Topic: «Collecting and analysis of Weibo data».

Task: Development of scalable framework for gathering data from Weibo.

Framework requirements:

1. The framework should allow you to search for posts and collect information about search results by specified criteria: post content, comments, user profiles.
2. Scalability. The framework should be able to be deployed on several compute nodes to increase performance.
3. Recursive data crawling. In case when during processing of any entity of the framework, references to other entities described in the framework appear, these entities must be processed, and a connection must be established between them. The recursion depth must be maintained.

Theoretical requirements:

1. Analogues review. Description of existing solutions for data collection from Weibo, such as libraries and API. Must include the key features of the tool, an example of work, the main pros and cons.
2. A review of the literature on the topic of analyzing data from social networks. Must include the review of at least 20 foreign scientific articles with at least 10 citations. The analysis of each publication should include a description of the key idea of the article, the main provisions, the authors' articles for discussion, the advantages and disadvantages of the work, and how the results of the article can be applied to the main task.
3. Architecture of the framework. A detailed description, including the advantages, bottlenecks of the proposed solution, as well as an architecture diagram in the UML format.