

## From Provenance to Explanations

Data-intensive systems effectively rule our lives. Highly complex data analysis is employed in smart cities, in scientific experiments, to support medical decisions, for marketing, and in many other contexts. An artifact of this great progress is that decisions and actions, possibly with crucial effects, are often being made through complex computation, whose logic is usually not publicly available. This raises significant concerns of various flavors, such as: is a result based on private data? If so, was the data used only for legitimate purposes? If a result is of high importance, is it based on highly reliable data? Is the underlying computation correct? These and similar concerns are all related to the provenance of data. Provenance intuitively explains a computation result by capturing the way in which different parts of the data are used, combined and manipulated by a query or an application. I will explain our research on provenance solutions for data analytics and the way it addresses questions such as the above.