

Rashmi Mohan: This is ACM ByteCast, a podcast series from the Association for Computing Machinery, the world's largest educational and scientific computing society. We talk to researchers, practitioners and innovators who are at the intersection of computing research and practice. They share their experiences, the lessons they've learned and their own visions for the future of computing. I am your host, Rashmi Mohan.

If you accidentally discovered your new favorite song while out on your morning run today, you'll have our next guest to thank. Mounia Lalmas is a director of research and head of tech research at Spotify, where she leads a team of researchers across the globe solving problems in the domain of content personalization and discovery. She has a rich career in studying user engagement and who holds an honorary professorship at University College, London. She is an author and a regular committee chair on many top tier conferences like SIGGRAPH and WSDM. Mounia, welcome to ACM ByteCast.

Mounia Lalmas: Thank you.

Rashmi Mohan: Mounia, I'd love to lead with the question that I ask all my guests. If you could please introduce yourself and talk about what you currently do and also give us some background and insight into what drew you into this field of work.

Mounia Lalmas: Okay, thank you. I'm Mounia based at Spotify. I'm a researcher, I've always been a researcher first in academia, now in industry. My passion has always been evaluation and user engagement, this is a very important problem in many online industries especially around personalization. Why this excites me is it's a health problem, everything is becoming more and more online, personalization is getting bigger and bigger and doing it right remain a health problem. And not just that knowing whether we're doing it right or not, so that's why user engagement is something that I get up every morning, oh, how can I solve a few things?

Rashmi Mohan: That's super exciting. But I'm wondering if I could go back even further. What drew you into computing?

Mounia Lalmas: This was a long time ago. I was just good at math, I love math. I love the obstruction level that you can get with math, for example with respect to algebra, logic, geometries and so on. At that time, if you weren't good at math, you were going more to a career related to become a teacher. And I was not sure this is what I wanted to do, and there was this opportunity, it was called where I grew up in Algeria informatic. People that were quite good at math, many of us ended up into this area of informatic computing and we are still there.

Rashmi Mohan: Oh, that's fantastic. Did you always imagine that you would get into computing research or as you delved into your academic pursuits, you found that there

were problems really that you wanted to solve because you didn't have the answers to them?

Mounia Lalmas: Yeah, I didn't know much about the whole area of research at that time when you start as an undergraduate student. I started to get into it, in Algeria when your study is about five years and at the end of those five years, you do your master, is calendar of the master project, is one year research.

At that time, the hot topic was actually the expert system and this is where I got really interesting, is started to think about what challenge and how to approach those challenge. So not just applying things but okay, we don't know how to do this, this is a problem we need to solve and how to go on about it by being rigorous, by doing a lot of reading and understanding what are the latest into state of the art and so on. And this is what drove me more into research as part of this final year project, which then brought me to do master. This is where I moved to Scotland to do my master of applied science at the University of Glasgow and then I ended up doing the PhD. So it was not planned but it followed this excitement I got at the end of my breaking of a bachelor degree for example, in the UK.

Rashmi Mohan: I like how you call out those traits of doggedness, the need to be rigorous about the kind of problems you're trying to solve, the need to read more and investigate more. Such classic traits for anybody whether they want to do a PhD or just is interested more in researching the areas that they're working on.

Going back to the point that you made earlier which is around information retrieval, and you were saying that the problems are just becoming bigger. How do you feel like your research interests have evolved over time?

Mounia Lalmas: When I started information retrieval, actually it was a coincidence. I did my master at Glasgow University, I did it actually on formal lectures, again, the strong link with my interest in math and in particular, logic. And it was like okay, I'm interested in a PhD and it happened that the University of Glasgow, there was a very strong information group. And at that time, one of the big topic is how to use logics into building better information retrieval system. So that is how I ended up in information retrieval, and I stayed there. At that time, we didn't have the big search engine. Information retrieval was still for example a big intersection with library, information studies and so on.

And then with search engines, things got really, really big. You have suddenly everybody with an information need, it can be very precise, it can be very vague and so satisfying a user is not easy. It's very, very, very wide and so on. And so this is where I got more and more interested about the evaluation. So less about the algorithm, but it's always easy to return result there maybe for some extent, good enough, but is how do we even know if it's good enough? And so I moved more into the evaluation of the results and evaluation is a big part of information retrieval research. Anybody that is into this area as a researcher or

applied researcher or engineer will always say okay, how do we know that what we're returning to the user is good? And so on. And this have been more my path and less on the algorithmic side.

There's been a lot of progress on the algorithmic but I'm always been interested, what does it mean for the user? Are we retrieving the right things at the right time? And this journey is actually not finished. Why are we growing the algorithm? We still keep on asking the questions. What does it mean to the users? Because search has grown now so much. Intranet, internet, e-commerce, music, shopping and so on, and internets are very different, satisfaction are very different. So there's a lot of research still waiting to be done.

Rashmi Mohan: It sounds so exciting and such a valid point that you bring up that evaluation of the results almost is what will feed back into the algorithm to make it better. And one of the things that I know when I was looking up your previous work, the need to measure was there even in a lot of the previous early work that you did, I know that you did a lot of work around measuring user engagement and that was again, pioneering at that time, just like you said, when search was just coming up, the idea of providing content to users, consumption of content was growing. And so understanding how does a user stay engaged was really important and you did some very, very incredible work around that. Do you remember the key innovations at that time that felt like a paradigm shift in the field?

Mounia Lalmas: That's a really interesting question, I could write a book about it. There are two parts. There's the evaluation which is the offline evaluation with very precise metrics, for example, area under the curve, precision and recall. Then there is the online evaluation where actually systems are running and we get actual feedback from the user with key time spans and so on. Earlier in the field of information, there was a lot of progress with the offline evaluation also in the sense that precision and recall does not capture everything related to engagement, to satisfaction and so on.

And not just me but a number of people have been working around that, trying to build better metrics especially with respect to offline. Then going back to the online, it is different. Here is you have those concrete feedback from the user. And there's always this thing is if everybody click, it's a notion that the results are good and so on. And this worked for a while because it can be viewed more as a proxy of user engagements, a lot of the metrics actually that people refer to are proxy of engagement. But then there's always this question, what does it really mean in reality, the value of a click and so on? This is where I started to be much more interested. So you have the offline, you have the online, there is a connection between those two high precision, does not mean that is going to lead to a long-term engagement.

So those are the questions, not just myself but many people at company and also at universities start to really understand the connection. And to some

extent, what was maybe the breakthrough is to ask those questions and to not just rely on metrics that everybody used and said well, I'm doing right, I've got the right precision and so on. So with the breakthrough, you just say actually, what does it mean to have high precision? What does it mean to have high recall? What does it mean to have high click through rate and so on? And by trying to ask this question, myself and others have come up with maybe better metrics to really understand well, satisfaction means often user returning. And then if you agree with that, you're trying to find what are metrics that correlate with this? So it's really the thinking.

Another thing that is also important is metrics. When people talk about metrics both online and offline, they have to be very careful that the task that is being looked into has an impact of what the metrics mean. For example, social media, a good metric is people spending quite a lot of time on it, why search? It could be that the people just click on result and spend as little time. If the search results are good, the user find what they're looking for and just leave. So again, this idea of metrics of evaluation and so on is very specific to the task at hand. And the field of information retrieval I would say was one of the earlier ones to really look at this and asking those questions. Precision and recall are still used but we are using also a lot of other metrics.

There's this new area which is I'm not so much involved by, which is conversational search. What is the metric? What is success? And so on. And being brought out as an information retrieval person, we constantly keep asking. So as long as we ask, we will make progress.

Rashmi Mohan: Thank you for that actually, because what you bring up is not something that we may always think about if we're not in the field, which is really the metric of success is different depending on the type of problem that you're trying to solve. And similarly also, looking at the impact that a metric has on what you're trying to solve for.

For example, one of the things that I was thinking about which I know in the past, I've also looked at is there are so many other dimensions that are changing with how a user is interacting with the service that you're providing, whether that is search or whether that is reading content, what you initially may think of as the mode of engagement, maybe the reading on a desktop versus now moving over, not now but maybe a few years ago, moving over to a device like a mobile phone. Suddenly the kind of metrics that you're looking at are very different and you may not get the same signals that you do with a desktop computer in terms of clicks, et cetera. Do you remember some of those shifts and what do you see as trends that are changing now in terms of the kind of metrics that people have to evaluate?

Mounia Lalmas: A big change that people trying to incorporate when they measure success is user behavior. For example on the desktop, it's very well known there is above the fold, people don't go below. So everything has to be optimized for quite at

the top of the pages or what is visible on the desktop screen. This is not the case on mobile phone. Scrolling down is much, much more common. So if you just look at scrolling down on desktop and compare this to the phone, if you just compare metrics without being aware those are different user behavior, then you just get completely the wrong results. And this change a little bit what is viewed as success, for example when reading news on desktop and reading news on mobile.

For us for example, just in the context of Spotify, we have the mobile experience where we try to find a way to combine familiar content that the user wants to listen of particular artist, genre and so on, and the other content which is more about discovery and we take into account that people would browse to some extent up and down, and this allow you to investigate the value of metrics differently. So again, what I'm trying to say is user behavior has also a strong effect on the metrics. And even if it's the same product, and sometimes you may keep the same metric, you may keep okay, let's just keep at the moment clicks real rates but it has to be interpreted differently, knowing it is desktop and knowing it is mobile or other kind of device.

Another thing also, it's related to again some of the work we have at Spotify. So we have playlist, Spotify has a lot of playlist [inaudible 00:15:06] into them. And what is success of a playlist? And we have playlists that are made for people to fall asleep and some are made for people maybe to do party type environments. A sleeping playlist success means the user stopped listening for that playlist and does not do much, that's the whole point of a sleeping playlist. While other playlists about songs that you may want to use to build a party playlist, there's going to be a lot of interaction and a lot of scape and so on. And this is success, the user is trying to act like a DJ and try to extract some track to create their own playlist. So again, even the same product but two different part of it, two different playlist with different intent, success is very different.

And we are looking into how this is helping actually building a better personalization at Spotify. Hopefully it gives you an idea that there's not one answer, it's user model, the application, and also sometimes going down to the item itself.

Rashmi Mohan: Absolutely, I think that gives a lot of clarity. And the two things that you said there that I'd like to maybe talk a little bit more about is one, of course, you're talking about personalizing the experience for the user, understanding the user's intent and possibly the context in which they're using the product. And one of the other things that you said which is in some cases, you are trying to optimize for discovery and in some cases, you're trying to optimize for surfacing maybe new content. How do you strike a balance between those two?

Mounia Lalmas: That is a good question, that is what we're working every day. There's various way. Again, I'm talking about work we're doing at Spotify, not just a researcher work with engineers, product manager, designer. So you can organize for

example the front page in the way this is for discovery, this is for the kind of things you're looking to do now. That is one way you try to understand various needs. And because we have so many users and other online services in shopping and they have so many users, they get a good understanding what a general big needs. Needs of something novel, need of something to solve a problem that is just now, if you have enough of those either identified intent need is easy to even organize the front page, the homepage accordingly. Now, the other thing is, what do you have this personalization work?

Personalization is about trying to return to the user what is the most relevant to them, and of course the definition of relevance is something that evolve. It could be this is what you tend to listen to, we're going to give you more of this. But in the context of entertainment, and particularly in the context of music, we know that this is just not good enough. We have to feed the user of course with what they want to listen now, but this is always a journey. User will just evolve into that listening. And we can just decide okay, there is a playlist. We may try to find way to attract to it that is related to other tracks and that place is related to what we think the user is likely to listen or not. And then you get signaled back, did it work or it didn't work? And then from there, you can build better algorithm.

It's a mixture of what is called the bandits area, kind of exploits although the definition there a bit more technical, it's like trying to provide to the user what they need now but try also to show other things. And it could be something as simple, maybe 10% of the time I'm going to give the user, something that is a big difference. And from then, you can start to get some signals. Does it work? Does it not work? And then you can build also better algorithms to incorporate those signals back. It's an exploration by itself.

In the context of music, I think we're lucky that many users are on a discovery journey. They change into what they listen. They may listen to something quite a lot now but at some point, they will want something else. So I think we have more room to push more discovery. It doesn't mean it is easy, it's not just I'm going to give you something completely random. No, it's just to find the right level. We're also trying to understand how our user ready to receive more different diverse content. Some people are just pretty happy to listening to particular type of things and this is perfectly all right, some people are more open.

And we can look into this by just looking at the listening behaviors. Do they listen to very diverse sets of for example, Java and so on? Or is it very, very specific? This is this notion of going to understanding the users, understanding the content, understanding and acknowledging that we have to find a way to give the right content but somewhat injecting into it something a little bit different. We can do that explicitly. For example at Spotify, we have Discover Weekly which is this playlist every Monday, which is a new content, or we can do it while user is listening. Okay, maybe having some track or some song that

are a little bit different, they haven't listened to that particular artist that is close enough to for example what the session is about.

So we try those barriers way and we're learning from that. And again at the end, those are signals then we take back, understand feedback into algorithms, allow us to build better algorithm and so on.

Rashmi Mohan: It's funny because I heard from, of course I have teenage daughters who use Spotify and they absolutely enjoy it. And you're often doubted to be able to "read the mind of your user" and what you're talking about is incredible because it seems like a lot of the work you're doing is really to understand what does my user need? And providing them that. And also like you mentioned, music is probably an area which lends itself well to a little bit of exploration and discovery, but there's going back to this idea behind reading the user's mind, feels like Spotify is able to give the user what they're seeking at that time. How do you achieve that? Is that a very conscious mission that you are using as your north star as you go down your parts of research or product ideas?

Mounia Lalmas: Of course, we're not going to tell you our secret but again, music is not new. People relate to music, group of people relate to music, people listen to music together. We have editors that are expert in some particular type of music. So we know that what people want in the morning on Monday is very different to what they want on Friday evening. So there's a lot of knowledge that comes from experts in music. And this is why we have playlists that are created exactly for that. We have playlists for again lullaby, sleeping, yoga, the gym and so on. And working with experts, they just know what they're doing.

Now, the second part is imagine the playlist about happy music or sad music. What makes me happy may not be what another person feels oh, this is a happy song or this is a sad song and so on, and this is where the personalization come into account. So from this, oh, this is sad music, this is happy music, this is running music. So how do we personalize to for example artists, genre or beats and so on, that is specific to the user? So it's combining a little bit what we refer to as human-in-the-loop and the algorithm to bring this together. And by doing this, it looks like we're reading the mind of the user, which is good.

Rashmi Mohan: Yeah, that's great. I like the phrase that you say, human-in-the-loop and to bring that extra level of intelligence into these recommendations. In terms of personalization research, Mounia, what are the common computing problems that you're trying to solve and what maybe the industry is really looking at right now?

Mounia Lalmas: There's a number of them. For example, a lot of firm algorithms are not yet scalable. So we may come with the best algorithms, whether us in the research or in academia and so on. And then it has to be scalable, users do not wait much for a search result. And so this notion is really, really, really good algorithm that works really, really well, we need to make it scalable. And by the way scalable is

also to be reactive. It's like okay, suddenly we have a lot of signals that are a little bit different, how does the algorithm react to it? So this is also the aspect related to the scalability.

A big area that is at Spotify but also elsewhere is explanation and interpretability, transparency, those algorithms are really optimizing in general, always for the next click. Sometime not much explanation, there's a whole area with buyers and there's a lot of research around this at Spotify, but also elsewhere. By just letting the algorithm running on their own, there's a lot of problems happening and it's important to address them. The other thing which is I think a bit related to this is to move from again, looking a little bit more machine learning jargon is, what those algorithms are trying to do is to optimize for a metric for an objective function and to be like click through rate, just to click optimize for the next click. And it's very well known that this is good for the moment but it's not good for long term.

It's a hard question, how do we know that what we're trying to optimize for now is good long-term? Some of the research we have done is people that have a more diverse listening tend to stay longer for example on Spotify. So it's important for us again, going back to discovery. So it's not just about optimizing for the next click, it is optimizing for what we call the long-term user satisfaction. And we're not there yet, which is kind of exciting because this is often also going back to my passion which is metrics and user engagement.

And at Spotify, how we are proposing to go into this is to rethink how we do the optimization. And we are investing in one particular technology which is reinforcement learning because we believe it will allow us to do that while also allowing us to interpret the various models, transparency, and those are challenges, but they are challenges that many are trying to address now and this one of our focus from the personalization perspective, not just the next listening behavior, the longterm listening behavior, and this is what will make personalization more successful.

Rashmi Mohan:

I like how you tied that together in terms of the work that you're doing in personalization, but taking it back to the work that you do in user engagement, not just looking for user engagement in the near term but really the longterm behavior that you're trying to optimize for. That's great. One of the things, Mounia, also wanted to touch upon is just in terms of your career, you've spent a lot of time as you had said earlier, doing research in both academia as well as in industry. And in industry, when you're working for a research organization or heading one up like you do, how do you strike the balance between optimizing for what is bringing business value versus researching for the ability to actually do groundbreaking work? I know it's a discussion that has happened often times which is like, how do you balance the two worlds? What is your philosophy around that?

Mounia Lalmas: That's a very, very, very good question and it's a question that many research organization always keep asking themselves, revisit and so on. The way we're doing it... and my answer is going to be maybe very Spotify-specific. We are trying to address challenge that is relevant to Spotify product at the moment, this is what we were doing. For example, how do you optimize for long-term and short-term? We have the support of the product, the engineer, and so on, so we're working with them. So we are trying to make better product, better algorithm, better methodology for evaluation purpose. So we do research to improve the product or to build better product. And as a byproduct of this, we for example publish, also we publish also with our colleagues that are not necessarily in research.

Maybe it is a bit of a challenge, it's a bit of a balance. We are lucky there are a lot of really, really interesting research problem at Spotify, so we can jump on many and this has allowed us to really help on various occasions and so on. So maybe it's a right time the way we work now and our contribution has been very much valued. At the same time, because those are often maybe not necessarily brand new problem but they're new in the context of audio listening. That's why there's a good choice of research that needs to be done and this is very much valued by the product team, by the business. That is our current philosophy, whether it will be the philosophy in two years time, I don't know.

But I can still add that there is this investment in reinforcement learning, we know this is going to be not tomorrow but it's going to be a journey a few years. And what we're trying to understand is where do we want to be for example in five years, which technology and how do we progress towards for example, this vision? So we always try to define the long-term research needs and steps. And this allow us to come up earlier with proof of concept, okay, this is good, this is less good, okay, let's move this way and not this way. And finally is to continually discuss with the business what the research is doing and not just to work on our own with no communication with product team. Hopefully it gives you a little bit how we try to make it work, but it's likely that we evolve as we grow. But at the moment, this is how it works and it's working pretty well.

Rashmi Mohan: I like how you see that. Iterate through the ideas that you have but also get the validation working closely with a product or an engineering team to see if you're actually moving in the right direction, but also fuel the needs of research in this area which is so nascent in and of itself. But I know, Mounia, that one of the other things I wanted to definitely talk to you about was you have a lot of interactions that happen in the community overall, right? You do participate in conferences, you're on various committees. One of the things I had heard about the work that you've done around the initiative for XML retrieval, you co-led that project. I'm just curious, why was it important for you to do that and what do you think is the value that you get from these industry engagement and participation?

Mounia Lalmas: This is going back to me, information retrieval researcher interested in evaluation. Evaluation is very big again. In information retrieval, we have the track initiative which is every year with a number of task people building test collection, how to evaluate and how to compare approaches. So we know how everybody or how the state of the art is advancing in the number of area.

For this one, I like we were discussing a particular problem which is at that time, it was XML was the big thing, everything was going to be represented with the XML format. And then there was this notion, we don't need to return the whole document but we need to return just a bit of the document. And we found that pretty quickly, that the way we did evaluation with precision and recall just didn't work out. And it was interesting, a lot of people got interested into that area like research, they're always going through phases at some point, this is a popular problem people are trying to solve. So we had the opportunity to build a group which was international across the globe, coming both from industry and academia to try to solve this problem.

I work on XML retrieval, we call it focus retrieval, how do I know that my approach is doing well? By doing this help us really to make progress together, so it was not one person deciding that's the way to do it, but they also allowed to build a strong community and more people interested in a particular research area. It will also allow master's student and PhD student to take a topic for their dissertation, contribute to it. We're able also to validate their work and so on. Without that, they would not have been able to validate.

Having a mixture of academia and industry is always good. Industry bring perspective that maybe in academia we're not very much aware of. Those are the constraints we're having, those other questions we are having and so on, those are sometimes the data assets that we're having and it's to try to... again conversation is everything.

And academia often come with very, very strong models but scalability become an issue and all those kinds of things. So having those conversation allows us to really make a good progress in this particular area, but also to bring a strong community of researcher that now fill all over the world in various places, academia and industry. We should also view research as an education, especially people based in academia is growing the next researcher and those initiatives like track and IMX and so on allow also to do that. There's also the education part which I'm also passionate about.

Rashmi Mohan: I can tell just by the passion that you speak of this and the tremendous value that it brings not just to you but the community overall. How does one find these opportunities? How does one engage, whether it's somebody early in their career or somebody who's in industry or academia, what would you suggest?

Mounia Lalmas: Be open. I don't like this word but it is an important word, is find a way to network to understand what are the opportunities. So it's again networking, but networking with a purpose. It's important to know what is out there, identify important research area. And if it's students, this will happen partly by being part of a group in a university or research group. If it's an industry, it will be a combination of what are the needs of the business and what is happening outside and so on. So again, it's go back to this conversation, ask questions, attending important events like some of the conferences and discuss and discuss and discuss and discuss.

And things which we used to do in my early, early, early age is to organize workshops. So if there's a particular area that is of interest, see if it's a big metrics it's bringing in, a bit of metrics, a bit of machine learning, a bit of design and so on, organize workshop around these because then one can bring experts and really also start to build an understanding of what one can do for the career maybe the next one to three years or longer term and so on.

So again, it's a lot going back to this conversation, talking to people and organizing workshop is a great way to really learn a lot because also it pushed us to do this networking maybe in a more constrained way.

Rashmi Mohan: Thank you for those, those are very practical and actionable tips. I'm sure that our listeners will really appreciate. Mounia, what do you do outside of work? What are your hobbies or what are your passions?

Mounia Lalmas: I have mostly two. I used to like doing a lot of weight training in the past but I damaged my back and it took me a while to be able to replace it, so I have started yoga a bit more than a year ago. And especially with the pandemic, you can't do much and a lot of the yoga is all online. So I really took it seriously and this is becoming a hobby to the point that I'm starting reading books about the value of yoga. It's both about the actual yoga exercise, but the kind of well-being spirituality that comes into it. And people who knows me know that I really like Prosecco. I won't call it as a hobby or passion, but it's something that I like very much.

Rashmi Mohan: That's great. And the part that you bring up about wellbeing is so important in these times especially. This has been an excellent conversation, Mounia. For our final bite, I'd love to understand, what are you most excited about in the field of personalization research or information retrieval, the areas that you're interested in?

Mounia Lalmas: What excites me is, so you have information retrieval, you have this whole area which is very related but still different recommender system, you have also voice, how people now interact with our live system. i you put this into ecosystem. So personalization is very much about the user. The user has a need or user wants to get things done or want to listen to something. We forget the

content provider and for example in the context of Spotify or the artists, and so that's what I call the ecosystem.

And all this, it's all related to what I referred to downstream interaction. Interaction is not just the click, it's a relationship in the context of at least Spotify, but it's also elsewhere between the user and content. And there's various way the interaction is evolving. We have now the whole area of conversation. So like now, we're trying to not evaluate a click or an approach, we're trying to evaluate how users interact with content during that journey and this is just fascinating because success is not just now, it's a success of a journey.

And I'm looking very much forward and I'm already starting to look into this. What does the success of a journey mean? And this is super, super exciting at least for me.

Rashmi Mohan: I think what you say is relevant of course in the field of personalization research but overall as well, there's so much depth of that statement, success of a journey. Mounia, thank you so much for speaking to us at ACM ByteCast, we totally enjoyed it.

Mounia Lalmas: Thank you very much.

Rashmi Mohan: ACM ByteCast is a production of the Association for Computing Machinery's practitioners board. To learn more about ACM and its activities, visit acm.org. For more information about this and other episodes, please visit our website at learning.acm.org/bytecast. That's learning.acm.org/bytecast.