Juan Miguel de Joya :

This is ACM ByteCast, a podcast series from the Association for Computing Machinery, the world's largest education 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, Juan Miguel de Joya.

At its core, search is about helping people find what they need in a world overflowing with information. So it sounds simple, but it's actually a deeply difficult problem. How do you understand what someone is looking for, sift through enormous amounts of information, and decide what is most useful, relevant, and trustworthy in that moment? And once you think about it that way, you realize that search is pretty much everywhere. It's in the questions we type into a search engine, the videos and articles recommended to us, the products we discover, the maps that guide us, and increasingly the AI tools we turn to for answers.

Behind all of these everyday moments is a deep field of research called information retrieval, the science of helping people find the right information at the right time from an overwhelming amount of data. Few people have helped shape that field as deeply as Ricardo Baeza-Yates. Ricardo is a Chilean computer scientist whose work in information retrieval, algorithms, web search, and data mining has influenced how modern search and recommendation systems are built and understood. He is currently the search chief scientist at You.com and holds part-time professor appointments at KTH Royal Institute of Technology in Sweden, Universitat Pompeu Fabra in Spain, and the Universidad de Chile. His career has moved across academia, industry, and entrepreneurship. He previously served as vice president of research at Yahoo Labs, holds 14 patents, and has co-founded several startups in Chile and Spain, including Theodora AI, which focuses on mitigating technological bias. He has earned his engineering degree and master's degree in computer science and electrical engineering from the Universidad de Chile, and his PhD in computer science from the University of Waterloo.

But Ricardo's impact goes beyond technical foundations. He has spent his career mentoring researchers, building institutions, and helping expand the visibility of the Latin American computing community around the world. And that combination of pioneering research and commitment to lifting others up is what has led to his recognition with the ACM Luiz André Barroso Award, which honors fundamental contributions to computing from researchers in historically underrepresented communities. So it's my honor and real pleasure actually to spend time with you again, Ricardo. It's been quite a while.

Ricardo Baeza-Yates:

Thank you. So long without seeing you.

Juan Miguel de Joya :

Yeah. When was the last time we saw each other? Was it ITU?

Ricardo Baeza-Yates:

In Geneva, in the AI event that they do every year.

Juan Miguel de Joya :

Oh yeah, that's true.

Ricardo Baeza-Yates:

Exactly, Yes.

Juan Miguel de Joya :

Yeah. Will you be going this year?

Ricardo Baeza-Yates:

No, no, I only went that year when I met you.

Juan Miguel de Joya :

Oh, okay. Yeah, that's a lucky year.

Ricardo Baeza-Yates:

Yes.

Juan Miguel de Joya :

Yeah. So before we continue, I wanted to ask you, I know we talked a bit about your background, but would you mind introducing yourself and talking about what you're currently doing?

Ricardo Baeza-Yates:

Yes, no problem. So currently I'm sharing my time between industry and academia, mainly industry in You.com, as you mentioned, in San Francisco, but also I do have these part-time commitments in several places. And physically I have to be a few months in Stockholm because of my KTH commitment that was even before I started working with You.com. And so I'm right now in Stockholm doing research on responsible AI. So mainly with my PhD students, mainly in Barcelona. And my work here in KTH, touching different problems on responsible AI, maybe the main one is evaluation and looking at how to evaluate harm and not success, because things should work by default. I mean, why we evaluate when things work? We need to think what happens when things doesn't work and how much harm that creates, because all errors are not equal. So there are errors and they're more harmful than others. And we are looking at that. So different ways to minimize critical errors, and also to try to measure. And because of regulation of the use of AI needs that. For example, the EU regulation is based on harm and risk, but really we don't know how to measure risk too well yet.

Juan Miguel de Joya :

I totally agree, and I think that's a very fair point. I think one of the challenges we are living in now currently in this space is that both the technology for discoverability is so pervasive. We take it for granted that search did not exist decades ago. And so the work that you were doing in that space is really vital, but at the same time, that space is also evolving in conjunction with a lot of different technologies like machine learning and how that impacts our ability to disseminate and understand information.

Before we go in further, I think one thing I really wanted to ask is that sometimes when we have audiences listening to this podcast, they may not necessarily completely understand all the technology. And so I was very curious if you could explain what makes both either search or responsible AI difficult as a topic to tackle.

Ricardo Baeza-Yates:

In some sense, I'm combining the work of those two fields because at You.com mainly does search APIs for AI agents. And what you want if you want to succeed with AI agents or what is called the agentic AI, is to have relevant and true information. And this is not easy if you use a language model that will predict everything. And in these predictions, some mistakes can occur. I don't like this word hallucination because really they're not hallucinating, they're more making mistakes. Sometimes mistakes are created, and it could be interesting mistake depending on the context. But if you make a mistake in a legal case or in a medical case or even in other settings related to people, this could be really harmful. And then you need to combine responsible AI, for example, to find the right information. And that's why we're building the best possible search APIs for agents to make sure that they have the correct information to start with.

Of course, on the way down, mistakes can happen, but at least starting with true information. And today it's very hard to check anything that the chatbot will give to you because you're not an expert. So if it's well-written and looks like correct, why it's not correct? And very few people will find the mistakes. And this happening all over again in many contexts, in many news, and many people of course don't mention it, but I know cases every day about this problem.

Juan Miguel de Joya :

I can totally see that happening. And we're seeing some of that also appear in the news. There's very different cases just going back even years from now since even when we first met, where you're seeing how we are beginning to really grapple with the impact of the technology. Is there something that you think about search or AI that we as engineers would know about, but that the public usually misunderstands?

Ricardo Baeza-Yates:

I think even some computer scientists don't understand very well the limitations of data, machine learning, and even us, because I think the most important problem is not the technology, it's us, how we use it. So for example, if we don't understand that really the system doesn't understand anything and it's only a very good prediction, we run into trouble. So the first problem I think that's true for everyone is that we humanize this technology. We talk about thinking, reasoning, seeing, writing, reading. And for all these things to work, you need to have real understanding. So first we should use the right language. So we should use, "Okay, this system process data, generate data," but that's it. Whenever you have a system that looks at a cat image and transform the cat image to the word cat, it's only changing the representation. It's not understanding what a cat is like we do.

So I think this is even within some computer scientists, it's not well seen because they use words that they shouldn't use to explain how this system works. So we should be careful on that. Weizenbaum, 30 years ago, the person that designed the first chatbot, ELISA, said we should never confuse the computers with humans, but today we're doing it every day.

Juan Miguel de Joya :

Right. And it is a bird line nowadays, even with search, for example. Let's go with search. Search as a field has gone from keyword matching and ranked links to semantic search, neural systems, RAGs or retrieval augmented generation, and AI assistance that synthesize answers. And are we still doing some sort of information retrieval when we're giving synthesized answers by AI? Are those valid information retrievals, you think?

Ricardo Baeza-Yates:

For me as an information retrieval researcher, I would say no. I know that many people is moving away of that standard of search engine that are using AI summaries because AI summaries many times are wrong. And some people look only at that. Well, some people even use chatbot tools instead of search, and there are many consequences there. One is that they're not really searching, they're predicting information. That's one problem. But also, for example, they're using 10 times more energy than searching. So they're wasting energy. So I think that's another concern I have, the use of resources that are important for us. And there we need to be fair, it's not only chatbots. I mean, we're using a lot of energy with TikTok, Instagram, and many other apps that are helping us to waste time, but also they're wasting energy. So I think many people is not conscious of that, and someone is paying that. Probably you are paying part of it, but I think we are not paying part of all our waste of time that we do every day.

So I think we need to be very careful on this balance on how to use these tools for the right thing. For example, I believe this can be good tools to translate information, to help you in other languages, to summarize information, to extract information. But if we use it for just writing, basically we have all this problem of cognitive delegation, cognitive offloading. Basically, if we do it, it's not too much trouble because we have learned how to do it. But for young people, that they haven't learned yet to how to do it, I wonder what will happen in 10 more years when people cannot write anything, not because they don't know how to write, it's because they don't know how to think, which this is the important part. They will be like, I would call it cognitive zombies. It's not evolution, it's involution. And I wrote a paper about that, about how we are not evolving in search because we are using too much AI for the wrong things.

Now, if we combine AI the right way with search, we can get very interesting solutions. For example, if you have relevant information and you want to summarize it, then a language model can create the best answer possible, but you know that it's all true because it started from true information. But when you start predicting information, then you run into trouble. And this is something that it's very hard to fix because it's part of the architecture. The architecture was designed to predict the most probable text, was not designed to look at knowledge database and give you the real fact.

Juan Miguel de Joya :

So one of the things that's really interesting now obviously, I think, is retrieval augmented generation or RAGs. So for folks who do not know what a RAG is, to put it simply, high level, it means that an AI looks up the relevant information first and then uses that information to give a more accurate answer instead of relying what it already "knows" based off of the model. And to me personally, it does sort of connect modern AI systems back to information retrieval. Would that fit what you're envisioning, or would you see it in a more different direction?

Ricardo Baeza-Yates:

Yes, but you know that even using RAG, the system sometimes make up answers. So this is not enough. This improves. So mistakes are less probable, but because of this prediction approach, this mistake can even happen. So there are some weird results show that if you put some random results on the RAG, the result may improve. So this is counterintuitive. If you put things that are not interesting, then maybe the signal-to-noise ratio improves and then you get a better answer. So we still don't understand how systems that have not one million, but billions of parameters really work. We can build them. We know what is architecture, but really we cannot imagine in our minds how these things interact when you're using one billion formulas to compute something, right?

Juan Miguel de Joya :

That's right.

Ricardo Baeza-Yates:

No, very hard.

Juan Miguel de Joya :

Absolutely. And I think as search becomes more semantic and more personalized and AI generated, those stakes are going to just get higher. To your point, these systems don't just find information, they can shape what people see and trust and believe. And that makes a lot of the work that you're doing with biases and search and recommender systems and responsible AI especially important now.

Ricardo Baeza-Yates:

Yeah. One problem that really worries me is not only the bias, it's the cultural colonization behind the systems. So my standard example, if I asked you how many continents there are.

Juan Miguel de Joya :

Yeah. Are you asking me?

Ricardo Baeza-Yates:

How many continents there are for you?

Juan Miguel de Joya :

I'm pretty sure there's seven, but is there a new one that I didn't know about?

Ricardo Baeza-Yates:

No, but you're studying in the U.S., right?

Juan Miguel de Joya :

Right.

Ricardo Baeza-Yates:

That's why for you it's seven.

Juan Miguel de Joya :

Yeah.

Ricardo Baeza-Yates:

But in the European tradition, it's only six because America is only one.

Juan Miguel de Joya :

Oh, okay.

Ricardo Baeza-Yates:

The American culture is already colonizing the rest of the world. That's why, for example, the Olympic flag has only five circles. And Antarctica is not there because it doesn't compete in the Olympias. But for the U.S. and Canada and many other countries now, in Latin America, for example, it's not seven because North and South America is two different continents. So that's a simple example of cultural colonization, but there are more complicated examples where you don't see all the possible points of view. If we want to give true answers, we need to give more than all the possible points of view that are valid. Some topics are not about truths or not, they're about beliefs. And for example, if people believe we should control guns or not, you should have both views and so on. And that's happened with many, many problems. But you will not get all possible points of views because you are trained with biased datasets. And also you have guardrails against maybe some thoughts or some beliefs that you think are wrong, but maybe another culture don't think the same. We are getting this colonization, the next digital issue.

Not only that, I think together with this is that the best language model supports around 200 languages, but we have more than 7,000 languages alive. They did some analysis, and 10% of the people in the world don't speak one of these supported languages. Now, if you add that all the people that don't have internet, all the young people that cannot use a computer or a device, all the old people that will never use it, and if you do the right overlaps, you get between 50 and 55% of the people that cannot use these technologies. So when people say, "We are democratizing AI," I really hate it, because we are not democratizing AI, we are increasing the digital gap, and much faster than before between people that cannot use this technology and people that can use it. Now, maybe the winners, and I will try to be provocative here, maybe the winners will be those people because those people will keep thinking, they also have more privacy. So maybe they're the poor people today, but if these things don't go well, they're the future of the world.

Juan Miguel de Joya :

In a very good perspective, actually. I think people won't see our faces on this, I imagine, but we're both from the Global South. And I think we have had these conversations on the discrepancies between what technology means for people in developed countries versus developing countries. So what I'm going to try and do actually in this conversation is let's take a step back from the technology aspects for a bit.

Ricardo Baeza-Yates:

You are from the Global South, but you are from the North Hemisphere?

Juan Miguel de Joya :

I was raised in the North Hemisphere, it's true.

Ricardo Baeza-Yates:

But which country is your Global South?

Juan Miguel de Joya :

I was born in the Philippines.

Ricardo Baeza-Yates:

That's what I'm saying, but this is in the North Hemisphere.

Juan Miguel de Joya :

Is it the North Hemisphere?

Ricardo Baeza-Yates:

Yes.

Juan Miguel de Joya :

Obviously this is a good example of where the technology bias comes in.

Ricardo Baeza-Yates:

Yeah, but I say it because this is just a convention to talk about the Global South because most of the developing countries are in the South, but even includes Mongolia, which is north of Philippines. So-

Juan Miguel de Joya :

Yeah, that's absolutely true.

Ricardo Baeza-Yates:

... it doesn't make much sense for me, because I love geography. So [inaudible 00:17:06]. But remember the equator goes through Indonesia, which is south of Philippines.

Juan Miguel de Joya :

Okay, so that's definitely a bias for sure, because they told me, "Oh, we're so close to the equator," but you're absolutely right. Let's amend that, and you're absolutely correct. I think to your point as well, just from the language that we've used, Global South is a misnomer to some degrees when we talk about developing countries. But I really want to step back and talk about your origins, it informs a lot of your opinions. And to understand why those opinions exist, if I understand correctly, you started out in Chile and then moved in computer science. Chile's not necessarily at that point in time, if I understand correctly, not one of the traditional centers for computing. So I was just curious, what incentivized you to get into computer science actually?

Ricardo Baeza-Yates:

That's an interesting question. I think I never thought about it.

Juan Miguel de Joya :

Really?

Ricardo Baeza-Yates:

Because I never plan things, so I take opportunities. For example, for me, I think there is another reason that maybe was the main driver of my career is that I learned to read and write very early by my maternal grandfather. So that opened my mind to the world. And then I had a great amazing British teacher in primary school that had a course called General Knowledge, because she was not a teacher, but she wanted to teach. So she teach general knowledge. And this opens even more the world. General knowledge includes many things. So this was a very interesting class for me. You will not know what you will learn every week, and this is nice.

And then, I think the teachers help a lot on the site what you will study. So I think a person can do many things, but depending on the best teachers, they will drive you to that path. So I had very good teachers in math and physics in high school. So I decided to be an astronomer or an engineer. So I started engineering. I had never seen a computer when I entered university. And then I had to do the first course on programming. I discovered that I love the logic behind algorithms, and really the combination of math and logic captured me. I was studying electrical engineering. I did finish that, but really I realized that what I really liked was computer science. And then I had an amazing professor of algorithms, maybe that's the reason I started computer science. So I did the master with him, Patricio Poblete. I think he was one of my best teachers I had because he did his PhD at Waterloo, and other people in Chile did it for historical reasons. I went there, I got a scholarship. It was not easy at that time to get scholarship, especially from Chile. And I found also a great supervisor, Gaston Gonnet, that even make me love more algorithms.

And they were working on the Oxford English Dictionary project. So basically at that time, that was one of the largest files was 570 megabytes in 1986. So that was a very large file. It didn't fit in one single CD-ROM. At that time it fit in four CD-ROMs. Later, the CD-ROMs improved and then you could fit into one. But the problem was how to search that. And I started with sequential search algorithms in my PhD thesis. And then naturally when the web arrived, I finished my PhD in '89, so when the web came in 91, 92, 93, although the idea is from 1989, naturally I said, "Okay, I can apply all these search skills to the web." And then I asked to do information retrieval, which is basically searching algorithm, but using an inverted index, that's the classical approach.

And then I started to work on that with a Brazilian friend, which was a friend of André Luiz Barroso, because he worked at Google too. And Berthier Ribeiro-Neto, we decided one day to do something crazy, "Why we don't do a good book referring to information retrieval? Because seems that there's no modern book." And they said yes. And that became in '99, this Modern Information Retrieval, the most cited book on the topic. So that's not a standard scene, that two South Americans write the most cited textbook in a topic. This is not true, I think in any other topic. So sometimes for crazy ideas, you get interesting results. So I think this was the beginning. And then that's maybe because of that reason I became the VP of research of Yahoo in 2006 when I started the lab in Europe in Barcelona. And after 10 years, Yahoo was sold to Verizon, so I left that. I became CTO of Intent, which was another semantic search company. Then I went to be director of research of the Institutes for Experiential AI of Northeastern University. Interesting, my manager there was Usama Fayyad, which was the creator of Yahoo Labs, and he was my skip manager at Yahoo.

And by the way, thanks to Prabhakar Raghavan that works at Google. I started in Yahoo, so I should say thanks to him too. And then when the things didn't went very well for responsible AI because of current governments, not mention any political thing, I decided to move to, again, back to industry and also to do my responsible AI research in Europe, where it's more appreciated than in the U.S.

Juan Miguel de Joya :

You've had a very interesting journey, I would say. I didn't know all of it, so it's me also discovering it with everyone else. And I find it incredibly fascinating that you seem to say that things are just happening clandestinely. But I'm very curious if the experience you have growing outside of where computing was happening, a lot of that research influenced how you approach the application of responsible AI and how it approaches where your interests are right now.

Ricardo Baeza-Yates:

As I said before, I never had any plans. So this is not a planned life. That's better than a plan, because I think plans limit yourself. Maybe if I had a plan, I would be in Chile, but because of that, I'm not in Chile, and that's good and bad for the country, because I can represent Chile abroad. But I think when you just take opportunities, and for me, any change is an opportunity. Even when it looks like a bad change, I think it's opportunity to redesign yourself. When I was at Yahoo, I was worried. I seen bias had been a topic that worries me always. And I've been working on bias now almost 20 years. So before this was popular. For example, my first paper on bias, which is about the genealogical tree of the web was what's the relation between search and the content of the web. So how search changes the content of the web, how it shapes the content. Because if all people take the top pages in a search answer to write something else in the web, then basically you are reinforcing whatever is the ranking of the search engine. And then the search engine will later say, "Okay, I was right. These are the best results." But basically it's a vicious loop. You are giving to the people what to use to produce content.

And it's very hard to mitigate that bias because it's very hard to tell people, "No, don't look at the top 10 results, look at the 100 results, look at all of them and choose the best information." People don't do that. They click on the two, three first results and then they take something from there. That first bias, which is called ranking bias, is mitigated by all search engines. So if people click more on the first position, partly because it's in the first position, you need to mitigate that, otherwise, you fool yourself and the rich get richer and the poor get poorer. But the second order bias that affects the content is much harder. And we show that this was happening, that basically there was this thing, and that was in 2008. And almost 30 years later, I think this is even worse, but we cannot measure, because at that time, maybe there were 100 billion pages in the web. Now there are more than a trillion, and they're mostly dynamic and it's very complicated.

And now there's a lot of AI-generated content, which is the next problem. And something that is not true maybe will be amplified. And maybe in some topics we have more false content than true content. And then any model trained by this data, we're seeing that the false content is the right content because it's the dominating content. And we don't even know how much of this is happening, and we don't even know in which topics that is happening. It's so complicated because the number of fake news and fake content today is much larger than before. And of course it's much larger every day.

Juan Miguel de Joya :

No, it's absolutely true. I guess when we're talking about generative-AI systems nowadays when we're using the application, even what we're seeing in terms of authentication research and application still needs to catch up with some of the advances as well. And going around those effectively, do generative AI systems make bias easier to detect if the output is explicit? Or do you think it's just becoming harder and harder nowadays?

Ricardo Baeza-Yates:

Yeah, that's a very good question. I think it becomes harder and harder because this system also have biases to themselves. For example, there are paper that showed that if you write your CV with ChatGPT, ChatGPT will find that better than if you write it with Gemini or Claud, they look like they recognize their generation, they prefer these systems. And of course, if people is using, for example, ChatGPT to evaluate people, which I don't think is a good idea because data doesn't represent well people, in fact, any problem where people is involved, data is not a good representation of the problem, you get a vicious loop again. So okay, the system chose mainly people that use the same model to write their CV, for example. And this is the case of CVs, but in many other cases it's the same. For example, there are no good tools to detect if something was written by a model or not.

In fact, a few years ago, we wrote a paper saying that one regulation that exists is that any model that is published should come with a tool to detect if that model did this text or not. And you can do it with digital watermarks and other things, so you can do that. And this will be very important because then you can say, "Okay, this was not written by human or partially written by human, but was mainly AI written." And this is important in many contexts like exams, CVs, and news and so on. So we don't have that. And we need that more and more every day.

Juan Miguel de Joya :

No, it's true. I do think it's funny, when you look at social media and you can totally tell when something has been written by ChatGPT, there's a certain cadence in terms of how something is written. There's always three points, but then they say something like, "But honestly, blah, blah, blah."

Ricardo Baeza-Yates:

This is true for everyone or only for us that they can detect that? This is for me what worries me, some people like that.

Juan Miguel de Joya :

Yeah, some people do. I see it on LinkedIn as well. So I mean, I'm like, "Oh, okay." Well, for some people, I guess there's an ease of efficiency to it. But at the same time, again, for us, there's a cognitive aspect of us where we're like, "Okay, maybe we can tell that's from ChatGPT," for example. But would you say there is one particular type of bias or some biases in search or in these type of things where us as users would never really notice?

Ricardo Baeza-Yates:

Well, I think the second-order bias that they explain, people will never notice. Basically that people is reinforcing the ranking beliefs of search engines because of this use of content coming from search to create new content in the web. This is a loop, and we prove it 18 years ago, and today it might be worse because now we'll be combined with generated content that maybe not even true. Well, always there's some content that's not true coming from people, but now generative AI amplified that by a factor of, I would say 1,000. I don't know what is the factor, but it's so easy to basically write a blog with your name, say whatever I want to say. And it's so easy. And people is doing that, especially in politics. In 2024, more than 100 countries had elections that fake news played a role. Sadly in country that very close to my heart, people that lie during politics and invented lies, and that lies were enough to convince people that if they were true, was complicated, I think that creates fear, people are easily manipulated.

Juan Miguel de Joya :

No, it's true. And we can see how bias and personalization can bring us into basically the feedback-loop problem, right?

Ricardo Baeza-Yates:

Yeah.

Juan Miguel de Joya :

Where a system can show us information, we can click, we watch, we ignore or share stuff, and then the system learns from that behavior and reinforces that behavior. In some ways that could make some

recommendations useful. Ideally, that's where we wanted to go with the technology, but it can also narrow what we see or amplify existing patterns.

Ricardo Baeza-Yates:

And one important problem related to that, I think, I forgot, because I think I mentioned all my worries, but one worry we haven't talked is mental health, because if you have a cognitive issue, if you have a mental health issue, you cannot distinguish the reality basically from the fiction. And then you have all these teenagers that using chatbots, being helped to commit suicide, or even we already have two mass murderers that were helped by chatbot on the planning stage. And this is the beginning of something that is much larger. We have millions of people basically interacting with this system, thinking that they have a digital friend that will know what is better for their future.

Was a recent study where they asked many people to basically interact 30 minutes with these tools and then decide if they follow their advice or not. And 70% of the people decided to follow the advice. And of course, after some period, they were not better because the advice was without context. After half an hour, you don't know a person, you don't know what is the right context. And basically in many cases these people were worse than before, but they trusted the system because we have an imposter syndrome with AI. We think that AI is better than us because they have processed the whole web, they must know more. People follow this advice. And I guess the best case is the Adam Raine, the teenager that died in California last year, 16 years old. The system mentioned six more times the word suicide. The system sent 300 warnings to OpenAI. And nothing was done because there was no human interloop. Although I think the human should be on the loop, not in the. He even showed the rope and said to the system, "Is this rope good enough?" And the system said yes. And he hung himself and he died.

This is an example, but this is the ones that we know. But how many? Maybe we don't know because not everything goes to the news and also not everything is recorded, it's found by someone. Maybe there are many cases that no one has found that are related to a bad use of chatbot. Of course, chatbots are not responsible, this is a person doing this. If the system sends 300 warnings, how many warnings do you need to send to stop the interaction? This is my problem. You should do something. And after this case, OpenAI did parental controls, which they should have done earlier, because this is obvious. I mean, you cannot have a teenager using this without any very specific guardrails to avoid bad usage.

Juan Miguel de Joya :

ACM ByteCast is available on Apple Podcasts, Google Podcast, Overcast, Podbean, and Spotify. If you're enjoying this episode, please subscribe and leave us a review on your favorite platform.

That's a really interesting point. And I guess there's a question here for me at least where we're looking at these systems, let's say the search system, let's say the machine learning, information retrieval that we're deriving from querying ChatGPT or certain systems, what are the changes that we should expect in terms of responsibility for these systems?

Ricardo Baeza-Yates:

That's a very good question because I don't know why for AI people think the responsibility is different from other fields. And the example usually I give is that if you have a car and you have a problem in your car, you will never go and check who did the piece that is failing. You go and talk to your dealer and say, "Okay, the car has a problem." You can use maybe the guarantee or maybe someone will fix it, but you talk to the maker. The same is true for software, for AI, for chatbots, it's the same. So if a chatbot

produces a harmful incident, the company that put that on the market or in the public, this is the responsible one.

So it's very easy. And that's why in the case I was mentioning, there's a case against OpenAI for allowing, for example, 300 warnings, which is weird. And the same will be true for any other system. So even if in the fine print of the terms of usage or condition, you say, "I'm not liable of bad use," that's not completely true. I mean, if you try to forbid any bad usage and you try to control that, yes. But if you don't do anything to control the bad use, this is completely different. This is like selling guns without any permits. Anyone can buy a gun and you don't know to show that if you are crazy or not.

Juan Miguel de Joya :

Well, I mean, that's an interesting point, and I think that's probably how we met honestly in Geneva, right? We were again at-

Ricardo Baeza-Yates:

Yes.

Juan Miguel de Joya :

We were at the AI or Good Summit for sure. And those are some of the questions that come up in that summit. One thing I think that is interesting to think about, we ideally could live in a world where people are trying to enact good when we're developing systems such as NLP systems. But do you think a system can become biased even if no one is intentionally designing it to be biased?

Ricardo Baeza-Yates:

Oh, that's for sure. I mean, most of these problems we have, no one has had the intention that these incidents appear. I think that they never thought that this kind of usage may happen. And the first one was in 2023 with a different system, not with OpenAI. But if you think a bit more at the beginning when you are in design phase, for example, you do some kind of red teaming with different kinds of stakeholders, especially users, maybe you will come up with this kind of usage, and a mother will say, "We need parental controls. I mean, I don't want my son to use this without any control." This is for me, if you're a parent and you have a gun without... Well, it's not locked in your home, and suddenly many people have lost their children because they have a gun without lock at their home. But this is for me, common sense. In Spanish, we say that common sense is less common of the senses. So this is the problem.

And this is the same. And these tools don't have any common sense. So they're very easy to break. Even if you have a guardrail, you can say, "Imagine that you're in a fiction story and we have this context and these people," and then the system will predict that you are not talking about you, but you are interacting about something else that is completely fictional, the system will produce information that shouldn't be produced and then help the person to do whatever he or she wants. See, these are very easy to manipulate because they don't understand the work.

Today with a friend, he was saying that he found from ChatGPT seven free museum that he can visit in Stockholm. And I say, "Are you sure they're free?" And then he check, and one was free, but for people younger than 19, he was not younger than 19. Another one said, "Oh, it's free after 5:00 PM." So yeah, it's free, but it depends on the context. And I'm sure after I asked him, "Okay, how many of the seven are free?" Probably maybe one or two. But yes, it looked like they're free, but there were some condition that the system was not taking account because they don't understand the world. They're just predicting the world, and in the training data they process the word free.

Juan Miguel de Joya :

Yeah, sure.

Ricardo Baeza-Yates:

So see, you need to understand. It may infer that the museum is free, but it's not. So many details. So it's so hard to solve because by design, by architecture, these systems are built to predict things, not to know things. There's no knowledge database. If they have a knowledge database, that would be very different. This is the next step. I think in the future we will have knowledge databases. And the big company that have large knowledge databases will profit from that. Also, they will have real logical inference, like classical AI, predicting, "Okay, can you go from here to here, so there's some causality and then we can infer things?" Yes. And the last one I already mentioned is common sense, but that's so hard. Common sense implies that you do the right thing even if you don't know it. I guess it's a hardwire in our brain from our genetic history. And if you see something wrong, you know what to do, most people. The people that don't have common sense, they die, but other people that have common sense, they save themselves because they do the right thing, "Okay, we need to run this way."

Juan Miguel de Joya :

No, totally, that's such an interesting point. And I agree with you. I was listening to a talk that Ken Perlin gave, and he talked about virtual reality, but the way he described how children learn is completely different. The semantic language, the things that they learn in terms of communicating with others is very different because of the advent of the technology existing during the period that they were young. For us, we did not have that technology. So even what common sense is or how we communicate in a common sense way or a pragmatic way is completely different based on how the technologies impact us. And I'm just curious, when you were starting out with the research that you were doing in information retrieval, did you ever imagine that we would be in a world where we're using this technology in this way?

Ricardo Baeza-Yates:

No, no, I don't think so. Some funny anecdote, we invited Don Knuth for famous student award, my role model in algorithms, to give a Q&A in the Latin American conference two years ago. Of course, he did it online because he doesn't travel now, he's more than 80 years old. All the people went to talk to him about AI. And he didn't want to talk too much about AI because, I guess he didn't like it or maybe he didn't know enough. But something he said was very interesting because he loves algorithms. I try to quote him exactly. He said, "I never thought that we will use algorithms that we don't understand." Completely right. I mean, we know how to build them, but we really don't understand what they are doing. And this was interesting thought, that we are losing control. That we can do so amazing things that predict amazing text most of the time, but we are losing control of the output. So before, I guess we could predict the output, today we can't.

In a classical algorithm, you can say, "Okay, what will be the output?" And you can tell the output and then you can check it if it's right or wrong. The numbers will be sorted and then you check that or whatever. And you will find the second largest of the set or whatever, so the classical outputs. But today we cannot do that. And maybe I never thought about that, basically that verification, we passed from the problem of solving problems to the problem of verifying solutions. So this is a change. For example, that is happening from, for example, with [inaudible 00:41:07] like, "Okay, you have all these possible security problems, and now we need to verify how bad it is, how you can fix it, if the fix that the system provides is correct and so on."

The problem before, we had too many problems to solve. Now AI will do that, even in math. And there were some interesting results in math recently. Even Knuth wrote an incident report about a math problem that he proposed that AI found solution. But now we don't have enough people to verify these things, like, "Okay, here there are 10,000 security issues. Here there are 100 CRMs from Erdős problems. Check if they're correct or not." So it's interesting, we still need humans, but in a different phase, and in a phase that in some sense may be more complicated, because if we put the level of knowledge higher, the verification will be more complicated because every time we're doing more complex things. Maybe that will improve humans. The best computer scientists will not be the ones that can program well, will be the one that can do a great architecture, a great orchestration, a great integration of agents, and mainly a great validation and verification of results.

Juan Miguel de Joya :

No, that's an interesting point. I would say if I were to step back, and I'm going to ask you this question just because we're on this topic, from your perspective, what do you think a healthier and more accountable system for an information feedback loop would be, let's say in search or in AI systems?

Ricardo Baeza-Yates:

I think we need to keep the classical search. So basically we need to know when something's relevant or not. We need to do fact checking, and for that we need to have a knowledge basis. The problem is that knowledge is also bigger and bigger. And part of this knowledge can only be verified by very few people. For example, things that you have done in your life, you are the only one that can verify that fast. And maybe if I try to verify that searching the web, there are half of them I cannot verify because they're not in the web. So this is the problem, the knowledge is becoming much larger, and also the people that can verify that knowledge is becoming scarce. There's one people, two people that can verify that. For example, about your life, maybe you are the only one. There's no one else in your family that knows everything about you. And that's true because they were not with you, they didn't have the same experience and so on.

So this is, I think, the problem. How we verify information is that we need to build different ways to verify information, and maybe we need to work better on some different kind of relevance. This is a research problem, I think, we don't know how to build relevance for this new world of agents that basically will generate something that looks true. Now, one very easy experiment is that if you have something that you are not sure, you can ask the same to many different models. And if they all disagree, probably no one is giving you the right answer. But if they all agree, probably they're giving you the right answer. The problem is what happens if you have partial agreement, maybe this will be the level of confidence. Okay, how many models agree will be the level of [inaudible 00:44:17]. But it still could be a lot of biases in training data, could be a lot of false information.

There's a very interesting example of a person that did the following. This person was a female researcher that basically published a few articles about an illness that didn't exist. And the paper, if you read it, clearly is not a good paper. The paper says it's fake, but this was basically used by all models. And if you ask about that illness, you will get an answer. And the answer is wrong because the illness doesn't exist. But they process these papers, something like bixonimania, and it's something that says that you get the red eyes because of watching too much a computer screen, from blue lights. This doesn't exist. But people will believe that exist if someone asks, maybe we can do the experiment later, "How do you call the illness of getting red eyes because of blue light?" maybe they fix it, but how many things like this are already learned by this system? It's not one, it's many, but you don't know them.

Juan Miguel de Joya :

Right. No, that's true. And that's a really interesting anecdote. I was just thinking, "Wow, I should put some eyedrops on." Yeah. No, these are very interesting points. And I think the question I have, just for anyone who's listening actually, who might be just interested in the field, be it AI, be it just classical information retrieval, what would you tell someone who wants to build these powerful systems, but also just wants to avoid causing harm?

Ricardo Baeza-Yates:

I would say first that you need to understand very well how these systems work. So you need to understand the limitations of data, for example, that data doesn't represent well people, the limitations of mature learning, that these systems don't really understand, but they predict things. And I gave a talk, a keynote at SIGMOD, and later maybe it's in YouTube, about the limitations of data mature learning and us. And also all the about uses of this, because if you learn examples that can make harm to people, maybe you will stop using them. But then if you really want to use this system well, I will say let's use the ACM principles for responsible algorithmic systems. I was one of the main causes of these principles together with Jeanna Matthews. For me, the first one is the main one, show that your system is legitimate and also you have all the competence to do it.

What means that the system is legitimate? Well, it's ethically sound, it's legal, and also it's based on science. There are a lot of pseudoscientific applications basically predicting things that really are not related to the data. And then competency means you know well AI, you know computer science, you know the domain expertise of the problem. For example, if it's a health system, you have doctors working with you and so on. And very important, you have the permission to do it, because many problems have been because of people doing things that they shouldn't supposed to do, because sometimes people think, "Oh, this is a great idea, let's do it," but someone else has to authorize them to do it. And I think this is the case of the childcare fraud model that was used in the Netherlands, that then implied that the whole government resigned in 2021. Probably the engineer that did that and the Ministry of Social Affairs never asked if the system was okay or not. They did it, and basically it didn't work.

Juan Miguel de Joya :

No, I mean that's very good advice. And to be honest, I think useful for all of us when we're thinking about being responsible professionals actually in computing. So-

Ricardo Baeza-Yates:

Yeah. And these principles are also translated to Spanish and soon to other languages so more people can use them, and they can find it in the ACM website.

Juan Miguel de Joya :

Nice. It's always good to know. It's really funny, because I latched onto something you said a while ago where you said you're representing Chile and the Latin American community outside of the Latin American community. But to be honest, you have had an impact on not just the representation side, but also building research capacity and technology capacity in Latin America. So that includes the Center of Web Research at the Universidad de Chile, and mentoring, I think about 34 PhD students, many from Latin America. So I wouldn't actually discount your impact on people and the community. And so I was just curious, just from your perspective, what does Latin America bring to computer science that the global field hasn't taken a look at yet?

Ricardo Baeza-Yates:

So another thing that I'm proud of, my 34 PhD students, half of them are women. So that's part of the-

Juan Miguel de Joya :

Oh yeah. Nice.

Ricardo Baeza-Yates:

... gender parity, affirmative action regarding bias. I think Latin America can help on giving a different view of things, because sometimes you have this first-world view of things that basically you think about the problem that is not a problem of the whole earth, it's a problem of a given rich country. You lose sight of the real problems. What I have said earlier that they say, "Okay, we're democratizing AI." "Yeah, but you're democratizing AI in your country where all people can use these tools because they have the same language and they have money and time to use it." But in other countries, it's completely different. So if you go to one country where the language is not supported and they don't have very good internet, that means that only maybe 20% of the people that speak English, if you're lucky, can use these tools.

This is something that you shouldn't forget, that you should look at the problem from a global point of view and not from a local point of view. So I think that's important. And from Latin America, I think we know that because we suffer that. Basically we are overlooked. I was the president of the Latin American Center for Studies that basically is the association of all the departments of computer science in almost all Latin America, and that's more than 100 institutions from Mexico to Chile. So that helps you to see the differences, to see, for example, just to say one thing. So I think there are more differences between Latin America than between the U.S. and Chile. So sometimes things are relative. So the levels of, let's say, number of PhDs in computer science, the quality of the research, if you put that, you will find more inequalities in the same region rather than with the rest of the world.

And it's something that I think in the U.S. and Latin America, the same thing, everywhere is the same, which is not true. I mean, one of the most important things we have on earth is the diversity of people, diversity of languages, diversity of opinions, and talent is everywhere. So we can bring some specific talent. We are not too many. Chile is only 20 million people, so nothing. And even the countries even smaller that produce great people, let's say Uruguay and Costa Rica. So we need to make use of that diversity. I think part of the power of humankind is the diversity. And suddenly some people doesn't like diversity. And this is a single problem for all of us because our future depends on diversity. Resilience depends on diversity. Real innovation, good innovation, remember innovation is not always positive, there are a lot of bad innovation, but when we use the word innovation, we have the bias towards innovation is always good. No, no, regulation doesn't stop innovation. Regulation stops bad innovation. And we need to have that. So we need to build a world where everything's better because we know what are the issues.

And for example, the financial world has a lot of regulations and they still can innovate. Technology can be the same. And this will be good because today I think the approach to innovate is brute force. Let's use more compute, more data, more energy. But this is not the best way to innovate. I love from DeepSeek [inaudible 00:52:01]. Because I said, okay, the Chinese are sinking because they have some restrictions, and they have to improve to match the same quality. So they use less energy, less parameters, less compute, and the results are not far. So this is what we need, we need to go back to thinking, we need to go back quoting Knuth in a different way, we need to go back to understand what is happening in the system to do it better.

Juan Miguel de Joya :

No, I think that's a great point. There's always going to be a question of responsibility.

Ricardo Baeza-Yates:

We need to be responsible. That's why I prefer to use responsible AI and not trustworthy AI or AI safety, that you don't take care of what happens. And of course I don't like ethical AI, because AI is not human, so cannot be ethical.

Juan Miguel de Joya :

Yeah, those are interesting nuances to bring up too.

Ricardo Baeza-Yates:

Yes.

Juan Miguel de Joya :

Yeah, absolutely. One of the questions I think I had when it comes to how you mentored students, and to be honest, when you gave me any advice as well, which I always appreciate, is your ability to elevate people to be inspired. And I was just curious if you have any advice on how to develop community to let students or any young professional believe that they can contribute at the highest international level.

Ricardo Baeza-Yates:

That's a complicated question. I think at the end means that you need to volunteer for a lot of things. And join or volunteer for ACM. I'm a volunteer for ACM, and a lot of people give a lot of time to volunteer either working in committees, either doing this podcast, many different ways, like organizing conferences, mentoring people. I think you need to volunteer to build these communities. And these communities are built by networking, by meeting people, by helping people. I still answer all my email. I know some people don't do that, but many times it's a student that is asking, "Can you share this paper? Can you help me with this?" And so I try to answer all of them, because that's the only way to improve the world, especially if this request comes from developing countries. But sadly, I don't see the same in many people. So most of the researchers are only interested in their careers and building only their success.

I think part of your success depends on what you do to others. At the end, you are not expecting any return, but life will return something to you. I believe that. And in my experience, that happens. 20 years later, because of something you did to someone, you get something back that is even better. So I think you need to be kind, to volunteer for these things and to help people, not only in computer science, in any aspect. So try to help people. Look at the world. I mean, you are not alone, because today I feel that most people think that they live alone and they don't care about other people.

Juan Miguel de Joya :

No, I think sometimes we can get that feeling as well. Personally, I think this is why I think your receipt of the Luiz André Barroso Award is a big deal, it really brings together two important parts of your career, which is a lot of your technical contributions, but also your broader impact in general to community. And so I think one of the things I was thinking about when I just heard about the news, I was wondering how you felt about it when you received it. What were you feeling at that time?

Ricardo Baeza-Yates:

It was a really great feeling. I never met him, but I knew about him. And it's sad that he died young because of an illness. As I said, Berthier, my co-author of the book, was his friend, so I knew about him. And in some sense, being South American, I feel proud of receiving an award that was given because of a South American researcher. So this is amazing. And when I will receive it in the ACM Award ceremony, I guess I will feel more things, because I will be surrounded with many people that I know, that I care, and also people that I admire, like two awards. So it will be a nice moment in my career.

Juan Miguel de Joya :

I think so as well. It's an interesting award to receive, I would say. And I think one of the things I'm curious about for sure is that because it does carry a deeper meaning about visibility and representation, and that's something we've been just talking about, of course, is how do you hold those two meanings together for yourself?

Ricardo Baeza-Yates:

Philosophical question. So I think I feel lucky that in this case I was born in South America, because otherwise I would not be able to win this award. And this has other connotations. So for example, everyone is born in a random place and could have been anywhere. And that's why, because I'm in some sense an immigrant, I have lived in seven different countries during my life. So I feel like I don't belong to any country really. I already have three passports. So I feel that I belong to many communities, one is computer science. But I feel that's part of the problem in the world, that we have these boundaries that are completely made up, that we have ownership issues. We have a lot of problems that we created because all the civilization is basically a fiction. And some people like Harari and other people point that well, we have invented most of the restrictions we have, money, property, basically belonging to a country.

And so all these things were not here before. And some of them are causing problems. I mean, we live in a time of growing inequality, and I think that's not sustainable. Also, I will not talk about climate change and other issues we have. So the question is, it's weird that we are the only animals that have taken that path. We can think something that very few other animals maybe are doing in the way that we are doing, but at the same time, we are stupid because we are creating a world that is not better for us. I will place myself there. I cannot do much, but work in responsible AI is my small contribution to improve the world.

Juan Miguel de Joya :

You may seem to think it's small, but I think everything that we do has meaning and value. And the ability to put that value to something that is meaningful for others doubles that quite a bit. We're coming towards the end, by the way. I don't know, unless you want, we can still keep talking after the podcast, of course, but I really wanted to just reflect on your career and your experience as a professional. You've been doing this for quite a while. You have seen a lot of different things, not just from the technology side, as we mentioned, but also the impact side. Do you have any thoughts on how your mentorship has shaped the way you think about impact?

Ricardo Baeza-Yates:

So, as you said, I mean, it's not only about computer science. For example, I love geography, so I travel a lot and I have seen many realities. I think understanding the world also changes you. So seeing the contrast of India, seeing how people live in Africa, seeing very happy people that don't have anything. I

mean, these things shows that maybe the culture we are creating probably is not the best for us, because we were not made that way at the beginning. So evolution is important. So I feel that maybe some tribe in Amazonas is much more happy than us, probably that's the case. And they don't have many of the problems that we have because they don't have created those problems, because at the end, those problems were created by us.

I will go back to what I said before. I think how we can shape the world in just ways that improve things in small ways, maybe at some point, maybe new generations will say, "Stop this." And you can see it in some movements, like changes on how you eat, on what things you do, on how much exercise you do. Maybe if more people will do this, and at some point the political elites go away, we will have better people governing and we will have a better world. But maybe I will not see that, but I think you and I can contribute a little bit to improve that. But I guess you have to be a bit idealistic in the spite that it's not being very realistic, because it's a very hard work to be against the dominant system. And it's not about politics. I think many people mix politics with having a better world. I think if all people wants to have a better world, maybe there are different ways to do it, but we should forget about politics when you want a better world.

Juan Miguel de Joya :

No, I think that's an interesting point. I guess from a technological side, what question about, maybe let's say search AI or even any technology do you hope that the next generation will solve or try to solve?

Ricardo Baeza-Yates:

Well, I think the main question today is where we should use AI or not. This is a question that we should carefully think, because if we lose cognitive skills, we are lost in some sense in history. We need to decide basic things about where and when to use AI. For example, when a kid should start using AI. For what? Until when? These are basic questions that don't have answers because we are going too fast, but if we don't answer these things soon, the inequalities will appear in different ways. Maybe in some sense in good ways, because the more developed countries are using more AI than the developing countries, and this could be in their disadvantage, not in their advantage.

Juan Miguel de Joya :

Yeah. No, that's a very good point. Ricardo, thanks so much for the conversation, by the way. First off, it's really nice to just see you and talk to you again. It's been a while.

Ricardo Baeza-Yates:

Yes.

Juan Miguel de Joya :

I do really appreciate it, because I think what I sometimes forget is that there is a way that we should see technology, and you're helping us see how search is not just a technology, AI is not just a technology, but it's also one of the ways that people can make sense of the world. And how can we improve the way we make sense of that world? We've talked about some of the foundation of information retrieval, we talked about how search has evolved, and the responsibilities that come with building these systems as well. And finally, I think we got a chance to talk a bit about some of your life, some of the stuff about mentorship, the different things that we're seeing and the opportunities there should be to elevating the Latin American community as well. So yeah, your work reminds us that

computing is not only about the systems we build, but also about the people and institutions we help bring forward. So yeah, congratulations again on the receipt of the award. Thank you for joining the ByteCast.

Ricardo Baeza-Yates:

Thank you for the opportunity of having this ByteCast. And maybe to remember, as you're saying, AI is a tool, and may help to solve our problems, but the only people that can solve our problems are ourselves. These are social problems. These are not technological problems. And technology can help, but also can create more problems. And we need to try to avoid that.

Juan Miguel de Joya :

Absolutely.

Ricardo Baeza-Yates:

Thank you.

Juan Miguel de Joya :

Thank you. ACM ByteCast is a production of the Association for Computing Machinery's Practitioner 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.