothesis, like the word *never* is often

evidence of an implicit question. "Oh, they would *never* do that." Or, "Our readers would *never* do that." Like, that's really a question, right? Not, "Would our readers do that?" Actually, we've found things that we assumed that our readers would never do or never want, it turns out that they actually want. Or habits. Like, newsrooms are full of habits, right? Documenting those habits has been like a personal crusade of mine. "Well, we do it this way because that's how we've always done it." Those things are rich for analysis insight and then action at the end of the day. And so that's what we do.

So, I'm going to show you an example of one instance in which we've sort of immersed ourselves in newsroom culture and come to some insights rather than reporting. And if you visit the newsroom of The New York Times, and if you're ever in New York, look me up and I'll show you around. I'm not an official tour guide. [*laughter*] I was a tour guide in college. I'm not an official tour guide of The Times. There's not much to show actually. It's just like three floors, and it's kind of like an insurance company when you go visit it. [*laughter*] Like, it's really quiet. When I joined, I thought it was going to be like people [yelling], "Get me a rewrite!" and yelling at each other. And that usually doesn't happen, but sometimes it does, and usually that's when news breaks or a big, important event is happening.

Then it gets really exciting, because people are running around. Like, the new pope is named, for example, and you know, we've got the lead all. We've got somebody rewriting the reports coming in from Rome. And then we've got a live blog as well. We have a very good live blog that we use for breaking news events. We have video. We have slideshow. The klieg lights come on and suddenly they are interviewing [people] and copy streaming in from everywhere. And that's really exciting!

But when I was standing there actually watching this unfold, you know, my first week downstairs in the newsroom, a question popped into my mind. It's like, "Is anybody reading all of the stuff that we're putting out?" Not, are people reading individual elements? Because we can tell that pretty clearly from Chartbeat or Webtrends or analytics or whatever, whatever we're using. But are people reading more than one thing? That's actually quite easy for us to do. We went down and we looked [at] out of these ten articles we put out for this event, are people reading it? I'm not going to give you real numbers, but I'll give you a sense of the order or magnitude. A million people were reading one thing and out. 100,000 reading two [things]. 10,000 reading three [things]. 1,000 reading four [things]. So on and so forth. Right?

That was kind of disappointing. The instinctive thing is, well, are we doing too much? Right? Should we stop doing it? But then we decided to look a little deeper into what was going on, and we actually looked at people's likelihood to navigate through the news package. And what we found there was really, really kind of enlightening, because it showed you not just those aggregate numbers, but what was actually contributing to the aggregate

numbers. What users were actually doing. How were they navigating through the package?

So, I'm actually going to show you an example of this. So, it's kind of special. We haven't shown this before. So, Mel inspired us to share when she came to visit, so we're sharing now. This is my one slide. Imagine this is a cocktail napkin, okay, so we're half way through the beer and I'm must drawing on the cocktail napkin. Okay. So, this is a tool we call Package Mapper, and this is an example of Package Mapper from the Golden Globes, which is a pretty harmless thing, so I got permission to show it to you. So, this is the Golden Globes—our coverage of the Golden Globes. [On] the evening of the Golden Globes, there are six things that we produce. The X's are not like the raciness of the content. I just redacted the numbers. I should use number signs. But basically these are the six things that we published the night of the Golden Globes—the six major things that we published. And you can see we decided to -- each ball is like a URL, right, or an asset. The size of the ball are the number of cookies who have visited that. The arrows indicate how many people navigated from one thing to the next, right? So, you'll see this guy over here is sort of like the hub page. It's like the starting point. Like, we like to think it's front pages. You have to have a front page, right? So, that's the front page on it. This thing here, this big red one, is our live blog, which was really popular. This green one is a slideshow about Golden Globe fashions. This little guy is like a preview piece, which is why it didn't get too much traffic. And then these two blue ones are style stories recapping red carpet fashion through the eyes of our fashion reviewers.

And this was like really eye opening when we did this, because we're like, wow, this is how people are actually navigating through it. And we noticed that nodes have different characteristics, right? So, you can classify them. So, this hub page is what we call a *giver*. Like, it doesn't take a lot of traffic to impact it. It doesn't take a lot of traffic from the homepage, but it's giving a lot of traffic in the package, which is what it's designed to do. Okay. But it's actually not quite as many people as we thought are using that as the front page, which itself is an insight. We don't just have one front page. We have six front pages, right?

This live blog is there. This guy is a dead end. The slideshow is a dead end. It is getting a lot of traffic from the package. It is not redistributing it out, which is upsetting. And this guy here is a wallflower. They just stand off to the side, don't join the dance. That's what we call those guys.

So, the immediate thing here is that you see that there are…. Oh, and we colored it by desk, right? So, the color here is…. Sorry, the section that it was published into. So, this green one is published off the movies desk. These two guys are off the styles desk, right? But they are about the same thing. This is a slideshow about fashion. And we have two stories about fashion. And yet, there is no navigation happening between these pieces of content. Why? It's a function of our CMS. It's a function of how we think about our

site. We think about it in terms of sections, not news packages, which is how our readers think about it, right? And we also have producers sitting at opposite sides in the newsroom who are each producing that content individually and probably not talking to each other, right, which is a problem.

So, the insight here is: 1) Our readers think about news packages differently than our newsroom is architected to deal with. 2) There's a very easy solution for this. You just get everybody in a room before we do the Golden Globes coverage. We give them—with high technology—[a] marker and whiteboard. We draw out what we are planning to cover [and] make sure that they are all linked. And if we are feeling like a little advanced and we want some extra credit, maybe we actually like hypothesize about the routes that people will take through the package. Wouldn't that be interesting? Right? And so, that's what we are starting to do.

We now have a tool, a version of this, which is updated up to the minute on every topic that The New York Times is covering that has three or more articles or assets about that topic. And we are able to call up on the spur of the moment and see exactly how readers are consuming things like the Olympics or a plane crash coverage or what-have-you, identify dead ends, [and] coordinate coverage across desks. And it's actually really turning out to be useful.

This is a prototype. One of the things we do is we build prototypes really, really, quickly. We get them in front of editors and journalists. We use it to spark a conversation, which is really important, and then we iterate, iterate, iterate, iterate. This thing keeps breaking unfortunately, but I'm starting to get complaints, which is great, because it means that people are using it. I think one of my goals this year should be like 200 complaints from the newsroom that the tools I've built and spots on the desk are not working. That would show that we're having great success. And since they are starting to complain, I think we are doing a good job.

So, this is one example of the many things that we are doing. I think I probably have limited time, but I do extend an invitation to discuss this further, because there are a lot of really fascinating things going on both with what we're doing and other companies, as Mel alluded to.

I will just sort of to wrap up just point out a couple of things that I think have been intrinsic to our success. And I would call it like reciprocal humility. What do I mean by that? I think data analysts tend to have the tendency to go to the data first and look at the data and then come back with these really kind of like fascinatingly cool, baffling presentations for their clients, which are not usable, because they are not connected to the jobs that they do and the values that they have. Instead, we prefer to take the approach that these people are the most important people in the world. And it's our job to help them do their jobs better, and so we need to understand their jobs, their

motivations, their values, all those things. And the questions have to start from them and then the data comes later.

On the other hand, I think people working in the newsroom have to understand that difference that I just explained to you about reporting and analytics—the difference between metrics and insights. Because the more you ask an analyst for metrics or reporting, the less time they will have for insights. And that's the most soul-crushing thing that analysts can have is to spend their time shoveling numbers back to people when they are not usable. It makes for a very frustrating experience on both sides.

So, thanks again for taking the time and look forward to talking more.

[*Applause*.]

Nicholas White: That is easily the best reaction that my great-great-grand-uncle has gotten in about 100 years. [*laughter*] He died in 1908. So, actually, you know, my ancestry while it's also interesting trivia, at least to members of my family, is I think part and parcel of the story the Daily Dot. So, I'm going to be coming back to that in just a second.

But first of all, I think the first question is, what is the Daily Dot? The Daily Dot actually began as most startups do—as an idea on a cocktail napkin. It was not mine. It was my co-founder, Nova Spivack's [idea], and he said, "You know, a lot of people cover the internet for geeks and for entrepreneurs and for investors, but who covers it for like, you know, people? [*laughter*] You know, like normal people. Oddly, he told me this idea. Actually, he told my friend Josh this idea, and then Josh told me this idea. And I thought, *What a cool idea! Somebody should do that! Why don't we do that?!* And so we did. And so, the Daily Dot began from that, but ultimately what it is, it is a newspaper about internet community, so we like to think of it as, if the United States has The New York Times and France has Le Monde, the internet has the Daily Dot.

The thing is, as a journalism startup, and partly because of my background, we did a lot of things wrong. First of all, we conceived of ourselves as a newspaper, which on the day that we launched we were told that while we were probably a good idea, we had a really bad metaphor, which was very odd criticism to get. And I wasn't quite sure what to do with it, but nevertheless, I took it and thought about it dutifully, but I still stick with the metaphor. The reason is, is because we care about news gathering first and foremost. One of the questions that we regularly ask prospective newsroom employees is, "If you had to pick between one, which would you pick, reporting or writing?"

And for us, the right answer is reporting. And that's because of the other thing that we did wrong. We started a new category. [There] really isn't anybody else who exclusively covers internet community news. And you can't

do that if you're not willing to actually generate original information, original news that no one else has found before. And when you have audience data and that's a core piece of your business, that's one of the key assets of your business, you're faced with a strategic question: Do you give the people what they want or do you give them what they need?

There's actually a really good business model for both, it should be said. In fact, arguably, the business model for pandering is far better than the business model for giving people what they need. You have to look no further [and] I don't want to call anyone out, but one name would perhaps rhyme with *Mox* News, or to be a little bit more current, another rhyme would be *Pupworthy*. [*laughter*] And we don't want to do that. We want to do something else. We want to give people not only what they need and what they should care about, but we want to give it to them in a way that they will in fact care about it.

So, okay, maybe that lead art is a little bit pandering, but the story of this article is that one of the early communities that we focused on a lot was Reddit. And what we discovered in our audience research was that the number one thing that people were searching for, along with Reddit, was NSFW search terms. They wanted porn on Reddit. I know that shocks many of you. But we didn't decided, okay, we're going to do a story, "The 11 Best NSFW Sub-Reddits," [or] "The 17 Best Porn Shots on Reddit." We didn't do that.

What we did, in fact, was a series of interviews with the most popular users of the r/gonewild forums. We tracked down all of the women who were exposing themselves on Reddit and said, "Why are you doing this? What is this all about for you? Why do you keep doing it?" And for those of you who watched any kind of cable television in the 90's, remember all of the ads about Girls Gone Wild? And I had to say, for me, what I always wondered was, *Who are you and why are you doing this?* Which for those of you, if that didn't make it obvious, I'm the worst person in the world to go to a strip club with, so don't invite me to do that later. [*laughter*] Because I'm always sort of wondering, *What was your career path?* [*laughter*] *And where did I go wrong?*

But more broadly, bigger than thinking, okay, we know that people are interested in this, but how do we cover it in a way that's interesting, that is engaging, that is surprising, that is insightful? The other question we ask is…. When we're going to go out, our job, our mission is to tell the story of the internet. That's what we do every day in all of the articles that we write. One way or another, we're telling the story about the internet. No matter how weird or esoteric or bizarre or niche that might be, it is our job, because we are a newspaper, to tell that entire story, every bit of it. And that presents a challenge from a distribution standpoint. But we think that the core insight— and this was one of our founding insights—was the notion that every single article, no matter how weird, no matter how bizarre or strange or internet

freaky it could be, there was somebody out there who would love that story, who would need to read that story.

And this was actually borne out for us in real life within a few months of launching. This was our third most popular story in our first six months. I'm going to do a little audience participation again. Raise your right hand if you know what a brony is. Ooh. Now, raise your other hand if you are a brony. [*laughter*] Ooh, not a lot of those. [*laughter*] Man! So, for those of you who don't know, a brony is often an adult and often male fan of the TV show, the cartoon, My Little Pony. There are actually more of them than you would think. I will point out that no one in this room raised their hand saying that they were a brony, but our audience data would suggest that at least one of you is. [*laughter*] Just saying.

So, why was this story so big? We wrote a story about bronies, which not a whole lot of people do, which makes me thing Gawker was right when they said, "Weird internet people -- crazy internet people now have their own newspaper." [*laughter*] But there was a website, there *is* a website, it's still there, called EquestriaDaily.com. I encourage you to check it out. It is, in fact, the news organization for bronies. And a wave of traffic hit our servers which stayed up. We were thrilled. But they were there. That audience existed. We didn't know they were there. We didn't even know at the time how to find them, but they were there. That passionate group of people for whom this was an essential story…was there, and they found it, and they read it, and they loved it, and hopefully a few of them came back.

The trick of it is, you know, I'm here to sort of contend that a core competency of the media company in the digital world is the ability to market every single piece of content to its particular perfect audience. That is actually a hugely daunting task. We are a tiny company. We are brand new. The New York Times, I think, puts out more than 300 pieces of content a day. We put out 55. We have, even at two-and-a-half years in, we have two different brands. We have 12 sections. We have those 55 stories a week. And somehow we need to figure out how to get each and every single one of those in front of the perfect people every day. That isn't easy, but it can be done. And to some extent, it's good news for journalism, because it does mean that not everybody can do it, because you have to have the will, and the commitment, and the resources, and the practices, and the know-how to be able to do that day in and day out.

So, in terms of requiring an institution, I think, the need for data actually favors and is the modern equivalent, maybe not of the printing press, but of the circulation department. And I think it's very odd that for all the sort of bemoaning of the fact that the printing press is no longer relevant, I really never heard a publisher talking about how sad it was that his circulation department was under threat, which I think is strange. Some of my best friends work in circulation. But I think that's actually where we are, is that

we still need circulation departments, and it's being completely ignored by the media.

Thank you very much.

[*Applause*.]

Joshua Benton: We don't have a ton of time left. I just want to start off quickly for anyone on the panel who might have a thought on this. We have lots of data. We have in some ways too much data and in some ways not enough data. What is the data that you wish you had that you don't have? If we had a reunion of this panel in two years or three years, what would be the measurement of your user behavior or your user's identity, whatever it may be, that you most wish will be in your hands sometime soon?

Melody Kramer: I would like to know if people are learning from our content. I think a lot of people throw around the word *impact*. I don't think anybody has really defined that in a meaningful way yet, but if my audience learns something then I would like to know that and find a way to measure that.

James Robinson: Yeah, so, what I'd like is…. So, we focused exclusively on behavioral data in my first year on the News Analytics Team. And although we have it on the business side—we have a very large audience research department—I would love to be able to bring down some primary researchers down to the newsroom to look at -- to start talking to people and understanding what's going on in their head, not just their fingers.

Todd Cunningham: I would say also, just knowing what I know from my past life at Viacom for 15 years, that set-top box data, the data that actually holds the key to like viewership, and the trends, and all that behavior is completely locked up in, you know, nonsensical reasons for not keeping it, not productive ones, I'm sure.

Higinio Maycotte: And I think, you know, from a sustainability standpoint, not an editorial one, I think being able to prove to sponsors that not only do we have the most interesting audience in the world, but the most valuable audience in the world, you know, for the brands and agencies that are paying for the advertising.

Nicholas White: And I think that part of that actually is, I want to know did the reader that read that article know what site he was on? You know, in the world where the homepage is no longer the homepage and so much traffic is coming through social and in-search and whatever, you can build huge numbers and it doesn't mean that you've built your brand. And, you know, I think that that's ultimately what journalists are contributing to, and that's what makes great journalists valuable is that people are going to care about that brand, they're going to care about that banner at the top of the page.

Joshua Benton: Actually, you reminded me of something I thought of during James's presentation. It seems that as so much attractive input shifts to social, there's tons of social oriented data, but it sort of has meant that our audiences in some ways have become more transitory. They are coming from many different points of input. And one of the big challenges I'll say we have at Nieman Lab, because we have a very socially-oriented traffic pattern, is that recirculation of getting someone to go from one article to the next to the next. I'm curious, you know, a perfect example, I know you talked about this, but what sort of best practices have any of you seen around getting someone who clicks on a link on Twitter, goes to your site, maybe they know your brand, maybe they don't, but to convince them to stick around and go click on something else? Anyone have any secret, magical tips on that?

James Robinson: A newsroom with 1,100 people who are constantly putting out great material.

Joshua Benton: Okay. Everybody get 1,100 journalists. [*laughter*]

Nicholas White: I think to some extent I would tweak the premise of the question a little bit, which is to say that, like, I think that you're always going to be fighting an uphill battle with somebody who's coming from social media, because what they're actually engaged in is their social platform, and then they're jumping out to get the thing, and then they want to go back and say whatever they have to say or react to the tweet that they saw that brought them there. So, I think that what you actually need to be worried about is converting them into people who are going to come from Twitter again. I made fun of Upworthy earlier, but what they are really, really good at --

Joshua Benton: You mean Pupworthy.

Nicholas White: Yes, Pupworthy, which is actually my next startup. It's going to be all cute puppies. [*laughter*] Because there's too many cats on the internet.

Joshua Benton: Aggressive puppies who make you feel good about society.

Nicholas White: Yes, right. Yeah, yeah.

Joshua Benton: Or outrage at the corporation.

Nicholas White: Yes, exactly. But they, you know, they are really good at converting. They are not interested in you recirculating on their site. They say, "That battle is lost, but we want to convert you into somebody that we can bring back tomorrow from the same social network."

Higinio Maycotte: And I think it's a really good question. I mean, I think somebody earlier today said that mobile is dead. If I had to have an equivalent statement, I would say that in three to five years, the URL may be dead. So it's about using data to optimize context. You're going to have to find and monetize your readers no matter where they are and it might not be on your own media assets. And I know that sounds scary, but if you don't know your readers, if you can't value and understand how they are reacting and engaged with your brand, it may be very, very hard in the future to ultimately make money.

Joshua Benton: Let me ask about personalization, which didn't come up as much as I was thinking. The Times has the recommendations engine that is gathering data on what I like at The Times for some number of years. NPR has the ability to geolocate so that they know that I'm in Boston, so they give me [local] content in some sort of fashion. I'm curious, for me, as an observer, personalization seems like something that has always been a few years away. They've never quite gotten there. Do you think that if we get back together in two or three years, that we'll have made significant advances in that?

Melody Kramer: I think that we'll know a lot more about specific types of readers. And what I mean by that is somebody who goes to our site to read a politics story isn't necessarily like someone who goes to our site to read a piece on Justin Bieber. And how do we better know the politics reader and what they want and provide that to them? And let's just say that everybody coming from social doesn't go through our homepage, but people who go through our homepage are more likely to comment or people who go through our homepage stay longer on our website. And currently, we're not connecting the dots in a way that we can do anything with the information that we have coming in. But I think in knowing about each type of reader and the way that they are engaging will be more meaningful.

Todd Cunningham: I think that, yes, we will be able to know a lot more in a couple of years, mostly because now we have dynamic ad insertion tools that will allow…. There are more people in the game who have skin in the game in trying to understand who that reader or that audience member is. The part that I see is lacking is like a relationship-oriented outlook. Let media outlets continue to do the one-time-only thing and like re-up every single time with the reader or viewer or whoever the audience member is, instead of trying to build a relationship, a longer-term relationship with them. That's in essence what the audience member wants to do and it's also what the business wants to do. Is there some kind of like polar thing going on?

Higinio Maycotte: And to bring that back to business terms, I think if you are an e-commerce company, you really understand things like lifetime value of a new customer. But what is the lifetime value of a new visit? Is it worth investing in audience acquisition and audience building? If so, how much should you invest? And how does that pay out over the long term? I mean, I

think those sound like scary words, but it's all part of what we're going to have to adopt in the future.

Melody Kramer: You know, I think NPR has for many years, when people listened on the radio, it became a habit. And so every morning, you get up and you drive to work and you listen to NPR. And one of the most interesting things that I heard about in this entire two-day session was the WNYC sleep app, because I thought, *Oh, my goodness, they are now going to create a habit for people to tap into WNYC when they go to sleep [and] when they wake up in the morning.* People are logging their behavior and they are doing so through WNYC, which is absolutely brilliant. If they put in push notifications to that, I mean, they get people to use their app and view their content in a different way. And they are doing that by having people monitor their own sleep patterns. That's personalization.

James Robinson: Yeah, to answer your original question about personalization, I think it's an important element, and I think it will continue to evolve. I don't think for us it's necessarily the end-game, you know, the completely personalized New York Times. I don't think our readers would really want that. I think our editorial judgment is really important to them. And the sense of serendipity that you get when you discover something you didn't know you were interested in is also very important. I will say that from a data analyst perspective in the newsroom, we're working very closely with our recommendation engine folks, because we're doing sort of this same task from a different angle in the newsroom. For them, it's about finding the right collection of stories that a given reader would be interested in. And for us, what was really interesting, the direction that we're going in is, for a given story, how can we model out the readers that will be interested in that story? Which is really the other side of the same coin. So, we're sort of working together to build the systems that allow both of us to do those things.

Joshua Benton: Great. We have a few minutes left for some questions. A reminder, state your great-grandfather's name [*laughter*] and affiliation. And a question is something that ends with a question mark, not a "so there." [*laugher*]

Jonathan Groves: Honestly, I can't go all the way back to my great-grandfather. Jonathan Groves at Drury University. So, my great-grandfather Groves, I'll just say that. Great panel. I learned a lot. I was glad to hear so much focus on high-quality content. One of the things that I was curious about in thinking about the insights that we get from analytics, there's a lot of focus on the page views and the uniques, and we see those spikes in content typically short term. You know, everything is about the here and now. We see that either it's a day or a week. I'm wondering about your content that lasts for two weeks or a month. What kinds of content are you doing that gets that kind of traffic? And how do you measure…? Or not measure. Sorry, wrong word. This idea of longevity.

Nicholas White: I'm trying to remember who originally I heard this from, and I can't remember. But the notion is that there's three types of content and we are good at thinking about two kinds and not the third. And the first one is news, and the second one is evergreen content, and then there's a third category which is 'news to me.' And so, I think actually if you can do personalization or if you are very good at distribution, there are efforts that can be made around saying, "What is that set of information?" I can remember my brother like three months into the Great Recession, he was just out of college and not real big on actually keeping up with current events, and he sent me this text that was like, "So, I understand there's like a recession going on. What's that all about?" In text form. And I explained it to him as short as I could. Actually, no, you know what I did? I found the link to The New York Times story explaining the recession, and I just texted it to him. But anyway, like, you can find that interaction with readers. So, we have processes where we create curated feeds of the stories that we think are still out there that could be new to the right people. And then, we try to disseminate those and constantly are monitoring which ones are still finding new audiences or whatever. But it takes its own sort of process to get those in front of people.

James Robinson: Yeah, I mean, we have a lot of newsy content and everybody content. It's pretty easy to tell what has an expiration date and what hasn't. I mean, we had a huge amount of traffic to our archives going back all the way to 1850. We have a great new app called Times Machine where you can actually look at articles in their original form in the paper, like read the paper from 1972, which is awesome. So, I think, you know, it's what the reader is interested in at that time. There is news that is perishable, but there's a lot that isn't.

Melody Kramer: I'm really interested in evergreen in terms of how you can tag or mark content for being available on either an annual basis or, say, "I'm going to publish this piece, and I think it's going to be relevant for the next four months." "I'm going to publish this piece, and it's going to be relevant today and also a year from now and also five years from now, because it will be an anniversary." We haven't done this. This is just something that I'm thinking about. How do you take your content that you're publishing and say, "I know that this will be relevant again because of this"? Is there a calendar that you can build? Is there a way to tag so that you can pull things out of the archives automatically? I'm currently working on an archive project at NPR. We also have a Tumblr. It's NPRchives.tumblr.com. And we pull out a piece from 1984 every day and we're using it as an experimental space for the NPR library to play around with different kinds of technologies that will potentially allow us to bring up evergreen content on a more regular basis.

Joshua Benton: There was a great story on Tumblr about the radical trend of women wearing sneakers to work.

Melody Kramer: It's really good.

Joshua Benton: Which is a very interesting window into social morays in 1984. Just related to that, I think there's an extension to that, which is, how can you declare content no longer useful? Like, I hate it when I go to local news and search about an event, let's say, a sporting event, and I get the preview story, even though it's over, because it still exists as a unique URL.

Melody Kramer: I used to work at Fresh Air with Terry Gross, and all of our interviews were taped and all of the intros were live. And the reason for that was when somebody would pass away, we were able to take the interview, slap on a new intro that Terry would read live on air, and then use the same interview. So, it was content that part of it was archival and we could use that again and part of it was new.

James Robinson: We have the opposite experience. We have obits that are already written by people who have passed away. [*laughter*]

Joshua Benton: Other questions?

Robert Hernandez: Over here in the back. Robert Hernandez, mobile is dead guy. [*laughter*] I have a question for Nicholas.

Joshua Benton: That's your Twitter bio.

Robert Hernandez: Yeah, that's not in my bio. Question for Nicholas with the Daily Dot. How do you not view an investigative journalism piece into gone wild as pandering? How do you not connect it to pandering? And second, this is just out of curiosity. The back story of your Twitter handle "Desert Bunny."

Melody Kramer: Yes.

Robert Hernandez: What would your grandfather think? [*laughter*]

Joshua Benton: Attack of the softball. I like it. [*laughter*]

Nicholas White: So, I think that it's not pandering because that wasn't what they were looking for when they searched for that. They were not looking to find out what Nina 17's background was, where she went to school, why she did what she did, what she got out of it, how it affected her self-esteem. I can guarantee you that most of the people searching 'Reddit nude pics' were not worried about anyone's self-esteem. [*laughter*]

Joshua Benton: Certainly not their own. [*laughter*]

Robert Hernandez: They searched 'gone wild.' They didn't search self-esteem. They searched 'gone wild' and they saw a breast greeting them at the start of the article.

Nicholas White: Yeah. So, you're saying, like, okay, so they searched for 'gone wild,' and then they saw our shot with breasts at the beginning of the article. Well, I mean, if they were searching, they were probably searching on Google, in which case, I guarantee you and article on Gone Wild was not the first thing that came up. Gone Wild was. And second…. Hopefully, it was on the first page though. And second, like, the image wasn't there. You know, so, and the traffic for that story wasn't search traffic. You know, the traffic that we actually got was social sharing. It was picked up by, you know, women, actually, wanting to talk about the experience and actually kind of the sex-positive movement and that sort of thing. So, I think that you have to take into account where people are in order to…. I think somebody was talking about this on one of the other panels. But if you understand where the other person is, they are more likely to listen to you. And I think that that's really where we were coming in. And we got a story that I really don't think the people that were searching ended up being the audience for that story, but it gave us the idea. And I'm happy to tell you the story of my handle over drinks later.

Joshua Benton: All right. We have a question here.

Emily Allen: Hi. Emily Allen here. What are the ethical implications of this data?

Joshua Benton: Can you be more specific?

Emily Allen: Like, if your analytics that you're getting, like, what are the ethical implications you're getting from it like of your audience.

Higinio Maycotte: I'll say I think it's a complicated discussion, because when you ask the question of, who owns your data, just imagine being at a football game on Facebook taking a photo. You know, that data is passing through the facility, you know, the stadium itself. There's the teams on the field. There's the tel-com company that's serving up the bandwidth to make it all happen and then ultimately maybe the networks where that goes. Who owns it? You know, who's the owner of that picture? So, I think at the end of the day while those questions remain very complex, I think the key is absolute transparency in choice. So, I think as the person seeking the data and putting data to use, you just need to follow the rules of those who you're participating with. You know, the networks have very complex rules, but that's more a legal issue. And then from a consumer side, tell them exactly what you're going to do with the data, give them the choice to opt out, and also don't make them -- you know, don't force them to do one thing. Always give them options.

James Robinson: Yeah, ethical issues are really at the forefront in our mind a lot. They come up a lot. One example is we have a series called High Cost of Healthcare, where we actually ask people, you know, "Have you had a procedure?" "Are you a doctor?" We don't use that data for analytics. That's newsroom data. It's collected for news gathering. We're not going to use it. There's other ethical issues that have arisen at other organizations like Bloomberg, you know, looking at the terminal data. There's like newsroom ethics issues that arise. We once proposed an AV test, where we were going to serve half the people a three-paragraph summary of something and then the other half wouldn't get anything. And we vetoed that on the grounds that we didn't want half of our people not to know the news. Right? So, they come up all the time, and we try to handle them as sensibly as possible, and we keep our customers, our readers like foremost in our minds.

Nicholas White: And I would just add really quickly that just as much as like we're all worried about the ethics of like the data that we maintain, and it is something that I think everybody on this panel probably worries about, and everybody in this room would worry about. There's dozens of other companies who just blithely went right past that, who the average person has never heard of. And on the advertising side, they are generating *huge* amounts of this data, and it's already in a database somewhere, and they are completely invisible. And that, to me, is the much bigger story that we should be paying attention to.

Higinio Maycotte: And to add to that, I think it's all of our responsibilities to understand the analytics packages we're placing on our sites, the ad networks and ad exchanges we're working with. And most of those are the ones that are stealing all of the data [and] not asking for permission [and] sharing across sites. So, you know, tracking cookies are a really big deal. And I think a lot of media companies now are on to that. You see media companies like Turner and News Corp. kicking out the ad networks completely in a panic when they have these realizations. So, I think you need to be very careful, too, about the leakage and the data that you are inadvertently giving away as well.

Nicholas White: We've totally contributed. It's the stuff that we are aware of, were ethical about it, then we've completely—completely contributed to the creation of unethical data.

Todd Cunningham: Having a place to park your data. We're building our data listserv with a safe data harbor for everyone.

Joshua Benton: It's a trap. No. Question over here somewhere.

Chris Amico: Yes. Sorry. I'm Chris Amico. I'm one of the co-founders of Homicide Watch and two of my great-grandfathers fought in World War I for Italy. One was captured by Austrians.

Melody Kramer: Is that your comment? [*laughter*]

Chris Amico: That's my great-grandfather comment. My question is about what metrics have you stopped gathering? What did you once gather that you found no longer produces useful insight?

Joshua Benton: What are the numbers you no longer look at as much as you used to?

Nicholas White: I can tell you we don't look at the traditional engagement metrics anymore—pages per visit, pages per visitor, and stuff like that. Those are essentially irrelevant now, because we've gotten past the page view. We've gotten past so many things. And the nature of traffic now, I don't think those are relevant ways to talk about engagement anymore. There are other ways to do it, but the traditional ways are gone.

James Robinson: Just a riff off of that. We don't really look at averages so much. We look at distributions, you know, because we want to understand segments.

Todd Cunningham: Yeah, same here. Sentiment scores [are] just not reliable. For social media data, sentiment monitoring.

Joshua Benton: Great. Another question here.

Su Kong: My name is Su Kong from the University of Texas at San Antonio. I have a quick question about Package Mapper…to change.

James Robinson: Next segment, I hope.

Su Kong: Yeah. You showed several circles in different colors. Then you mentioned…. Kind of would you give us some relations on that? Like, is it the color and also the size? Does it shows view clicks first or view counts? And second, does it also show then why some pages have smaller circles and the others have larger circles there? So you can tell them, the news team, newsroom, why they got really small clicks in those.

James Robinson: Sure. Sure. I spoke really quickly when I was doing it. So just a short recap. The size represents the audience for that piece of content. The arrows represent the clicks from one piece of content to another. The color represents the desk or the section that it was published into. And the way that we communicate it to the journalists…. What I showed you was really an analyst tool. It's something that is very intuitive to somebody like me, but no so intuitive to somebody who's trying to like cover a breaking news event. So, one of the next steps, I think, for Package Mapper will be to automate alerts based on things that we see happening. If we see that something is not linked, we might have it go out in an email. We might have

a chat bot pop up into like our homepage producer's chatroom. We want to give people the information they need to do their jobs, not necessarily blind them with science. So, that's how we use it.

Nicholas White: What were to the two concentric circles?

James Robinson: Artistic license. At one point, they represented something, but they no longer represent it. [*laughter*]

Joshua Benton: Got an artist here. All right.

Ryan Single: Ryan Single. I run a content recommendation company called Contextually. Question: Have you guys found anything in the data that you can actually use to give to writers so that they can actually change their behavior or to figure out what kinds of topics or things they should be doing? Things they should be doing more of? Things they should be doing less of?

James Robinson: I'm the least qualified person in the newsroom of The New York Times to tell people what to write. That's an editor's job. It's not my job. And the way I approach it is that what they choose to write about sort of comes down from the mountaintop. And my job is to help every story find the audience it deserves. So, that's not really my role at The Times. I know other companies do it differently, but that's not something that we're doing right now.

Melody Kramer: We do the same thing.

Joshua Benton: One more.

Janine Warner: Janine Warner. I'm one of los gringos que habla Español here. And my great-great-grandfather signed the Declaration of Independence.

Joshua Benton: Damn.

Janine Warner: My grandfather was very proud. I did some research. He was a total renegade. I don't know. Are there any questions you can't answer today that you think you'll have the data for soon?

Melody Kramer: We're four weeks into a six-week development process, so we don't actually have a dashboard yet. I'm the product owner. I pitched the thing to the NPR vice president, where now I have three developers. I have a UX guy, Scott Wright, is on my team. We'll make it public. Like, we're not hiding anything. It'll be out in like two weeks [when] I think I'll cycle will end. And then we'll have something.

Higinio Maycotte: Maybe a slightly adjacent answer. Not the exact answer to the question is, I think it's also okay to ask your users questions. Surveys

are okay, but don't spend your survey questions asking things that the data already answers. You know, so many surveys ask us for a lot of the questions that we've answered on every other survey we've ever done, so there are definitely certain things today that the data cannot answer and maybe never will be able to answer, but back that up with real knowledge before you go out and ask the question.

Joshua Benton: I'm going to take moderator's privilege and just ask one question very quickly. It's great to talk about having data teams analyzing data. Most people in this room probably don't have anything like that. They probably have Google Analytics, maybe something else. I'm curious from the tools that the average news website might have, let's assume they have Google Analytics, is there one or two numbers that come out of that or one or two measurements that you think would be particularly valuable [that] people should be paying more attention to?

Melody Kramer: So, there's a…. If you're on GitHub and you search for Google Analytics Python, there's a program that allows you to write very easy Python programs to pull things out of Google Analytics. I could get the name for it and tweet it out. That's really helpful for smaller shops, because you can just literally manipulate the Google Analytics data the way that you'd like. And you don't have to buy anything. You don't have to do anything major.

James Robinson: I think we tend to look at the things that are doing really well and not at the things that aren't doing so well. I think if you reverse sort lists, sometimes you really see interesting things. If you have like a heat map tool on your homepage, turn the opacity all the way up to 100 and look at the things that aren't getting any clicks. I think if you look at the bottom of the ladder rather than the top, I think there's often insights to be found there.

Todd Cunningham: If you layer over that data like what you're selling it for or what its street value is, you might find out that some of the things that you are seeing is of low value are like actually of really high value and vice versa.

Nicholas White: Depending on how deep you are into Google Analytics, in the release about six months ago, Google Analytics started allowing you to use user segments as well as visit segments, and those are incredibly powerful.

Higinio Maycotte: And I think without, you know, access to the complex analytics packages or the analysts or the PI experts to help you sort of put together these types of questions, I always like new versus returning and being able to analyze that on a per author basis or a per content or category type basis. I mean, I think as you're building an audience, that gives you a

really easy indication of where your audience is growing and developing over time.

Joshua Benton: All right. Please join me in thanking these wonderful panelists.

[*Applause*.]