

*PERBANDINGAN METODE MAPREDUCE BERBASIS SINGLE  
NODE HADOOP PADA APLIKASI WORD COUNT*

*Diajukan Sebagai Syarat Untuk Menyelesaikan*

*Pendidikan Program Strata-1 Pada*

*Jurusan Teknik Informatika*



Oleh:

Elvira Oktaviani

NIM : 09021281823062

**Jurusan Teknik Informatika**  
**FAKULTAS ILMU KOMPUTER UNIVERSITAS SRIWIJAYA**  
**2023**

**LEMBAR PENGESAHAN SKRIPSI**

**PERBANDINGAN METODE MAPREDUCE BERBASIS SINGLE NODE HADOOP  
PADA APLIKASI WORD COUNT**

Oleh:

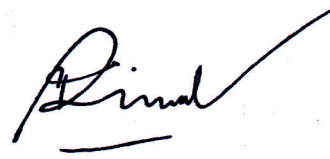
Elvira Oktaviani

NIM : 09021281823062

Indralaya, 7 Juli 2023

Pembimbing I

Pembimbing II,



Alvi Syahrini Utami, M.Kom.

Mastura Diana Marieska, S.T., M.T.

NIP. 197812222006042003

NIP. 198603212018032001

Mengetahui,

Ketua Jurusan Teknik Informatika



Alvi Syahrini Utami, M.Kom.

NIP. 197812222006042003

## TANDA LULUS UJIAN KOMPREHENSIF SKRIPSI

Pada hari Jum'at tanggal 23 Juni 2023 telah dilaksanakan ujian sidang skripsi oleh Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya

Nama : Elvira Oktaviani


NIM : 09021281823062

Judul : Perbandingan Metode *Mapreduce* Berbasis *Single Node Hadoop* Pada Aplikasi *Word Count*

dan dinyatakan LULUS.


1. Ketua Penguji

Yunita, M.Cs.  
NIP. 198306062015042002

  
.....


2. Penguji I

Rifkie Primartha, M.T.  
NIP. 197706012009121004

  
.....

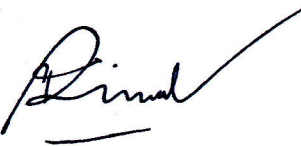
3. Pembimbing I

Alvi Syahrini Utami, M.Kom.  
NIP. 197812222006042003

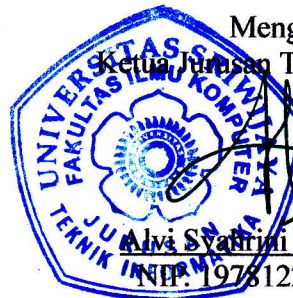
  
.....

4. Pembimbing II

Mastura Diana Marieska, S.T., M.T.  
NIP. 198603212018032001

  
.....

Mengetahui,  
Ketua Jurusan Teknik Informatika



Alvi Syahrini Utami, M.Kom.  
NIP. 197812222006042003

## HALAMAN PERNYATAAN BEBAS PLAGIAT

Yang bertanda tangan dibawah ini:

Nama : Elvira Oktaviani  
NIM : 09021281823062  
Program Studi : Teknik Informatika Reguler  
Judul Skripsi : Perbandingan Metode *MapReduce* berbasis *Single Node Hadoop* pada Aplikasi *Word Count*

Hasil pengecekan software *iThenticate/Turnitin* : 7%

Menyatakan bahwa laporan projek saya merupakan hasil karya sendiri dan bukan hasil plagiat. Apabila ditemukan unsur plagiat dalam laporan projek ini, maka saya bersedia menerima sanksi akademik dari Universitas Sriwijaya sesuai dengan ketentuan yang berlaku.

Demikian, pernyataan ini saya buat dengan sebenarnya dan tidak ada paksaan oleh siapapun.



Indralaya, Juli 2023



Elvira Oktaviani  
NIM. 09021281823062

## **MOTTO DAN PERSEMBAHAN**

Motto :

“Yang harus itu bukan berhasil, tapi berusaha”

Kupersembahkan Karya Tulis ini kepada :

- Kedua Orang Tuaku
- Ayuk dan Kakak
- Keluarga Besar
- Sahabat dan Teman Seperjuangan
- Para Guru dan Dosen
- Fakultas Ilmu Komputer Universitas Sriwijaya
- Diri Sendiri

## ABSTRACT

Word Count is a type of task used to count the occurrences of unique words in a text file. Processing time is an important factor to consider in meeting the standards of Big Data processing. In the context of Big Data processing, Hadoop MapReduce serves as a framework used to develop software and process large-scale data in parallel. The conducted research involved the processing of text files using the MapReduce method on the Hadoop Distributed File System (HDFS) using a single node, comparing the results of word count processing with and without the use of MapReduce. The research findings indicate that the implementation of Word Count without using MapReduce offers better speed and scalability in processing Indonesian language text data on a Hadoop single node. Additionally, the comparison of processing time between the Word Count program with Hadoop-based MapReduce and the Word Count program without MapReduce shows that the latter has faster processing time. A significant reduction in processing time, up to 95% for a 5 MB file, can be achieved by using the Word Count method without MapReduce, although the level of reduction decreases with increasing file size.

**Keywords** : Big Data, Word Count, MapReduce, HDFS, Hadoop Single Node

## ABSTRAK

*Word Count* merupakan jenis pekerjaan yang digunakan untuk menghitung kemunculan kata-kata unik dalam sebuah *file* teks. Kecepatan dalam pemrosesan data merupakan faktor penting yang harus dipertimbangkan dalam pemenuhan standar pemrosesan *Big Data*. Dalam konteks pemrosesan *Big Data*, *Hadoop MapReduce* menjadi *framework* yang digunakan untuk mengembangkan perangkat lunak dan melakukan pemrosesan data yang besar secara paralel. Penelitian yang dilakukan melibatkan pemrosesan *file* teks menggunakan metode *MapReduce* pada *Hadoop Distributed File System (HDFS)* menggunakan *single node* dengan membandingkan hasil pemrosesan *word count* dengan dan tanpa menggunakan metode *MapReduce*. Penelitian ini memberikan hasil bahwa implementasi *Word Count* tanpa menggunakan metode *MapReduce* menawarkan kecepatan dan skalabilitas yang lebih baik dalam pemrosesan data teks bahasa Indonesia dibandingkan dengan menggunakan *Hadoop single node*. Selain itu, perbandingan waktu pemrosesan antara program *Word Count* dengan metode *MapReduce* berbasis *Hadoop* dan *Word Count* tanpa *MapReduce* menunjukkan bahwa program *Word Count* tanpa *MapReduce* memiliki waktu pemrosesan yang lebih cepat. Reduksi waktu yang tinggi, hingga 95% pada *file* berukuran 5 MB, dapat dicapai dengan penggunaan metode *Word Count* tanpa *MapReduce*, meskipun tingkat reduksi tersebut menurun seiring dengan peningkatan ukuran *file*.

**Kata kunci :** *Big Data, Word Count, MapReduce, HDFS, Hadoop Single Node*

## KATA PENGANTAR

Puji dan syukur atas kehadiran Allah SWT yang telah memberikan hidayah, Rahmat, dan petunjuk-Nya sehingga penulis dapat menyelesaikan Laporan Skripsi dengan judul “Perbandingan metode *MapReduce* berbasis *Single Node Hadoop* pada Aplikasi *Word Count*” dengan baik, disusun guna memenuhi salah satu syarat menyelesaikan Pendidikan Strata-1 di Fakultas Ilmu Komputer Program Studi Teknik Informatika, Universitas Sriwijaya.

Dalam penulisan ini penulis menyadari bahwa tanpa adanya bimbingan, bantuan serta dukungan maupun petunjuk dari semua pihak, tidak mungkin laporan ini dapat diselesaikan. Oleh karena itu pada kesempatan ini penulis menyampaikan rasa terima kasih kepada :

1. Allah Subhanahu wa Ta'ala yang telah memberikan kesehatan serta kemudahan sehingga Skripsi ini dapat diselesaikan.
2. Nabi Muhammad Shallallahu 'Alaihi wa sallam yang telah memberikan jalan yang terang dan ilmu yang bermanfaat pada umatnya.
3. Kedua orang tua saya, Bunyamin dan Darnisa, yang telah memberikan dukungan materil dan moril serta doa restu, dan saudara saudara saya Eka Anggraini, Emilia, dan Taufik Hidayat yang memberikan dukungan demi kelancaran penyusunan skripsi ini.
4. Ibu Alvi Syahrini Utami, M.Kom. selaku dosen pembimbing I dan Ketua Jurusan Teknik Informatika Universitas Sriwijaya.
5. Ibu Mastura Diana Marieska, S.T., M.T. selaku dosen pembimbing II yang telah membimbing, mengarahkan dan memberikan motivasi kepada penulis dalam proses perkuliahan dan penyusunan skripsi ini.



6. Ibu Yunita, M.Cs. selaku ketua penguji dan Bapak Rifkie Primartha, M.T. selaku dosen penguji yang telah memberikan saran dan masukan untuk laporan skripsi ini.
7. Ibu Desty Rodiah, S.Kom., M.T. selaku dosen pembimbing akademik yang telah membimbing, mengarahkan, dan memberikan motivasi kepada penulis selama masa perkuliahan.
8. Seluruh dosen Program Studi Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.
9. Kak Ricy Firnando selaku admin Program Studi Teknik Informatika Indralaya, Kak Hafez selaku kepala admin lab Indralaya, serta seluruh staf dan pegawai di Fakultas Ilmu Komputer yang telah membantu dalam kelancaran administrasi dan akademik selama masa perkuliahan.
10. Melinia Rinda Afrilia yang telah banyak sekali membantu selama proses pengerjaan skripsi ini hingga selesai.
11. Kemas Fachir Raihan yang juga telah meluangkan waktu dan tenaga untuk membantu penulis dalam menyelesaikan program yang dikembangkan pada skripsi ini.
12. Dita, Ani, Aldiko, Yodies dan Rafi yang telah berbagi keluh kesah, motivasi, semangat, canda tawa dan atas segala bantuan selama masa perkuliahan.
13. M. Rafli Alamsyah Ratu yang telah banyak memberikan hari hari yang menyenangkan selama masa perkuliahan.
14. Semua teman – teman seperjuangan Teknik Informatika 2018.

Penulis menyadari dalam penyusunan skripsi ini masih banyak terdapat kekurangan disebabkan keterbatasan pengetahuan dan pengakaman, oleh karena itu kritik dan saran yang membangun akan sangat diharapkan untuk kemajuan penelitian berikutnya.

Akhir kata semoga skripsi ini dapat berguna dan bermanfaat bagi kita semua.

Indralaya, Juli 2023



Elvira Oktaviani

## DAFTAR ISI

HALAMAN JUDUL.....	i
LEMBAR PENGESAHAN SKRIPSI .....	ii
TANDA LULUS UJIAN KOMPREHENSIF SKRIPSI.....	iii
HALAMAN PERNYATAAN BEBAS PLAGIAT .....	iv
MOTTO DAN PERSEMBAHAN .....	v
ABSTRACT .....	vi
ABSTRAK .....	vii
KATA PENGANTAR .....	viii
DAFTAR ISI.....	xi
DAFTAR TABEL.....	xiv
DAFTAR GAMBAR .....	xv
BAB I PENDAHULUAN .....	I-1
1.1    Pendahuluan .....	I-1
1.2    Latar Belakang Masalah.....	I-1
1.3    Rumusan Masalah .....	I-3
1.4    Tujuan Penelitian.....	I-3
1.5    Manfaat Penelitian.....	I-4
1.6    Batasan Masalah.....	I-4
1.7    Sistematika Penulisan.....	I-4
1.8    Kesimpulan.....	I-5
BAB II KAJIAN LITERATUR .....	II-1

2.1	Pendahuluan .....	II-1
2.2	Landasan Teori .....	II-1
2.2.1.	Word Count.....	II-1
2.2.2.	Apache Hadoop .....	II-4
2.2.3.	Map Reduce .....	II-7
2.2.4.	HDFS .....	II-8
2.2.5.	RUP .....	II-10
2.3	Penelitian Lain yang Relevan.....	II-12
2.4	Kesimpulan.....	II-14
BAB III METODE PENELITIAN.....		III-1
3.1	Pendahuluan .....	III-1
3.2	Pengumpulan Data .....	III-1
3.2.1.	Jenis dan Sumber Data .....	III-1
3.2.2.	Metode Pengumpulan Data .....	III-2
3.3	Tahapan Penelitian .....	III-2
3.3.1.	Menentukan Kerangka Kerja .....	III-2
3.3.2.	Menetapkan Kriteria Pengujian.....	III-4
3.3.3.	Menetapkan Format Data Pengujian.....	III-4
3.3.4.	Menentukan Alat yang Digunakan dalam Penelitian.....	III-6
3.3.5.	Melakukan Pengujian Penelitian.....	III-6
3.3.6.	Melakukan Analisa dan Menarik Kesimpulan Penelitian.....	III-8
3.4	Metode Pengembangan Perangkat Lunak .....	III-9
3.5	Kesimpulan.....	III-10
BAB IV PENGEMBANGAN PERANGKAT LUNAK .....		IV-1
4.1	Pendahuluan .....	IV-1

4.2	Rational Unified Process (RUP).....	IV-1
4.3	Kesimpulan.....	IV-16
BAB V HASIL DAN PEMBAHASAN.....		V-1
5.1	Pendahuluan .....	V-1
5.2	Hasil Penelitian.....	V-1
5.2.1.	Konfigurasi Penelitian.....	V-1
5.2.2.	Data Hasil Pengujian.....	V-2
5.3	Analisis Hasil Penelitian .....	V-4
5.4	Kesimpulan.....	V-6
BAB VI KESIMPULAN DAN SARAN .....		VI-1
6.1	Pendahuluan .....	VI-1
6.2	Kesimpulan.....	VI-1
6.3	Saran.....	VI-2
DAFTAR PUSTAKA .....		xvi

## DAFTAR TABEL

	Halaman
II-1. <i>Sample MapReduce Mapper Function Output</i> .....	II-2
II-2. <i>Sample MapReduce Reduce Function Output</i> .....	II-3
III-1. Pembagian Ukuran Data Uji .....	III-2
III-2. Format Data Pengujian <i>Word Count</i> dengan <i>MapReduce</i> .....	III-5
III-3. Format Data Pengujian <i>Word Count</i> tanpa <i>MapReduce</i> .....	III-5
III-4. Hasil Perbandingan Pengujian .....	III-10
IV-1. Kebutuhan Fungsional.....	IV-2
IV-2. Kebutuhan Non-Fungsional .....	IV-2
IV-3. Definisi Aktor.....	IV-5
IV-4. Definisi <i>Use Case</i> .....	IV-5
IV-5. Skenario <i>Use Case</i> Menambahkan Dokumen Data Uji .....	IV-6
IV-6. Definisi Implementasi Kelas .....	IV-15
IV-7. Perancangan Pengujian Unit .....	IV-17
IV-8. Implementasi Pengujian <i>Word Count</i> dengan <i>MapReduce</i> .....	IV-18
V-1. Implementasi Pengujian <i>Word Count</i> tanpa <i>MapReduce</i> .....	V-2
V-2. Implementasi Pengujian <i>Word Count</i> tanpa <i>MapReduce</i> .....	V-3
V-3.Reduksi Waktu Pengujian <i>Word Count</i> dengan <i>MapReduce</i> dan tanpa <i>MapReduce</i> .....	V-3

## DAFTAR GAMBAR

	Halaman
II-1 Implementasi <i>Word Count</i> dengan <i>MapReduce</i> .....	II-4
II-2 Arsitektur <i>Hadoop</i> .....	II-5
II-3 Arsitektur <i>Hadoop Single Node</i> dan <i>Multi Node</i> .....	III-6
II-4 <i>Proposed Hadoop Model for Big Data Implementation</i> .....	III-7
II-5 Proses <i>MapReduce</i> .....	III-8
II-6 Arsitektur <i>HDFS</i> .....	III-9
II-7 <i>Secondary NameNode Task</i> .....	III-10
II-8 Tahap Pengembangan RUP .....	III-11
III-1 Tahapan Penelitian .....	III-3
III-2 Kerangka Kerja Penelitian .....	III-7
IV-1 Diagram <i>Use Case</i> .....	IV-5
IV-2 Diagram Aktivitas <i>Word Count</i> tanpa <i>MapReduce</i> .....	IV-12
IV-3 <i>Sequence Diagram Word Count</i> tanpa <i>MapReduce</i> .....	IV-14
IV-4 Implementasi <i>Diagram Class</i> .....	IV-14
V-1 Grafik Perbandingan Performasi Waktu Pengujian <i>Word Count</i> dengan <i>MapReduce</i> dan tanpa <i>MapReduce</i> .....	V-5

# **BAB I**

## **PENDAHULUAN**

### **1.1 Pendahuluan**

Bahasan yang dikaji dalam bagian ini memuat latar belakang, perumusan masalah, manfaat dari penelitian, tujuan dari penelitian, batasan-batasan masalah dalam penelitian yang dilakukan serta sistematika penulisan, yang diawali dengan uraian terkait implementasi *Word Count* menggunakan metode *MapReduce* berbasis *Hadoop Distributed File System*.

### **1.2 Latar Belakang Masalah**

Menurut Prabowo dan Abdurohman (2017) *Word Count* merupakan jenis *job* yang digunakan untuk menghitung kemunculan sebuah kata yang terbilang unik yang terdapat pada tipe data *file* teks dengan keluaran berupa jumlah berapa kata yang muncul pada pengujian tersebut. *Word Count* sendiri dapat mempengaruhi akurasi dalam analisis sentimen terkait pengelompokan kata positif, negatif dan netral pada *text mining* yang merupakan teknik mengekstraksi informasi dari data teks.

Data dalam jumlah yang sangat besar, variatif dan sangat cepat pertumbuhannya disebut dengan *Big Data*. Faktor kecepatan pada komputasi tersebut berpengaruh pada komponen yang memenuhi standar pemrosesan terkait. Maka dari itu, untuk menunjang proses komputasi tersebut sangat penting untuk mengaplikasikan suatu algoritma tertentu yang membuat pemrosesan data tersebut menjadi efektif dan efisien (Taqwin et al, 2018).



Adapun definisi *Big data* yakni perpaduan sistem teknologi memungkinkan kita untuk mengelola sejumlah besar data berbeda dengan kecepatan dan akurasi yang wajar dalam menganalisis. Terdapat 3 atribut pada *Big data* yakni *variety*, *volume*, serta *velocity* (Hurwitz et al., 2013). Utilitas teknologi yang dimiliki dari proses *big data* meliputi melakukan koneksi data penyimpanan dengan memanfaatkan *key-value store (KVS)* juga melakukan komputasi secara paralel dengan mengaplikasikan *MapReduce*. (Sujana, 2013).

Menurut Padhy (2012) volume data yang akan diproses oleh aplikasi *cloud* tumbuh jauh lebih cepat daripada daya komputasi. Pertumbuhan ini menuntut strategi baru untuk memproses dan menganalisis informasi. *Hadoop MapReduce* telah menjadi model komputasi yang kuat untuk mengatasi masalah ini. *Hadoop HDFS* menjadi lebih populer di antara semua alat *Big Data* karena *open source* dengan skalabilitas yang fleksibel, lebih sedikit mengeluarkan biaya & memungkinkan penyimpanan data dalam bentuk apa pun tanpa perlu memiliki tipe data atau skema yang ditentukan.

Berdasarkan uraian tersebut, *Hadoop MapReduce* adalah suatu framework untuk mengembangkan suatu software, serta mampu melakukan pengolahan kumpulan data yang berukuran besar dengan pemrosesan paralel pada *node* komputasi (Padhy., 2012 ). *MapReduce* dan HDFS merupakan bagian inti dari teknologi *Hadoop* di mana HDFS menyediakan penyimpanan data yang banyak dan beragam, dan *MapReduce* menyediakan pemrosesan data yang cepat. *MapReduce* biasanya mendukung *batch-based* paralelisasi skala tinggi dan pemrosesan yang andal. Kerangka kerja *Hadoop MapReduce*

menggunakan sistem *file* terdistribusi untuk membaca dan menulis datanya. Oleh karena itu, kinerja I/O dari *Hadoop MapReduce* sangat bergantung pada HDFS (Sahane et al., 2015).

Kajian yang akan diteliti melibatkan pemrosesan *big data* menggunakan metode *MapReduce* pada *Hadoop Distributed File System* (HDFS). Hasil pemrosesan akan dibandingkan dengan pemrosesan *word count* tanpa *MapReduce*.

### 1.3 Rumusan Masalah

Perumusan masalah yang sesuai dengan penjelasan pada latar belakang sebelumnya yaitu bagaimana melakukan implementasi *Word Count* dengan metode *MapReduce* berbasis *Hadoop Distributed File System* serta membandingkan hasil kecepatan pemrosesan tiap file pada *Word Count* dengan metode *MapReduce* berbasis *Hadoop Distributed File System* dan tanpa *MapReduce*.

### 1.4 Tujuan Penelitian

Adapun pengujian yang dilakukan bertujuan untuk:

1. Melakukan implementasi atau penerapan program *Word Count* dengan metode *MapReduce* berbasis *Hadoop Distributed File System* dan *Word Count* tanpa *MapReduce*
2. Membandingkan waktu pemrosesan yang dihasilkan oleh program *Word Count* dengan *MapReduce* berbasis *Hadoop Distributed File System* dan program *Word Count* tanpa *MapReduce*

### 1.5 Manfaat Penelitian

Adapun manfaat pengujian yang dilakukan yaitu:

1. Mengetahui efektifitas waktu pemrosesan *word count* tanpa *MapReduce* dibanding *word count* menggunakan *MapReduce* dengan *Hadoop Distributed File System*.
2. Dijadikan rujukan untuk penelitian lain terkait *MapReduce* dan *Hadoop Distributed File System*.

### 1.6 Batasan Masalah

1. File yang digunakan dalam format .txt.
2. Data yang digunakan berupa *file* teks dengan ukuran 5 MB, 10 MB, 25 MB, 50 MB, 100 MB dan 200 MB
3. Implementasi *Hadoop File System* dilakukan pada *Single Node*.

### 1.7 Sistematika Penulisan

Sistematika dalam pembuatan tugas akhir ini meliputi:

## BAB I. PENDAHULUAN

Mengkaji seputar penelitian yang memuat latar belakang hingga sistematika penulisan.

## BAB II. KAJIAN LITERATUR

Mengkaji teori yang diimplementasikan diantaranya; *Word Count*, *MapReduce*, *Hadoop* serta penelitian yang berkaitan dengan pengujian yang akan dilakukan.

### **BAB III. METODOLOGI PENELITIAN**

Mengkaji uraian tahap pada pengujian yang dijelaskan terpecinci serta menguraikan kerangka kerja yang diimplementasikan dan membangun perencanaan yang berkaitan dengan keberlangsungan pengujian yang dilakukan.

### **BAB IV. PENGEMBANGAN PERANGKAT LUNAK**

Pengembangan perangkat lunak akan dibahas dalam bab ini berdasarkan metode yang telah ditentukan.

### **BAB V. HASIL DAN PEMBAHASAN**

Pada bab ini akan diuraikan hasil dari proses pengujian perangkat lunak yang telah dikembangkan dan melakukan analisis terhadap hasil pengujian.

### **BAB VI. KESIMPULAN DAN SARAN**

Sebagai bab penutup maka pada bab ini akan ditarik kesimpulan dari proses yang telah dilakukan dan terdapat saran yang dapat digunakan untuk penelitian berikutnya.

#### **1.8 Kesimpulan**

Bab ini mengkaji seputar penelitian yaitu pengujian terhadap implementasi *word count* menggunakan *MapReduce* berbasis *Hadoop Distributed File System* dan pengujian *word count* tanpa *MapReduce*.

## DAFTAR PUSTAKA

- Adawiyah, R., & Munir, S. (2020). Jurnal Informatika Terpadu. *Jurnal Informatika Terpadu*, 6(1), 14–19. <https://journal.nurulfikri.ac.id/index.php/JIT>
- Arief Prasetyo, Sofyan Noor Arief, & Rokhimatul wakhidah. (2021). Optimasi Pemrosesan Enkripsi Dan Dekripsi Rsa Pada Single Board Computer (Sbc) Dengan Pembagian Beban Komputasi Dalam Sistem Terdistribusi. *Jurnal Informatika Polinema*, 7(4), 7–12. <https://doi.org/10.33795/jip.v7i4.523>
- Engineering, C., Sahane, M., Sirsat, S., Khan, R., Mahal, N. S., & Baug, R. (2015). *Analysis of Research Data using MapReduce Word Count Algorithm*. 4(5), 8–11. <https://doi.org/10.17148/IJARCCE.2015.4542>
- Ghazi, M. R., & Gangodkar, D. (2015). Hadoop, mapreduce and HDFS: A developers perspective. *Procedia Computer Science*, 48(C), 45–50. <https://doi.org/10.1016/j.procs.2015.04.108>
- Huang, W., Wang, H., Zhang, Y., & Zhang, S. (2017). A novel cluster computing technique based on signal clustering and analytic hierarchy model using hadoop. *Cluster Computing*. <https://doi.org/10.1007/s10586-017-1205-9>
- Hurwitz, J., Nugent, A., Halper, F., & Kaufman, M. (2013). *Big data for dummies* (Vol. 336). Hoboken, NJ: John Wiley & Sons.
- Maurya, M., & Mahajan, S. (2012). Performance analysis of MapReduce programs on Hadoop cluster. *Proceedings of the 2012 World Congress on Information*

*and Communication Technologies, WICT 2012*, 505–510.  
<https://doi.org/10.1109/WICT.2012.6409130>

Prabowo, S. (2017). Implementasi Algoritma Penjadwalan untuk pengelolaan Big Data dengan Hadoop. *Indonesian Journal on Computing (Indo-JC)*, 2(2), 119.  
<https://doi.org/10.21108/indojc.2017.2.2.189>

Somantri, M. (2005). Membangun Sistem Komputasi Terdistribusi Dengan Pemrograman C++. *Transmisi*, 9(1 Juny 2005), 26 – 30.

Sujana, A. P. (2013). Memanfaatkan Big Data Untuk Mendeteksi Emosi. *Teknik Komputer Unikom*, 2(2), 1–4.

Taqwin, P. A. T., Osmond, A. B., & Latuconsina, R. (2018). Implementasi Metode Mapreduce Pada Big Data Berbasis Hadoop Distributed File System. *E-Proceeding of Engineering*, 5(1), 1013–1020.  
<https://doi.org/10.1109/ACDT.2017.7886152>

Technologies, N., Pal, A., Agrawal, P., & Jain, K. (2014). *2014 Fourth International Conference on Communication Systems and Network Technologies A Performance Analysis of MapReduce Task with Large Number of Files Dataset in Big Data Using Hadoop*.  
<https://doi.org/10.1109/CSNT.2014.124>

Tia, T. K., & Kusuma, W. A. (2018). Model Simulasi Pengembangan Perangkat Lunak Menggunakan Rational Unified Process (Rup). *Teknika: Engineering and Sains Journal*, 2(1), 33. <https://doi.org/10.51804/tesj.v2i1.226.33-40>

Yunadian, Y., Nuha, H. H., & Prabowo, S. (2020). *Pengolahan Data Polling berbasis Media Sosial menggunakan MapReduce pada Framework Hadoop*. 7(1), 2581–2591.