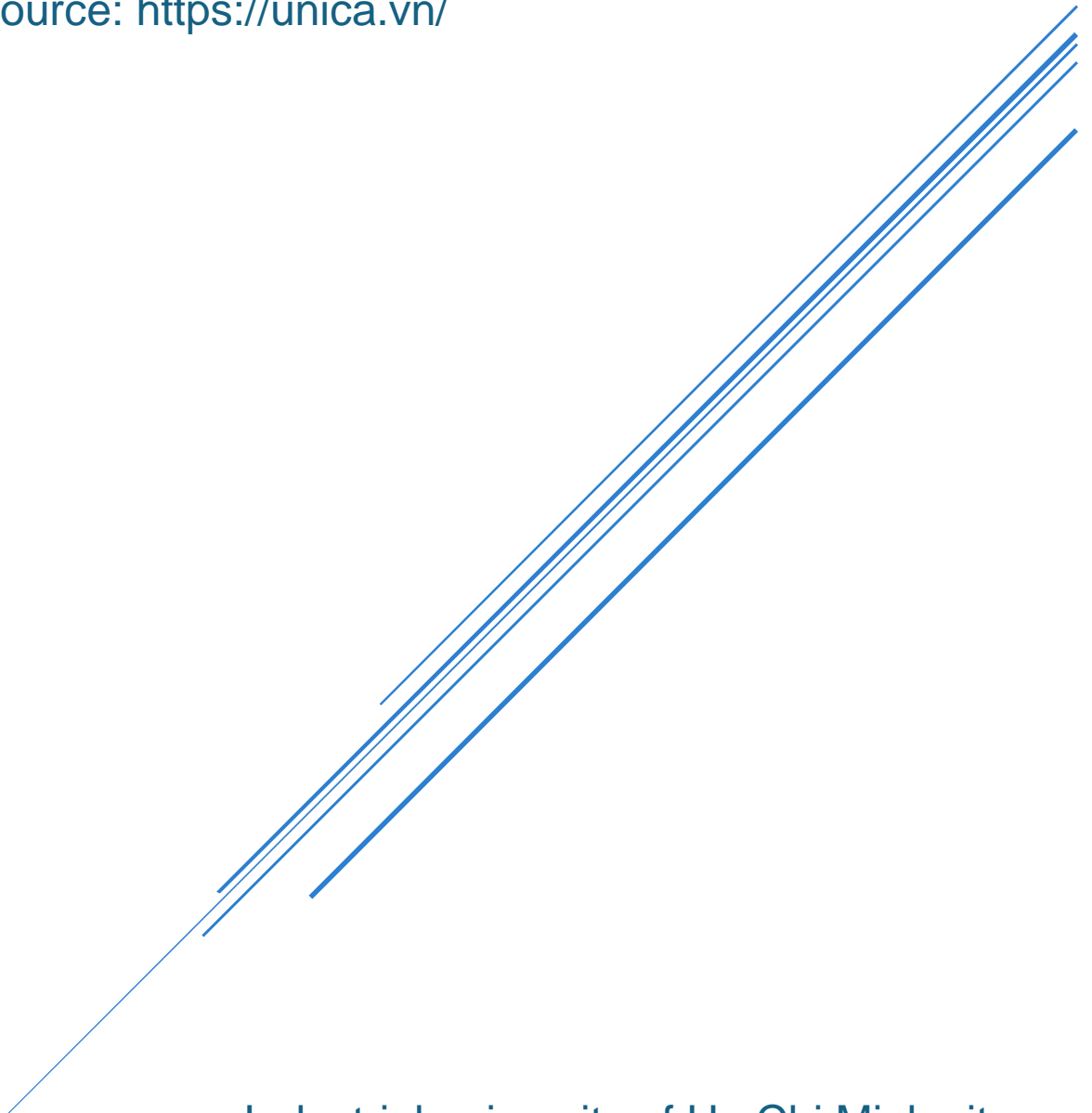


TITLE: BUILDING DATA PIPELINE TO COLLECT DATA FROM THE COURSE SALE WEBSITE

Data source: <https://unica.vn/>



Industrial university of Ho Chi Minh city
for Data Engineer

Student ID	Fullname	Task	Contribution rate (%)	Score (lecturer)
22695761	Nguyễn Thành Đạt	<ul style="list-style-type: none"> • Crawl data (technology, construction design) • Source code control • Data processing and cleaning 	36%	
22665311	Thái Văn Sơn	<ul style="list-style-type: none"> • Data crawler design • Crawl data (marketing) 	14%	
22691511	Đỗ Viết Thanh Bằng	<ul style="list-style-type: none"> • Crawl data (foreign languages) • Merge crawled data (remove duplicates) • Push data to MongoDB 	12%	
22693501	Trần Tuấn Sang	<ul style="list-style-type: none"> • Crawl data (finance and accounting) • Design ERD model 	13%	
22696961	Mai Hoàng Tuấn	<ul style="list-style-type: none"> • Crawl data (health and beauty) • Create tables and insert cleaned data into PostgreSQL 	13%	

Trello URL:

<https://trello.com/invite/b/6707e985a3e2c825354f3699/ATTI6405a2570e6ed98c25ad2b7d87a123e9BB4A1DAA/unicaproject>

Established Date: 09/09/2024

Github URL: https://github.com/3002tad/Unica_Project

Established Date: 09/09/2024

Data Pipeline System Architecture Design

