AI 4 Science Discovery Network+

AI4SD Interview with Professor Zoheir Sabeur
30/11/2021
Online Interview

Michelle Pauli
Michelle Pauli Ltd

12/08/2022

AI4SD Interview with Professor Zoheir Sabeur
Humans-of-AI4SD:Interview-38
12/08/2022
DOI: 10.5258/SOTON/AI3SD0234
Published by University of Southampton

Network: Artificial Intelligence and Augmented Intelligence for Automated Investigations for Scientific Discovery
This Network+ is EPSRC Funded under Grant No: EP/S000356/1

Principal Investigator: *Professor Jeremy Frey*
Co-Investigator: *Professor Mahesan Niranjan*
Network+ Coordinator: *Dr Samantha Kanza*

Contents

1 Interview Details

Title	AI4SD Interview with Professor Zoheir Sabeur
Interviewer	MP: Michelle Pauli - MichellePauli Ltd
Interviewee	ZS: Professor Zoheir Sabeur - Bournemouth University,
Interview Location	Online Interview
Dates	30/11/2021

2 Biography

Figure 1: Professor Zoheir Sabeur

Zoheir Sabeur: 'I was among the earliest generation of data scientists'

Zoheir Sabeur is Professor of Data Science and Artificial Intelligence at Bournemouth University, Department of Computing and Informatics, Faculty of Science and Technology (2019-present). He is also Visiting Professor of Data Science at Colorado School of Mines, USA (2017-present). Zoheir was Science Director at the School of Electronics and Computer Science, University of Southampton (2009-2019). He led his data science research teams in more than 30 large projects as Principal Investigator with research funds exceeding £7.5M. Prior to Southampton, Zoheir worked as Director and Head of Research, at BMT Group Limited (1996-2009), where he led his teams in the development of advanced environmental information systems for the UK Government and the Energy Industries. Prior to that, Zoheir held several academic appointments in Computing as Senior Research Fellow at Oxford Brookes University (1993-1996), SERC Research Fellow at University of Leeds (1991-1993) and University of Strathclyde (1990-1991), and Research Scientist at Wuppertal University (1987). Zoheir graduated with a PhD and MSc in Theoretical Physics from the University of Glasgow (1985-1990).

In this Humans of AI4SD interview he discusses the early days of data science and its importance to this century, how to understand, detect and track human behaviour to make urban spaces safer, the fascination of mathematical theories, and how data science came into its own during the pandemic.

3 Interview

MP: What's been your path to where you are today?

ZS: It all started with a wonderful physics teacher I had during secondary school. I was around 15 or 16 years old, and she was the first teacher who I felt connected a particular discipline to real-life problem solving. She brought up rich examples of how we could use what we were learning in physics, and how it related to the world around us. It fascinated me. I knew from then that I wanted to be a physicist, but in relation to something connected to real-world problem solving. I went on to get my baccalaureate in mathematics, before taking physics at university, specialising in theoretical physics. My interest was in how this world works and how the universe started and expanded.

These interests continued in my Master's and PhD degrees in theoretical physics at Glasgow University. My PhD was related to lattice quantum chromodynamics and at the time, I didn't even realise I wasn't strictly a theoretical physicist but in an early generation of data scientists using intensive computing. In lattice quantum chromodynamics, we represent field theory using a lattice hypercube, representing all the strong field and gauge-field interactions numerically. In the mid-80s, when we were doing it, it was quite a new branch of thinking among us theoretical physicists. Most of the theoretical physicists, in our group at Glasgow University were also numericists, applying all sorts of numerical techniques to converge our algorithms.

This approach was unusual at the time. My supervisor, the late Dr Ian Barbour, was one of the first pioneers in the UK, who brought data science into theoretical physics. He was among the early physicists who explored hadron matter phase transitions, using intensive computer processing in order to understand the behaviour of the represented lattice QCD system. I was indeed very privileged to have him as my PhD supervisor. Together we looked at the mechanisms of how free hadron matter formed earlier in the universe, and became the matter we know now with its confined quarks and gluons. We explored whether the thermodynamic transitions encountered by hadron matter were due to a fundamental symmetry which we knew about that broke down in order to lead the existence of the current state of hardon matter. These were questions which could not be answered theoretically, but only numerically. However, it required running the lattice systems on the most powerful supercomputers of the time. Most of my PhD was spent simulating lattice systems, under various scenarios or what called them at the time "under various jobs". With those long running simulations, I was able to analyse my data and investigate on the relationships between chiral symmetry breaking, or restoration, are connected to quarks-gluons confinement, or deconfinement. In fact, I didn't realise at the time that I was among the earliest generation of data scientists.

MP: What did you go on to do after your PhD?

ZS: After my PhD, I carried on using computer power for research, because that was what drove me most to investigate on the behaviour of natural phenomena and processes. There was a turning point around 1987, when I went to a Scottish universities summer school at St Andrew's University entitled "Computational Physics". There I met not only with physicists, but mathematicians and people doing meteorological sciences. During the three weeks running of this schools, we talked about artificial neural networks and cellular automata and fractals too!! All of this was a mind opener to me and for the possibilities of understanding natural phenomena using computer power.

I went on to work at Strathclyde University, doing computational molecular physics. I was trying to apply finite element techniques in order to solve Schrödinger's equations of tri-atomic systems computationally. Then I moved to Leeds, where I studied the phase transition of molecular fluid at high-density walls and spheres. There, I published some good papers, which keep being referenced to these days. Then I moved to Oxford Brookes University, where I was interested in the impact of transient waves on coastal defences. That was quite a leap from fundamental physics to environmental sciences, which also interested me. There, I developed computational methods to solve the Navier-Stokes equations with free surface boundaries, which are quite tough to simulate and keep them stable. I did, however, manage to develop a numerical solver, which remained stable before, during, and after wave impacts on a solid wall. With this approach, I came to the realisation that if the defences are permeable rather than solid, you can actually control the way forces propagate in the structure. From here, you can attenuate the waves as you want, while it could be useful for green energy harvesting in the future.

In the mid-90s, I moved from academia to industry at a time when people were starting to think about web-enabled systems. In industry, I developed advanced information systems for predicting the short-term and long-term fate of offshore oil and gas operational wastes, including their impact on marine ecosystems. I spent 13 years in industry, while collaborating with various academic institutions around the UK and the world in order to perform my laboratory experiments with them and validate my new modelling systems. I felt working on real-world problems "big time" while I delivered important, and independent peer scrutinised, scientific solutions to the energy sector industries for conducting their offshore operation safely in the marine environment. Further, this was in full agreement with the governmental regulators involved for licencing their activities too.

In 2009, I decided to move back to academia and set up my research group at IT innovation Centre, Department of Electronics and Computer Sciences, University of Southampton.

MP: What prompted your move back into academia?

ZS: In industry, I felt like I had done all I could do. I had developed these systems for the industry, as well as the UK government department of trade and industry, who funded these programmes together with the UK Oil and Gas Alliance. The systems were put in place to routinely regulate offshore industries operations. There was then a set of complete information tools for minimising the environmental risks of operational waste discharges such as drill cuttings and production waters on the marine environment.

In 2000, I met with the Head of the Department of Electronics and Computer science, who introduced me to the concept of data fusion. While I was still in industry, we developed ideas as of how to use data fusion to mix data and predict particular processes where we do not have extensive spatial sensing. Data fusion uses heterogeneous sources of observation data from existing multi-modal sensing platforms and combines them to create a fused set of information of interest. Doing this work introduced me to various groups at the University of Southampton, and then a possibility for an appointment came up at one of their research centres. This was the IT Innovation Centre at the Department of Electronics and Computer Science.

My appointment at Southampton, enabled me expand my ongoing research work in data sciences well beyond the marine sector. While my research work in industry was focused on its application on marine environment problems, I managed to expand it to be applied to security and health problem solving. In the big data science group I set up, I worked with my researchers on more than 30 funded research projects, where we investigated on data science and artificial intelligence issues and validated them in problem solving areas related to environment, health and security challenges. This work continues to this day at Bournemouth University where I am based currently.

My speciality really is on advancing AI fundamentals to detect and understand behaviour and processes in nature and man-made systems. The use of intensive computing, combined with data-driven AI methods is really the case here, as it is part of fourth paradigm for scientific discovery. Science has really undergone various phases from the empirical, theoretical, computational, then data-intensive paradigms for discovery. In a way, I encountered the move from the third to the fourth paradigm during my career from research student to academic!

MP: Have these transitions been smooth, or were there particular points which accelerated them?

ZS: It hasn't been steady. In the eighties, we used to meet with various people who could appreciate what we were doing in theoretical physics computationally; these were people in aerospace, engineering, or meteorological sciences. There were few of them in most of the scientific domains with whom we could share our computational techniques and findings. It wasn't until a few decades later that we saw other scientific communities getting involved.

It went gradually first, but it jumped to high levels around 2004-2005. The reason for this was the sheer explosion of big data generation, smart sensing devices started becoming affordable to everybody, and cloud technology also became available in 2008. This created a new generation of data scientists whose work can be useful to any type of discipline from medicine, social sciences and more. Data science is important to this 21st century – no question about it. Especially, when artificial intelligence is proving to be the most powerful technology available to us, with its early convincing capabilities for exploring, detecting and understanding new and complex phenomena rapidly. The progress for scientific discovery will be unprecedented and it's an exciting era to be part of.

MP: Tell me about your current research into urban spaces

ZS: Nowadays, more and more urban spaces are equipped with monitoring devices, like cameras and sensors. Because there have been many attacks in urban spaces, it is important to monitor them for the safety and security of citizens. In the research I'm conducting with my group at Bournemouth University, we are combining in-situ physical observations with cyber observations. These are observed separately, but the aim is to fuse them such observations and make sense out of them when combined through encoded machine reasoning. We try to understand human behaviour in a smart space using vision and other means of observations. With that it becomes possible to understand, detect and track human behaviour in a smart environment.

But beyond that, you can go further in detecting intentional behaviour, which is important if you want to manage the safety and security of crowds in space. It's important to understand how behaviour propagates in space too. With our methods, we can actually anticipate the potential propagation of behaviour. It could start from what we call a seed in a crowd of people, and their behaviour can propagate across the crowd, which is an interesting phenomenon to observe. That's important for people in the security sector who want to be ahead of the game for maintaining safety and security in critical spaces such as motorways, stadiums, airports and so forth.

MP: What are the ethical implications in this research?

ZS: In our universities, all our research applications go through ethics committees, and do comply with their expert recommendations and regulations. We use our research data while we have the consent of volunteers who want to participate in our experiments. The data is also destroyed after our research work is completed.

MP: What has surprised you in your research?

ZS: The beauty of doing research across a career is that, with time and experience, you become more and more resilient. You don't give up on investigations when something surprising happens, you can build the capabilities step-by-step to find answers.

I've also found lots of surprises in mathematics, and they're there for humanity to discover. In data science, we regularly need to go back to mathematical theories to see if they can explain a particular phenomenon, which they often can. As human beings, we like to formalise what we understand through fundamental mathematics. Wearing my theoretical physicist hat, knowledge extraction fundamentally needs to be assisted by strong mathematical theories in the end. The beauty of mathematics doesn't just surprise me, it fascinates me as I also see it in full part of science discovery!

MP: How has the Covid-19 pandemic affected data science?

ZS: The work of data scientists during the pandemic really attests to how accelerated the digital revolution is. I'm grateful the technology was there before the pandemic started, and now we're seeing just how powerful it is. The monitoring of the pandemic has been amazing. It relied heavily on data analysis, which helped us make critical decisions. We've been doing this in record time, and this is because people could rally together and join forces remotely: exchanging data, experimental findings, working together, and producing a vaccine. All thanks to the digital revolution we're in.