

Visions, needs and requirements for (future) research environments: An exploration with ERC Grantee and UiB researcher Saket Saurabh

Katharina Flicker (TU Wien), Saket Saurabh (University of Bergen), Bernd Saurugger (TU Wien)

Researchers are at the very heart of the EOSC: So what do researchers really need to do cutting-edge research? How do they think the EOSC could support them in their endeavours? Let's see what theoretical computer scientist Saket Saurabh has to say.

“By joining efforts we might achieve more than as individuals”

TU Wien: What is your work currently focused on?

SS: We focus on hard optimization problems – in a technical language, these are called np-hard problems. In practice, such problems are very hard to solve, and hence these problems can be solved exactly for only small problem sizes. That is why we draw on e.g. finding approximate solutions or some subclasses of inputs to solve practical problems.

TU Wien: What are the things that drive your work? And why?

SS: Let me give you an example: Imagine that you would like to travel to several cities and you would like to travel in an efficient way. Thus, you are given the distances **between** each pair of cities in order to find the shortest possible route that visits each city and returns to the origin city eventually. This is called the *Travelling Salesman* problem. It is an np-hard problem and it is one of the most well studied problems in optimization. And these kind of np-hard problems appear everywhere. So you would like to design an algorithm for that. We are both driven by curiosity to solve these hard problems in different algorithmic paradigms as

well as its need in various day to day applications.

TU Wien: What datasets are you working with? And how do you analyse the datasets?

“It would be great if datasets were freely accessible”

SS: I am a theoretician. Thus, we do not need many datasets because we do not do any experiments at all. However, lately we started working with problems related to big data. So you have a big amount of data that is trying to communicate something to you. Now what I'd like to do is to find a small data sets which essentially have the same kind of features as the big data sets. So in essence, the small data, which you are storing, is representative of the original data. I need both data sets to check the algorithms we need. Theoretically, I can say that “look! If you give me data that satisfy this property, then the small set, which I am going to compute algorithmically, represents the data very well. Theoretically, that is ok, because it

would prove theorems. But I would like the data to really run these algorithms and see that it really does what I intend to do. Not only in theory, but also in practice.

Unfortunately, it is very hard to get these datasets. For this, you would need to collaborate with very different kinds of people to get these datasets. Of course, we can generate artificial datasets and be happy that our algorithm works on these artificial datasets. However, it is not very satisfactory. So in my opinion, it would be great if datasets were freely accessible.

TU Wien: Ok. What would it take to encourage you, or the researchers (coming from your field) to make your data open, or at least accessible?

SS: I thought about that quite a bit. I would say that competition is the biggest hindrance. So for example, putting together a good dataset takes a lot of effort and time. Thus, it is very natural that one wants to exploit every little aspect of it, before you share it so that no one else can use it before you do. So I think the way we think about research has to change. We have to start opening data for everyone. The more data that gets available for me to use, the more I will be able to contribute. But someone has to start at some point in time and say “I do not care. I open my data and I open my research”. Once one or two big fish in the field do it, others will be encouraged to do the same.

“I think the way we think about research has to change. We have to start opening data for everyone. The more data that gets available for me to use, the more I will be able to contribute”

TU Wien: So what are your suggestions exactly?

SS: I think you will need the most established people. They will have to take a lead on this, because they are the trendsetters. As a young researcher, I was always a little bit worried about how things would go. However, established people will be less insecure. Thus, they need to take the lead, as others will definitely follow. The ERC Advanced grantees are a very good example. Generally, they are people who are on top in their fields. If they share and open data, it is very natural for their working groups to do the same. Thus, a culture of data sharing and opening data is being adapted.

“If they share and open data, it is very natural for their working groups to do the same. Thus, a culture of data sharing and opening data is being adapted”

TU Wien: In general, how will research look like in five to 25 years, or 15 years? And what will be the effect and impact on research environments like infrastructures, services or policies?

SS: So if there is one thing I learned due to the Covid-19 pandemic, it is that in theoretical research distances will matter less. Before that, I didn't think that online-seminars, teaching, meetings and conferences could actually work. Given the situation, we all had to try and do our best and I was satisfied with what we were able to get. That would also help reducing costs. Take the culture of having to attend conferences as an example. You go there to disseminate your ideas after your paper was accepted. To go there physical however, is expensive: you will pay registration fees and you will be paying for transport and accommodation in order to have a very small audience listening to you. The

audience becomes bigger if you do it online. In addition, you save money that can be spent on e.g. administrative staff / scientific administrators. I guess this will also effect policies and funding. And I really believe that as you go ahead, distance will have less meaning. Of course physical meetings will definitely be helpful and everything, but I think we have just found alternatives.

TU Wien: Having said all this, what would you need the EOOSC to be to support you and your research?

SS: Help getting data, find people to collaborate with and identify people with similar goals. Then we could come together and come up with better output and better dataset. For example, ERC is funding so many projects. When you take a closer look, you see that some are similar in nature. Therefore, if they could talk to each other, not only would they succeed in doing what they are doing, but in a joint effort, we might achieve more than as individuals or in small teams.



Saket Saurabh received his PhD in Computer Science (2008) from the Institute of Mathematical Sciences (IMSc), Chennai. Saurabh spent two years (2007-2009) as a Postdoctoral Fellow at the University of Bergen, Norway, and is now a professor at IMSc and at the Department of Informatics at the University of Bergen. His main research interests are in graph algorithms, parameterized algorithms and complexity. He is a Swarnajayanti Fellow in Mathematical Sciences, Fellow of Indian Academy of Sciences and Academia Europaea. He got the ACM India ECR 2020 Award. He is also recipient of an ERC starting Grant and an ERC Consolidator Grant in parameterized algorithms. He is a co-author of two books: Parameterized Algorithms and Kernelization theory of parameterized preprocessing.