BANCA D'ITALIA
EUROSISTEMA

Data science in central banking

Part 2: Applications and Tools, 14-17 February 2022

Closing Remarks by Piero Cipollone
Deputy Governor of the Bank of Italy

Rome, 17 February 2022

Ladies and gentlemen,

This workshop on 'Data Science in central banking: Applications and tools' is now drawing to a close. Let me begin by thanking all the speakers, discussants and participants who have contributed to the success of this initiative. Thanks especially to our colleagues and friends from the Irving Fisher Committee of the BIS with whom we have worked hard over the last year. Of course, special thanks also go to Dr. Rashad Cassim, Chairman of the Committee, for his support in these initiatives.

Data science is not a brand new offspring of big data and artificial intelligence. Exactly sixty years ago, in 1962, John Wilder Tukey wrote the essay 'The Future of Data Analysis', widely recognized as a cornerstone in the history of data science. He also used the term *spotty data,* anticipating many of the limits you have described when dealing with big data from the web.

In his book *Concise Survey of Computer Methods* published in 1974, the computer scientist Peter Naur stressed that data science was the science of dealing with data, rather than that of establishing how data should be related to what they represent.

In those years, the first econometric and statistical software emerged. I am old enough to remember the first commercial products. These programs were quite expensive and could only run on mainframes. Only governments and big universities could afford the annual license. No one would have ever imagined that about twenty years later, in early 1998, shortly after the announcement of the release of the Netscape source code, the 'open source' label would be created.

In the last forty years, we have witnessed astounding technological and methodological developments, which have helped to shape data science as an interdisciplinary field at the intersection of economics, statistics, social sciences, mathematics, physics and engineering.

All these efforts have paid off. Today, two years into the pandemic crisis, operators in both the private and public sector have acknowledged the irreplaceable role of artificial

intelligence and data science in collecting and processing information. Their adoption has spread rapidly among companies and central banks, enabling them to develop new statistical frameworks (nontraditional data) to cope with the rapidly evolving new environment.

The 45 papers presented at this workshop focus on key issues in central banking activities concerning network analysis, natural language processing for extracting information from textual documents, innovative tools for assuring data quality, new platforms for storing data and using distributed computing to compile statistics of interest.

They have made the role and value of data science for economic analysis crystal clear. We have learned how in recent years data science applications have been used more and more frequently, in different areas of central bank activities such as monetary policy, the compilation of statistics, banking supervision, communication and many more besides.

The importance of the topics discussed is testified to by the number and quality of the presenters and the audience: the discussions were followed by more than 600 researchers from almost 70 central banks and 50 other organizations, between institutions and universities. I have no doubt that all the participants benefited from the very high level of the interaction.

Of course, this is not the first time we have met to discuss these issues. On different occasions in the last two years, economists, statisticians and other field experts have convened to share views and present solutions whose adoption could benefit different institutions, both private and public.

Last year, at the beginning of June, we had a three-day meeting on 'Big Data and Data Science Applications and Development at Central Banks' organized by CEMLA and the Bank of Spain

A few months later, in October, we held the first edition of this 'Workshop on Data Science and Central Banking' organized by the Bank of Italy and the Irving Fisher Committee of the BIS.

The Bank of Italy's participation is this field of work, from the very early days, is no accident.

We have always placed great value on granular data collection, from structured and unstructured data sources, as well as on exploring innovative statistical and machine learning methodologies. A few years ago we set up a multidisciplinary team, with a broad skillset, devoted to digging out different sources of online or nontraditional data, to improving the accuracy of our macroeconomic forecasts, and to evaluating financial systemic and idiosyncratic risks. We were interested in exploring the potential for digital regulatory reporting and compliance checks and in devising early warning indicators of new emerging risks.

Before closing, allow me to underscore once again the importance and value of good cooperation among public authorities and between them and private sector companies

who often own big data. This cooperation is key to relaxing the constraints on collecting, storing and analysing nontraditional data. I believe our workshop has confirmed the valuable results that can be achieved, in terms of knowledge sharing, by working together.

The Bank of Italy is firmly committed to carrying out research on these emerging topics and to cooperating with other national and international institutions with the aim of better serving our societies.

I am sure we will have many future opportunities to share the results of our analytical efforts and I hope we will soon be able to meet again in person.

Once more, thank you all.