

МОЛБА ЗА ОДОБРАВАЊЕ ТЕМЕ МАСТЕР РАДА

Молим да ми се одобри израда мастер рада под насловом:

Приказ развојног оквира *Hadoop Map-Reduce* на примеру анализе геномских секвенци

Значај теме и области:

У данашње време на дневном нивоу веома брзо се стварају огромне количине разноврсних података (*BigData-3V*), а врло мало тих података стигну да буду обрађени. *Hadoop Map-Reduce* представља рачунарску парадигму за обраду огромних количина података, дистрибуираних на различите сервере. Што је врло битно, то ради ефикасно и јефтино. То је скуп алата од којих ће неки од бити представљени у овом раду, попут: *HDFS*, *YARN*, *Hbase*, *MRUnit...* Већина њих су писани и користе се помоћу програмског језика Јава.

Специфични циљ рада:

Улазни подаци су геномске секвенце (нпр. ДНК). Задатак је пописати и пребројати све подсеквенце дате дужине. Процес је дистрибуиран на више чворова и дели се на 2 фазе: прва је мапирање (у коју је укључено и филтрирање и сортирање) свих могућих подсеквенци, а друга обрада и пребројавање тако сређених подсеквенци. Циљ је сагледати реалне добити од примене развојног оквира *Hadoop Map-Reduce*, у односу на уобичајен централизован приступ развоју.

Остале битне информације:

Пројекат ће бити отвореног кода и јавно доступан путем сервиса Гитхаб (*Github*).

Почетна литература:

1. Tom White (2015). *Hadoop: The Definitive Guide, 4th Edition*, O'Reilly Media
2. Donald Miner, Adam Shook(2012). *MapReduce Design Patterns*, O'Reilly Media
3. <https://hadoop.apache.org>

Немања Вељковић,
1007/2012, Информатика

Сагласан ментор: др Саша Малков

(својеручни потпис студента)

(својеручни потпис ментора)

(датум подношења молбе)

Чланови комисије

1. др Владимир Филиповић
2. др Јована Ковачевић

Катедра **за рачунарство и информатику** је сагласна са предложеном темом.

(шеф катедре)

(датум одобравања молбе)