Angela Bonifati Speaks Out on Research, Grants, and Collaborations

H. V. Jagadish and Vanessa Braganholo

Angela Bonifati
https://perso.liris.cnrs.fr/angela.bonifati/

Welcome to ACM SIGMOD Record's series of interviews with distinguished members of the database community. It is my pleasure to host Angela Bonifati today. She is a Distinguished Professor of Computer Science at Lyon 1 University, Head of the Database Research Group at the CNRS Liris Research Lab, and a senior member of the Institut Universitaire de France. I'm H. V. Jagadish, and I'm glad that Professor Bonifati is with us today. So, let's dive in. Welcome, Angela!

You have many affiliations, one of which is the French University Institute, which may not be well-known to people outside France. Can you tell us a little about what that is?

Thank you, Jag, for the invitation. I'm really honored to be part of this series and to answer your questions today.

The French University Institute is a service of the French Ministry of Education that recognizes university professors across all domains in France. Only very few individuals are recognized. Around 2% of French university faculty are members of this institute. I'm a senior member, which means the share of members who are older than 40. What I receive as benefits from this institute is a waiver of part of my teaching service for the next 5 years, and also a national award of scientific excellence. There is also a small research grant from which I'm now hiring a postdoc.

Thank you. I was thinking about your research accomplishments (which you have many and we won't have time to talk about all of them today). In my mind, I put them in three main buckets: XML, schema matching, and graph databases. Of course, what you've done is not limited to these three, but this is just to create some structure. Could you tell us a little bit about how your research interests have evolved across these and other areas?

Yes, of course. I'm really happy to talk about this. I started to do research on data management during the XML hype, and this was during my PhD. At the end of my PhD, I began working on schema matching and mapping, and started to become interested in data integration problems. Later on, I was interested to explore the interplay between those areas and I also considered combinations of schema mapping with heterogeneous XML data. Starting from 2010, there was the explosion of Graph Data Management Systems, and I also started to work on these new NoSQL data systems. That was the start of companies such as Neo4j. I then continued working on data integration, schema mapping, and matching.

XML was over, if you remember, around those years, so I kept working on these two areas of graph databases and data integration. Recently, I worked on transforming graph data, where the latter can be obtained from any kind of format (even relational data and time series), to actually integrate this into a property graph. As you can see, the three areas are not siloed, but I always kept studying the interplay between them at an abstract level, with concrete contributions.

I guess in particular, graph databases and more broadly, graph processing and curation, is what you're really focused on lately, as far as I can tell. Can you tell us a little bit about your perspective about these areas? What's interesting and cool about them, what makes you passionate about these?

They are a very interesting new technology for data management, in my opinion. There is a wide landscape of graph engines and graph data systems that have appeared since the inception of graph databases in 2010. What happened is that there were so many tools and so many engines, and also so many query languages behind these tools. And these languages were not unified. That's why, about two years ago, there was this effort, driven by academia and industry, to standardize these languages. In particular, in 2023 and 2024, the first standardized Graph Query languages, namely SQL/PGQ and GQL, appeared as a first version of the standards for query languages for property graphs. And these are the first standards after SQL, after the standards that we know for relational databases.

That's why I see this area as really flourishing now. And I was lucky, because during the pandemic, I was part of the LDBC working group, which was working in liaison with the standardization body. I was working on property graph schemas and constraint languages, and these were published[1] in the Industrial Track of SIGMOD in collaboration with people from the industry. We have to think about the fact that some of these standards will now be implemented in the tools. So all these companies that are part of this effort will take their own language and implement some of the features of the standard. So, that's why I'm really excited about this area. And then, of course, I'm a professor, so I had grants on these topics.

And as you know, there are a lot of things that one can do on top of these standardized languages. I was working on graph transformations using standardized languages, and now, lately, I'm also exploring the boundaries between graph databases and artificial intelligence, especially for curation processes.

In terms of grants and grant-making, you've recently obtained an ERC Advanced Grant, which is a high recognition for leading researchers in Europe. I believe it's the largest kind of grant that researchers in Europe strive for. It's a very competitive thing, so

[1] Renzo Angles et al: PG-Schema: Schemas for Property Graphs. Proc. ACM Manag. Data 1(2): 198:1-198:25 (2023).

congratulations! Can you tell us a little more about your vision for this grant and the research that you are about to start, and what you think this will lead to?

Thank you! Yes, I'm really excited about this grant. As you said, this is the most prestigious individual grant that European researchers can get in their careers, and it's very competitive. So, I was extremely happy to receive this grant that will start on December 1st this year. So the name of this grant is ERC GO-Y (pronounced "Go Why"), and guess what? This grant is on unifying graph databases and causal models used in AI.

It starts from the observation that in our data systems (and this observation can also be made for relational databases), we have been focusing mainly on processing and pre-processing. But here in this grant, I envision a radical shift: having data systems that are actually guided by causality, by being able not only to return results, or to be able to do pre-processing, processing, or transformation of the data, but also to return the explanation behind the processes. And what I see behind the explanation is to be able to return the causes that are behind the effects.

In this 5-year project, I intend to address the development of theoretical foundations and also tools to be able to understand the causes behind the data and to be able to execute and evaluate these causal graph operations in our data systems. I work on graph databases, so when I wrote the grant, I realized that the theory of causation uses directed acyclic graphs to encode cause-and-effect relationships, and this can actually be encoded within a graph database. And this opens up a lot of research topics and data management tasks that, right now, are not there because people do these things with *ad-hoc* scripts. Additionally, they do it for one dataset or one causal question at a time, whereas in data systems, we can work with any dataset, and we can work with any query. Therefore, I think there is a lot of potential there to unify these areas that have been kept separate so far. The GO-Y grant will eventually explore the interplay between data management and causal AI, leading to implement causal processes in a database.

So, you mentioned that this was an individual grant, so it's all you? Are there other collaborators involved?

No, it's just for me and my team in Lyon. It's actually a huge grant, it's almost 2.5 million euros, only for my team. There is no consortium behind this grant.

No wonder it's so prestigious and so difficult!

Turning to some non-technical questions, when you have a big grant like this, the odds of getting it are very low, even for very good people, doing very good work. It's highly competitive, as we were talking about. So, what's your philosophy behind this, and what would your advice be to others who might be thinking about this? Should researchers go after these big prizes with low odds, or is it better to put energy into things that are more likely to succeed, even if they're not as big? How do you balance these?

It's a difficult question. I don't know whether I can give credible advice there, because it depends on the individual experience. You can apply to an ERC depending on the stage of your career. They have the Starting Grants, then they have the Consolidator Grants, which are targeted for mid-career researchers, and then they have the Advanced Grants, which are for senior people like me. I never applied for Starting or Consolidator grants. I was working with smaller grants before, also with the European grants, in a consortium with other people.

I think one needs a bit of training with other grants that are probably much easier to get before going to ERC. In my case, it was useful to have been funded by other agencies before ERC. But, I would certainly recommend it to people who are at a certain stage of their career, and even junior researchers, if they feel they have reached the peak of their career, that they have a lot of contributions, and they are doing a lot of research. Even if the chances are low, nothing is lost, because we know that writing a grant brings us to do some brainstorming, think about our future research topics, and the grant can always be funded by other agencies. So I would definitely recommend researchers who feel they are ready to go for an ERC, to go for an ERC.

I have a bare minimum requirement [for collaborating with Young PhD students and post-docs], that is I want to work with people who can meet the deadlines, and who are very serious about their work.

Can you tell us about your own professional arc? You started in Calabria, and are currently in Lyon, and have

been in multiple places along the way: first in Italy, then in France. How did you decide to stay in some place, or to look for better opportunities?

In my case, I was trying to choose a place based on the actual possibility of doing research in that place. I always wanted to do research. You know that, especially in France, but also in Italy, one can be consumed by other activities, right? It could be an admin or a huge teaching load. I tried to choose the place based on the possibility of doing research, and I always look for endowed research chairs as part of the job positions. These are very common in France, similar to the French University Institute. I had a couple of years in which each year I was basically waiving my teaching service. And this explains why I moved around some places, because whenever I felt that I did not have enough time to do research, then I told myself, "Maybe I should look for a new place."

In terms of countries, I've been lucky because I moved basically from Calabria, in southern Italy, to France, well, to northern France at the very beginning, and then to Lyon, which is in southern France. But the two countries are very similar, too. I also had a few extensive visits in North America and Canada.

I'm quite happy about what I did, because I've always been driven by the intention of doing more research. I think the various parts are part of the whole, right?

What I look for in my collaborations is to learn from others.

Even though you've been at multiple places, they've all, as far as I know, been academic, or at least quasi-academic, like INRIA, and so on. How do you feel about working in industry or industrial labs?

That is one possibility that I haven't explored. During my PhD, I did an internship in industry. I was at HP Labs in California, and I had a great experience there, both personally and professionally. I was working in Umeshwar Dayal's team for a 3-month internship. We published a paper[2] in the industrial track at VLDB 2001. But as you remember, the industry was hitting a crisis after the year 2000, laying off many people. And even though I liked this industrial experience, I thought I should stay in academia, and I looked for a postdoc. And then, starting from my postdoc, I went directly to academia, and I basically didn't regret it.

So I never felt the need to go into the industry, but of course, I don't exclude it for the future. I've always been collaborating with companies, so everything is open. I don't exclude it.

You've had many collaborators: juniors and seniors, from various places and various levels. But one thing that kind of struck me was that even when you were a junior researcher, you had already developed collaborations with many senior researchers. Could you say a little bit about your overall philosophy for collaborations, and then perhaps also specifically about how it came to be, and how you've had so many senior collaborators early in life? What are your views on that?

When I was at these visiting periods in North America, all the collaborations happened there. I also had collaborations with people in Europe. You are right, maybe my collaborations were always with people who are more senior than I am. What I look for in my collaboration is to learn from others. So, it was not intended, but you know, we learn from each other, and especially from people who are more senior than us; we learn a lot, because they have a lot of experience. This is something that we already experimented with during our PhD, because our PhD advisor has a lot of experience. It worked well. I learned a lot from all my collaborators.

I started these collaborations by chance, let's say. Sometimes I was visiting a place, and I was expecting to collaborate with someone, but this person didn't have time. So I just started talking to another colleague, and then we started to collaborate. You see, I would not say I always choose people who are more senior than I am, but it just happened, and in the end, the result is quite satisfactory. And what are the qualities that I value in my collaborators? If I also consider the collaborators in my team, like the young PhD students or postdocs, I have a bare minimum requirement, that is I want to work with people who can meet the deadlines, and who are very serious about their work. That's all.

That's very succinctly put. So, asking you a geopolitical question: as a senior researcher in Europe, how do you view research in Europe versus the rest of the world?

[2] Angela Bonifati, Fabio Casati, Umeshwar Dayal, Ming-Chien Shan. Warehousing Workflow Data: Challenges and Opportunities. VLDB 2001: 649-652.

Are there differences, both in terms of subject matter and in terms of style, that you'd like to comment on?

That's a difficult question. I'm a European faculty member, and as a researcher and educator, we are all confronted with generative AI and with AI in general. We can see that this AI landscape is rapidly evolving, and everyone is trying to understand how to work with it, so that we can only take the benefits of this new technology, and we can avoid the shortcomings.

Europe has been working on some of the regulations for AI. As you know, the EU AI Act was designed by the European Commission. But what I notice in Europe is that it is difficult to find agreement across European countries, and unless it comes top-down from the European Commission, then it is difficult to get all of the European countries aligned, whereas it would be good to have that. For instance, even in France, different universities or research institutions are deciding their own AI policies. Whereas we could decide together what the AI policy should be, right? This would actually let us control this technology.

It is very difficult to say what is the positioning that we should have in research. I think more investment is needed. In France, for instance, the government has been funding these Research Institutes on AI. But there were a few of them, and they were given to some cities at the detriment of the others. I think more investment is needed in AI if we want to be competitive in Europe with respect to other continents.

I think we need more discussion, too. As computer scientists, we need to discuss at all levels: as educators in our universities, as people who are in charge of deciding the future of conferences, etc. All the actors should be involved in discussions. I think this is bottom-up.

I also wanted to mention something else. I was reading an article about the latest United Nations Summit, where Yoshua Bengio mentioned the fact that now Gen AI is becoming even more critical in security and peace. And now in Europe we are there, right? I mean, we have two wars on our borders. So, I think that, as Europeans, we should really take it seriously. Having our own AI technology is important. With that, we will not only be competitive in the IT sector but will also take care of these critical issues, such as security and peace.

Talking about the research atmosphere more broadly. As a senior woman researcher in a community with very few women, are there comments that you'd like to make to the community as a whole? Both with respect to the database research community overall, and also specifically with respect to your position in France.

I think mentoring is very important. I'm a senior woman researcher in a community with a few women. What I do is I try to mentor my (male or female) collaborators, so that they learn how to navigate through their career milestones. I wrote an article[3] on advice to mid-career researchers in SIGMOD Record that appeared in July 2024. The gist of it is that, when somebody is in their mid-career, they have passed this stage of independent critical thinking. I was mentioning my own experience. I've always been driven by curiosity, curiosity about new research subjects, and then writing, writing down your ideas. These were the suggestions. I don't think this is a female-to-female suggestion. I mean, it could be from female to male, for young people or mid-career researchers. And I think mentoring. Mentoring both ways. I mentor other people, and I'm mentored by people who are senior than me. I think that's the key to having more people join this community.

[As Europeans,] Having our own AI technology is important. With that, we will not only be competitive in the IT sector but will also take care of these critical issues, such as security and peace

I think that's a very positive statement, and that leads me to ask you about your view of the community and your vision for it as the new chair of SIGMOD. What are your plans? What is your vision? What are the things that you'd like to make better for all of us?

Thank you for the question. I'm really honored to have been elected as the SIGMOD Chair. I started my term on July 1st, and in the last few weeks, we have already been discussing many issues. I see SIGMOD as part of the SIG ecosystem of ACM, apart from being an executive committee of the SIGMOD conference, and also co-sponsoring other conferences. The fact that it has these two aspects makes me busy on both sides.

The biggest challenge now on the ACM SIG landscape is the transition to open access. As you know, in January 2026, ACM will start adopting open access. This means

[3] Angela Bonifati: The Long and Winding Road to Mid-Career Academia. SIGMOD Rec. 53(2): 48-51 (2024).

that people who don't belong to institutions that have signed the open access agreement with ACM will have to pay the APC fees. This is a big transition for SIGMOD, and not only for SIGMOD, but for all the other conferences that are sponsored by SIGMOD and all the other ACM SIGs. So I'm really busy with this.

There are also other long-term challenges that are awaiting us, such as AI in data management, ethics in the review processes, and changes in the political landscape that affect our community. All these challenges are equally important, both at the ACM level and at the conference level. So, I intend to work on both aspects: SIGMOD as part of this ACM ecosystem, and SIGMOD as an executive committee of the ACM SIGMOD conference, and all the other SIGMOD-sponsored conferences.

Thank you. Best of luck in terms of what you're planning to do. I hope that's all very successful, and I look forward to it.

Thanks.

I wanted to talk about one of the things that our community has started to address in a very small way, which is helping people balance their professional work with family. A few conferences now have some help with childcare, for example. You have brought your daughter to some events, and you're managing all of the amazing things that you've been doing recently, while also having a young person that you're responsible for at home. So, can you talk about both how you manage your balance in terms of time, as well as are there things that we as a community could do to make this easier on others?

That's a difficult question. I think that the database community already has many initiatives to help young parents. There has been daycare, as you mentioned, in a few of the last editions of conferences. And this is very helpful for parents who want to attend the conference and bring their children. This is very important. This was not the case when my daughter was young. I remember the first years, I could not travel to conferences because I could just not bring her. The first time I brought her to an event, it was in a Dagstuhl Seminar. By chance, by going and checking the Dagstuhl website, I discovered that they offer a babysitter. And then I could take my daughter, because the babysitter would take care of my kid during the entire seminar. She was 5 years old at the time, and she had a great time. When we came back, she said, “Why don't we go back again to that castle?” And the reason why I brought her, of course, was to spend more time with her, because my partner could not keep her during that week, and also because I wanted her to learn about my job. And I think this actually happened.

Since then, whenever I can, if I have an event during summer holidays when she can come, I just bring her. And I think this is great, because she learns about my job, she knows why I'm traveling, why she cannot come, because she's at school, and I'm not at home. So, I would suggest that parents do this. And of course, the community can help there, as they're already doing, and maybe more, with solutions. Solutions for keeping children and having daycare during conferences.

So, going beyond work-life balance, I guess there's just a question of time management and juggling many different things. So, aside from having a school-aged daughter, right now, you're on the VLDB Endowment Board, you're a general chair of VLDB next year, you have a large ERC project that you're just starting, and you have been chair of SIGMOD now for the past couple of months. So, how do you do all of this?

I don't know. I'm a very organized person. I didn't ask myself this question before, but I think I'm very good at time management, and I think I even learned to be better at time management after becoming a parent, because I knew that in the evening, for instance, I could not work because I have to devote time to my kid. So, I just work on various tasks during the day. If I have to say, I do it in a round-robin fashion, in the sense that I dedicate fixed slots during the day to each of these activities. I mean, it worked in the past, so I'm very confident that it will work in the future. Of course, I have these, as you said, big challenges with VLDB 2026, and then ERC, and then SIGMOD chair. I haven't done this before, but I'm still optimistic. Let's see...

By the way, I also enjoy extra work activities, such as spending time with my family, going for long walks, and cooking, especially Italian cooking.

So, I guess we need to have another edition of Speaks Out, where you will share your recipes with us.

Of course! With pleasure, with pleasure.

But for today, I just want to thank you.

Thank you, Jag.