

<center><span style="font-family: verdana; font-size: 20px; color: #275c86;">Montenegro, Budva, Becici, 28 september - 02 october 2015</span></center>



Contribution ID: 76

Type: **not specified**

## Social Data Collection and Processing Framework

*Thursday 1 October 2015 16:20 (15 minutes)*

Modern information technologies have an impact in research in all possible areas of knowledge, and the humanities are not an exception. Some of them, such as psychology and sociology, can use observations of the human behavior and the opinions of individuals and communities as a base for research. One of the possible ways to acquire the data for the base is from social networking services, which in their present state may serve as a rich source of information about people. However, gathering, storage and especially processing of such data is a nontrivial task because of its large amount, which is only grows with time, complex and diverse structure, suitability of psychological and sociological methods for automatic application. This work outlines a framework for managing of large amounts of social data for applying psychological and sociological methods. Here it is described, how the framework handles gathering, storing, and processing complex, interconnecting data by using columnar database Hbase as a core storage. In addition, an example of the framework's work with performance results is presented, future improvements are discussed.

**Author:** Prof. DEGTYAREV, Alexander (Professor)

**Co-authors:** Prof. BOGDANOV, Alexander (St.Petersburg State University); Mr GUSCHANSKY, Dmitry (St.Petersburg State University)

**Presenter:** Mr GUSCHANSKY, Dmitry (St.Petersburg State University)

**Session Classification:** Computations with Hybrid Systems (CPU, GPU, coprocessors)