

*O. Stoianova, Saint Petersburg State University, Saint Petersburg, Russia, o.stoianova@spbu.ru*

*V. Bakotin, Saint Petersburg State University, Saint Petersburg, Russia, vasilibakotin@gmail.com*

# Text processing and machine learning methods in HRM applications: opportunities and features

New tasks of human resource management require the analysis of huge volumes of semi-structured text information. Methods of text processing and machine learning can significantly improve its effectiveness in case they take into consideration the features of tasks to be solved. The article describes actual analytical problems of human resource management, characteristics of information support of these problems, shortcomings and assumptions of frequently used methods of both classes in the tasks context. An example of applying text processing and machine learning methods in the task of compliance assessment is given in the article as well.

**Keywords:** text processing, machine learning, human resources management, semi-structured text data, compliance assessment, career planning, feedback analysis.

## Introduction

Human resource management (HRM) has changed significantly over the past few years. The competition for highly qualified personnel in many business areas requires new approaches to attracting top-skilled employees and forming internal and external personnel reserves. A good example is the high-technology industry, where ‘two-thirds of employees believe they could find a better job in less than 60 days if they only took the time to look’ [1]. In such conditions the tasks of employee retention and engagement become the key ones. To solve them many companies start using information systems for recruiting and career management based on personnel estimations. According to [2] 22% of companies survey employees quarterly or more often and only 14% never survey employees, but in many cases analytical survey results characterize only the overall situation in the company. For a deeper analysis, HR managers have to manually process large amounts of text data. This factor causes a great demand for ana-

lytical tools which can help HR managers to figure out employees’ mood and behavior patterns.

Another serious challenge HRM faces today is a human resource information security. In the modern business environment the boundaries between work and private life are erased from the point of communication means in use. The attempt to separate corporate and private informational channels fails because of necessity of rapid access policy which is provided by technical tools such as email redirection, common chats and so on, resulting in data loss threats appearance. 24/7 work style is the reason why it becomes almost impossible to control all the information flows from and to the organization. These huge amounts of data may contain a lot of relevant insides about possible destructive intentions the employees possess towards the company and about specific reasons of these intentions which can be analyzed on time thus preventing negative consequences. Since most of the data people send and receive daily turns out to be a semi-structured text data, methods of text mining can be applied to find possible solutions in informa-