

Characterization of Russian Deep Web

Denis Shestakov¹, Natalia Vorontsova²

¹ Turku Centre for Computer Science, University of Turku, Turku,
FIN-20520 Finland

² IM Sechenov Moscow Medical Academy, Moscow, 121019 Russia
Corresponding e-mail address: denis.shestakov@utu.fi

Abstract: The significant portion of the Web is hidden behind search forms and not indexed by conventional search engines. This part of the Web is known as the deep Web. Pages in the deep Web are dynamically generated in response to queries submitted via search forms. In this work, we studied the Russian part of deep Web. Our main goal was to estimate the number of deep Web sites in the Russian deep Web. The presented study is a first work devoted to the certain part of deep Web, which is formed on the basis of some particular language usage.