Abstract:

We study the complexity of query answering using views in a probabilistic XML setting, identifying large classes of XPath queries – with child and descendant navigation and predicates – for which there are efficient (PTime) algorithms. We consider this problem under the two possible semantics for XML query results: with persistent node identifiers and in their absence. Accordingly, we consider rewritings that can exploit a single view, by means of compensation, and rewritings that can use multiple views, by means of intersection. Since in a probabilistic setting queries return answers with probabilities, the problem of rewriting goes beyond the classic one of retrieving XML answers from views. In particular, for both semantics of XML queries, we show that, even when XML answers can be retrieved from views, their probabilities may not be computable.

About the Speaker:

Bogdan Cautis is a Professor of Computer Science at the University of Paris-Sud. Until recently he was with the Computer Science and Networks Department of Télécom ParisTech. He received his Habilitation (HdR) in March 2012, from Universite Pierre et Marie Curie and his Ph.D. in September 2007 from the University of Paris XI – working in the Gemo research team of INRIA Futurs, advised by INRIA DR Serge Abiteboul and Tova Milo from Tel Aviv University. His recent research interests lie in the broad area of Web data management and information retrieval: data management on the Web, data extraction, social networks, search, recommender systems, XML and semi-structured databases.