

## DATA-DRIVEN REVELATION?

Epistemological tensions in investigative journalism in the age of “big data”

As an increasing number of reporters see databases and algorithms as appropriate means of doing investigation, journalism has been challenged in recent years by the following question: to what extent would the processing of huge datasets allow journalists to produce new types of revelations that rely less on normative assumptions? Drawing on the analysis of a particular investigation by the San Francisco-based Center for Investigative Reporting, this article points out the existence of epistemological tensions in the making of journalistic revelations that involve the processing of vast amounts of data. First, I show that the design of data-processing artifacts can match the traditional epistemology of journalistic investigation, but only with great efforts and resources from the organization. Second, I point out that the use of these artifacts by journalists follows two opposite paths to produce the revelation: a “hypothesis-driven” path and a “data-driven” path. Such findings contribute to a better understanding of how news organizations produce justified beliefs, as data-processing artifacts become major components of the newsroom’s environment.