

# Ioana Manolescu

Senior researcher at Inria Saclay  
Professor at Ecole Polytechnique



## From Fact-Checking to Journalistic Data Integration

Wednesday, April 21, 2021  
11:30am – 12:30pm Pacific time



Scan or visit URL to register:  
<https://data.cs.sfu.ca/gWPF>

### Abstract

Professional journalism is of utmost importance nowadays. It is a main feature distinguishing dictatorships from democracies, and a mirror sorely needed by society to look upon itself and understand its functioning. In turn, understanding is necessary for making informed decisions, such as political choices.

With the world turning increasingly digital, journalists need to analyze very large amounts of data, while having no control over the structure, organization, and format of the data. Since 2013, my team has been working to understand data journalism and computational fact-checking use cases, to identify and develop tools adapted for this challenging setting. My talk will present highlights from several years of research in this area. First, I will describe efforts to improve the accessibility to reference data sources which serve as trusted background for verifying (fact-checking) statistical claims. Second, I will present ConnectionLens, a system for integrating very heterogeneous data sources as graphs, leveraging Information Extraction and Entity Disambiguation. Information needs on ConnectionLens graphs are answered through graph-based keyword search; I will discuss the characteristics of ConnectionLens keyword search which make it harder in our setting, and our solution to these issues. Finally, I will present an application we currently develop, in collaboration with Stéphane Horel, an investigative journalist from Le Monde.

### Biography

Ioana Manolescu is a senior researcher at Inria Saclay and a part-time professor at Ecole Polytechnique, France. She is the lead of the CEDAR INRIA team focusing on rich data analytics at cloud scale. She is also the scientific director of LabIA, a program ran by the French government whereas AI problems raised by branches of the local and national French public administration are tackled by French research teams. She is a member of the PVLDB Endowment Board of Trustees, and has been Associate Editor for PVLDB, president of the ACM SIGMOD PhD Award Committee, chair of the IEEE ICDE conference, and a program chair of EDBT, SSDBM, ICWE among others. She has co-authored more than 150 articles in international journals and conferences and co-authored books on "Web Data Management" and on "Cloud-based RDF Data Management". Her main research interests algebraic and storage optimizations for semistructured data, in particular Semantic Web graphs, novel data models and languages for complex data management, data models and algorithms for fact-checking and data journalism, a topic where she is collaborating with journalists from Le Monde. She is also a recipient of the ANR AI Chair titled "SourcesSay: Intelligent Analysis and Interconnexion of Heterogeneous Data in Digital Arenas" (2020-2023).