



## DAFTAR ISI

LEMBAR PENGESAHAN .....	iii
PERNYATAAN KEASLIAN PROYEK AKHIR .....	iv
KATA PENGANTAR .....	v
DAFTAR ISI .....	vii
DAFTAR GAMBAR .....	ix
DAFTAR TABEL .....	xi
DAFTAR LAMPIRAN .....	xii
INTISARI .....	1
BAB I PENDAHULUAN .....	2
1.1    Latar Belakang .....	2
1.2    Rumusan Masalah .....	3
1.3    Batasan Masalah .....	3
1.4    Tujuan Penelitian .....	4
1.5    Manfaat Penelitian .....	4
1.6    Sistematika Penulisan .....	4
BAB II TINJAUAN PUSTAKA .....	6
2.1    Lingkup Tinjauan Pustaka .....	6
2.2 <i>Data Processing</i> .....	11
2.3 <i>Data Quality</i> .....	12
2.4    Dataprof .....	14
2.5 <i>Cloud Composer</i> .....	17
2.6 <i>Great Expectations</i> .....	19
2.7    MongoDB .....	20
2.8    Bigquery .....	21
2.9    Hipotesis .....	22
BAB III METODOLOGI PENELITIAN .....	23
3.1    Peralatan .....	23
3.2    Bahan .....	24
3.3    Tahapan Penelitian .....	25
3.4    Rancangan Sistem .....	25



UNIVERSITAS  
GADJAH MADA

IMPLEMENTASI AIRFLOW DAN APACHE SPARK DALAM PERANCANGAN DATA PIPELINE UNTUK

DATA PROCESSING

AUTOMATION PADA CLOUD-BASED SERVER

Gigih Haryo Yudhanto, Yuris Mulya Saputra, S.T., M.sc., Ph.D.

Universitas Gadjah Mada, 2023 | Diunduh dari <http://etd.repository.ugm.ac.id/>

3.5	Instalasi dan Konfigurasi .....	26
BAB IV HASIL DAN PEMBAHASAN .....		50
4.1	Pengujian VPC <i>Peering</i> .....	50
4.2	Pengujian Performa Apache Spark .....	51
4.3	Pengujian Kualitas Data.....	53
4.4	Analisis Data Coingecko API .....	55
BAB V PENUTUP .....		61
5.1	Kesimpulan .....	61
5.2	Saran .....	61
DAFTAR PUSTAKA .....		62
LAMPIRAN .....		65