

SADRŽAJ

SAŽETAK	4
ABSTRACT.....	5
Uvod	6
1. Big Data.....	7
1.2 Pojam 'Big Data'	7
1.3 Karakteristike Big Data.....	8
1.3.1 Volumen.....	9
1.3.2 Raznolikost	9
1.3.3 Brzina	9
1.3.4 Vrijednost.....	9
1.3.5 Vjerodostojnost.....	10
2. Vrste velikih podataka.....	10
2.1 Glavne komponente velikih podataka.....	11
2.2 Prednosti velikih podataka	12
2.3 Istorija velikih podataka	13
3. Metode i tehnike za optimizaciju skladištenja	14
3.1 Kompresija podataka.....	14
3.1.1 Vrste kompresije podataka	15
3.1.2 Tehnike kompresije podataka	16
3.1.3 Algoritmi kompresije podataka.....	17
3.1.4 Najbolje prakse za kompresiju podataka	19
3.2 Indeksiranje podataka	21
3.2.1 Strategije indeksiranja zasnovane na stablu.....	22
3.2.2 Bitmap indeksi.....	24
4. Apache Hadoop.....	25
4.1 HDFS	26

4.2 MapReduce	28
4.3 YARN	30
4.4 Optimizacija skladištenja podataka u Apache Hadoop sistemu	34
5. BIG DATA alati za skladištenje i analizu podataka	38
5.1 Apache Spark.....	39
5.2 Apache Kafka.....	42
5.3 Apache Flink.....	45
5.3.1 Apache Flink arhitektura i ključne komponente.....	45
6. Bezbjednost i sigurnost u Big Data	47
6.1 Šifrovanje i anonimizacija podataka	48
6.1.1 Tehnike anonimizacije podataka.....	50
6.1.2 Tehnike šifrovanja podataka	50
ZAKLJUČAK	51
LITERATURA	52
POPIS SLIKA.....	54