Lecture series on the subject "The interplay between fundamental computer science and data science"
Chair: Prof. Dr. Florin Manea

Abstract and brief biography

2. Solon Pissis: On String Sanitization

Abstract: Strings are being used to represent individuals' data arising from a large range of domains, such as transportation, web analytics, or molecular biology. For example, a string can represent: an individual's movement history, with each letter corresponding to a visited location; an individual's purchasing history in a retailer, with each letter corresponding to a purchased product; or a fragment of the genome sequence of a patient, with each letter corresponding to a DNA base. Analyzing such strings is thus necessary in many different applications. To support these applications, string data must often be disseminated beyond the party that has collected them. However, individuals' data dissemination has led to justified privacy concerns due to the exposure of confidential information. Thus, it may be necessary to sanitize a string prior to its dissemination, so that confidential knowledge is not exposed. At the same time, it is important to preserve the utility of the sanitized string, so that data protection does not outweigh the benefits of disseminating the string to the party that disseminates or analyzes the string, or to the society at large. In this talk, we will visit several combinatorial algorithms, which offer privacy/utility trade-off guarantees within the process of string sanitization.

Brief Bio:
Dr. Solon Pissis read Computer Science at the University of Athens (2003-2007) before completing his M.Sc in High-Performance Computing at the University of Edinburgh (2007-2008). He obtained his Ph.D in Computer Science from King's College London (2008-2012). He was a Research Associate at the University of Florida and a Visiting Scientist at the Heidelberg Institute for Theoretical Studies (2012-2013). He was appointed Lecturer (2013) and then Senior Lecturer (2018) in Computer Science at King's College London. In 2019 he joined the CWI, where he leads the Algorithms and Data Structures for Sequence Analysis group. In 2020, he joined the Faculty of Science at the Vrije Universiteit Amsterdam as (part-time) Associate Professor. His area of research within computer science is algorithms and data structures. Research topics he is interested in include

algorithms and data structures on strings and trees for pattern matching, indexing, comparison, and finding regularities. Applications thereof he is interested in include data mining, bioinformatics, data compression, and information retrieval.