BỘ GIÁO DỤC VÀ ĐÀO TẠO

**TRƯỜNG ĐẠI HỌC CẦN THƠ**

**KHOA CÔNG NGHỆ THÔNG TIN & TRUYỀN THÔNG**



**BÁO CÁO THỰC TẬP THỰC TẾ**

**CHUYÊN NGÀNH CNTT – (Mã HP CT450)**

**Đề tài**

**TÍCH HỢP, PHÂN TÍCH, CHUYỂN ĐỔI DỮ LIỆU,**

**SỬ DỤNG CÁC DỊCH VỤ CỦA MICROSOFT AZURE**

**Cơ quan thực tập**

**Công ty TNHH Giải Pháp Phần Mềm Tường Minh (TMA Solutions)**

**Giảng viên hướng dẫn Cán bộ hướng dẫn**

*TS. Thái Minh Tuấn Ngô Tuấn Đạt*

**Sinh viên thực hiện**

*Nguyễn Kim Mai Vy – B1809432*

*Khóa: 44*

**Cần thơ, 07/2022**

# LỜI CÁM ƠN

Lời đầu tiên, em xin gửi lời cảm ơn chân thành đến ban lãnh đạo Công ty TNHH Giải Pháp Phần Mềm Tường Minh (TMA Solutions) Chi nhánh 6 (Lab 6) quận 12, TP HCM, anh Huỳnh Thái Trung, anh Lê Nhựt Trường và anh Huỳnh Quang Tín đã tạo cho em một cơ hội được đồng hành cùng công ty cũng như gắn kết em với nhóm, quan tâm, giúp đỡ và đưa ra những lời khuyên hữu ích những lúc em gặp khó khăn.

Em xin chân thành cảm ơn anh Ngô Tuấn Đạt đã tận tâm hướng dẫn, truyền dạy các kỹ năng, kiến thức quý báu để em hoàn thành tốt đề tài được phân công trong quá trình thực tập. Bên cạnh đó, em cũng xin cảm ơn các Anh Chị trong nhóm đã luôn nhiệt tình giúp đỡ, tạo mọi điều kiện tốt nhất giúp em nhanh chóng hoà nhập vào môi trường làm việc chuyên nghiệp, năng động, nhiệt huyết; cũng như được tham gia, học hỏi các kinh nghiệm từ những dự án thực tế, để nâng cao kiến thức, kỹ năng của bản thân.

Em cũng xin chân thành cảm ơn và sự tri ân sâu sắc đến Thầy Thái Minh Tuấn đã hướng dẫn, giải đáp những thắc mắc lúc em gặp khó khăn trong quá trình thực tập thực tế. Em luôn trân trọng và biết ơn những góp ý của thầy trong quá trình thực tập cũng như những kiến thức quý báo thầy đã truyền dạy trong những năm em học Đại học. Đó là những hành trang giúp em không ngừng tiến bộ và có thể hoàn thành được đề tài thực tập thực tế này. Em xin gửi lời cảm ơn đến với Thầy Cô, Gia đình, Bạn bè cũng như Nhà trường đã tạo điều kiện tốt nhất để em có thể hoàn thành khóa thực tập thành công.

Mặc dù, em đã cố gắng nỗ lực để hoàn thành tốt nhất nhưng chắc không thể tránh khỏi những thiếu sót. Em rất mong được nhận những lời nhận xét và góp ý từ Thầy và anh Đạt, qua đó giúp em hoàn thiện đề tài tốt hơn nữa.

Cuối lời, em kính chúc Thầy Thái Minh Tuấn và các Anh Chị ở công ty luôn mạnh khoẻ và thành công hơn nữa trên nhiều lĩnh vực trong cuộc sống. Em xin chân thành cảm ơn!

Cần Thơ, ngày 15 tháng 07 năm 2022

Người viết

**Nguyễn Kim Mai Vy**

# PHIẾU ĐÁNH GIÁ BÁO CÁO KẾT QUẢ THỰC TẬP

M-TT-04

**HỌC KỲ 3 – 2021-2022**

**(Dùng cho giáo viên chấm báo cáo thực tập)**

Họ và tên cán bộ chấm báo cáo: TS. Thái Minh Tuấn

Họ tên sinh viên thực tập: Nguyễn Kim Mai Vy Mã số SV: B1809432

|  |  |  |
| --- | --- | --- |
| **Nội dung đánh giá** | **Điểm tối đa** | **Điểm chấm** |
| **I. Hình thức trình bày** | **1.0** |  |
| I.1 Đúng format của khoa (Trang bìa, trang lời cảm ơn, trang đánh giá thực tập của khoa, trang mục lục và các nội dung báo cáo). Sử dụng đúng mã và font tiếng Việt (Unicode Times New Roman, Size 13) | 0.5 |  |
| I.2 Trình bày mạch lạc, súc tích, không có lỗi chính tả | 0.5 |  |
| **II. Phiếu theo dõi** | **4.75** |  |
| II.1 Có lịch làm việc đầy đủ cho 8 tuần | 0.25 |  |
| II.2 Số buổi thực tập tại cơ quan trong 1 tuần >=6; ít hơn 6 buổi 0.0 điểm | 1.0 |  |
| II.3 Hoàn thành tốt kế hoạch công tác ghi trong lịch làm việc.  Cách tính điểm = (Điểm cộng của cán bộ hướng dẫn/100) x 3.5 | 3.5 |  |
| **III. Nội dung thực tập (quyển báo cáo)** | **4.25** |  |
| III.1 Có được sự hiểu biết tốt về cơ quan nơi thực tập | 0.5 |  |
| III.2 Phương pháp thực hiện phù hợp với nội dung công việc được giao | 1.0 |  |
| III.3 Kết quả củng cố lý thuyết | 0.5 |  |
| III.4 Kết quả rèn luyện kỹ năng thực hành | 0.5 |  |
| III.5 Kinh nghiệm thực tiễn thu nhận được | 0.5 |  |
| III.6 Kết quả công việc có đóng góp cho cơ quan nơi thực tập | 1.25 |  |
| **TỔNG CỘNG** | **10.0** |  |
| Điểm trừ |  |  |
| **Điểm còn lại** |  |  |

Cần Thơ, ngày … tháng 07 năm 2022

GIÁO VIÊN CHẤM BÁO CÁO

*(ký tên)*

# MỤC LỤC

LỜI CÁM ƠN i

PHIẾU ĐÁNH GIÁ BÁO CÁO KẾT QUẢ THỰC TẬP ii

MỤC LỤC iii

PHIẾU GIAO VIỆC iv

DANH MỤC HÌNH v

DANH MỤC BẢNG v

**CHƯƠNG I: GIỚI THIỆU CÔNG TY TMA SOLUTIONS 1**

1.1 Tổ chức vận hành, nhân sự công ty TMA 1

1.1.1 Giới thiệu 1

1.1.2 Sơ đồ tổ chức 2

1.2 Hoạt động chuyên ngành và môi trường làm việc 3

1.2.1 Hoạt động chuyên ngành 3

1.2.2 Môi trường làm việc 5

**CHƯƠNG 2: NỘI DUNG THỰC TẬP 7**

2.1 Kế hoạch công việc được giao 7

2.2 Phương pháp thực hiện 8

2.2.1 Tìm hiểu lý thuyết và công nghệ liên quan 8

2.2.1.1 Scala 8

2.2.1.2 Apache Spark 9

2.2.1.3 Microsoft Azure 10

2.2.1.4 IntelliJ 10

2.2.2 Các bước thực hiện dự án thực tế 11

**CHƯƠNG 3: KẾT QUẢ 12**

3.1 Kiến thức lý thuyết được củng cố 12

3.2 Kinh nghiệm thực hành học hỏi được 12

3.3 Kinh nghiệm thực tiễn được tích lũy 12

3.4 Những đóng góp của bản thân vào dự án 12

# PHIẾU GIAO VIỆC

# DANH MỤC HÌNH

[Hình 1: Sự phát triển nhân lực công ty TMA Solutions (1997-2021) 2](#_Toc108105828)

[Hình 2: Sơ đồ tổ chức công ty TMA Solutions 2](#_Toc108105829)

# DANH MỤC BẢNG

[Bảng 1: Các hoạt động chuyên ngành công ty TMA Solutions 5](#_Toc108105823)

[Bảng 2: Kế hoạch phân công công việc 8](#_Toc108105824)

# CHƯƠNG I: GIỚI THIỆU CÔNG TY TMA SOLUTIONS

## Tổ chức vận hành, nhân sự công ty TMA

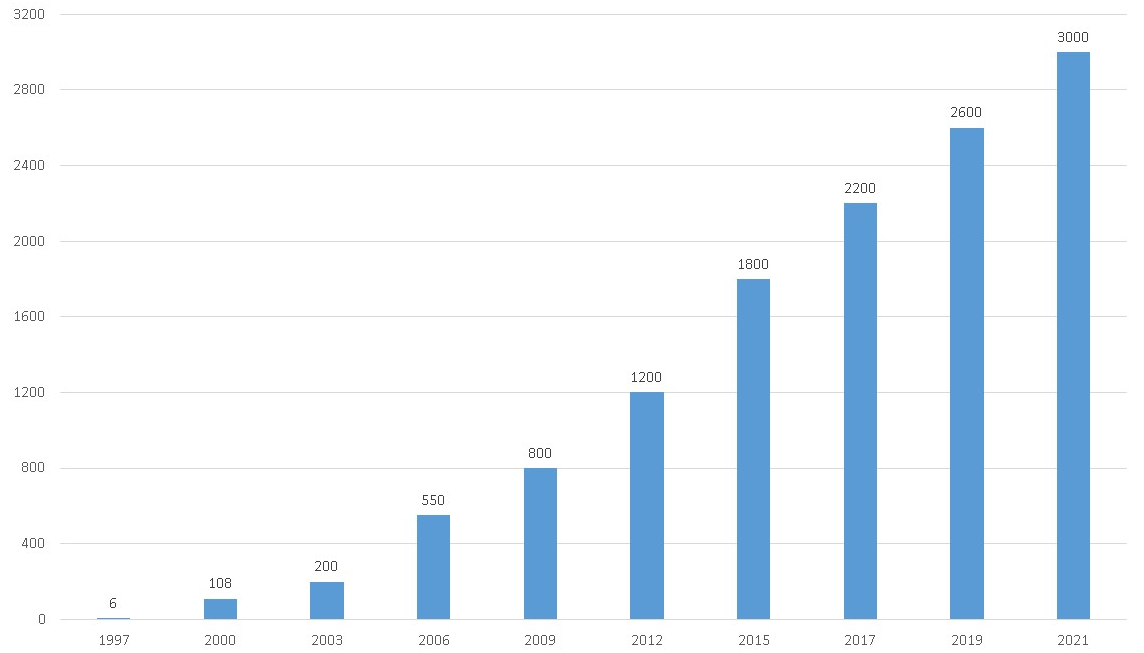
### 1.1.1 Giới thiệu

TMA Solutions được thành lập năm 1997 với sự phát triển vững mạnh trong suốt 24 năm qua, TMA Solutions nhanh chóng trở thành công ty phần mềm hàng đầu Việt Nam. Trải qua 25 năm hoạt động, công ty TMA Solutions nhận được rất nhiều bằng khen, giải thưởng và liên tục được báo chí, truyền thông đưa tin như:

* Bằng khen của Uỷ ban Nhân dân Thành phố Hồ Chí Minh
* Chứng nhận và cúp của Hội Tin học Thành phố Hồ Chí Minh (HCA)
* Huy Chương Vàng Xuất khẩu Phần Mềm (15 năm liền từ 2004-2019)
* Top 10 Đơn vị Xuất khẩu Phần Mềm, FinTech, AI và IoT
* Bằng khen của VINASA
* Là một trong 15 công ty hàng đầu thế giới trong việc áp dụng hiệu quả quy trình gia công phần mềm (Báo cáo của công ty tư vấn Mỹ Aberdeen, 09/2002)
* Đối tác chính thức của Microsoft từ 2007 đến 2021
* Đạt nhiều chứng chỉ quốc tế cao cấp (CMMI-L5, TL 9000, ISO 9001:2000, ISO 27001:2013)

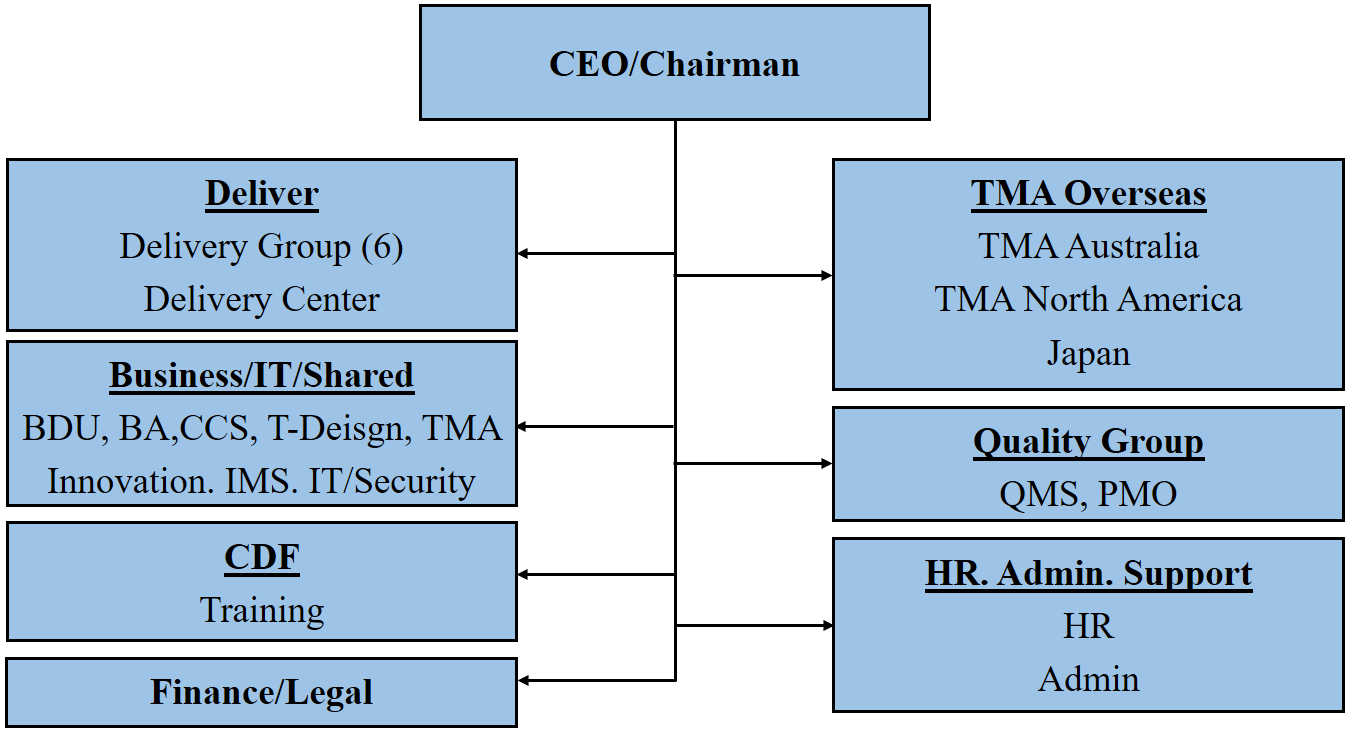
TMA Solutions hiện là đối tác tin cậy có nhiều năm kinh nghiệm thực hiện những dự án lớn và phức tạp với những công nghệ mới nhất với hơn 100 khách hàng đến từ hơn 27 quốc gia trên thế giới. Đến thời điểm hiện tại, TMA Solutions có hơn 2700 kỹ sư tài năng, nhiệt huyết đang làm việc trong một môi trường năng động và chuyên nghiệp và luôn nhận được sự tin cậy, đánh giá cao từ phía khách hàng.

TMA Solutions hiện có 7 chi nhánh tại Việt Nam (6 tại TP. HCM và 1 ở TP. Quy Nhơn) cùng 6 văn phòng ở các nước Mỹ, Canada, Pháp, Nhật, Úc, Singapore.



##### Hình 1: Sự phát triển nhân lực công ty TMA Solutions (1997-2021)

### 1.1.2 Sơ đồ tổ chức



##### Hình 2: Sơ đồ tổ chức công ty TMA Solutions

## Hoạt động chuyên ngành và môi trường làm việc

### 1.2.1 Hoạt động chuyên ngành

|  |  |
| --- | --- |
| Front-end | * HTML5, CSS3, AngularJS, Bootstrap, NodeJS * ReactJS, ReactNative, Redux * LESS, SASS, EnyoJS, Moonstone, NoSQL * Backbone JS, Underscore JS, Require JS, jQuery * Durandal JS, KnockoutJS, Ext JS * Sencha Touch, PhoneGap, Titanium, KendoUI |
| Điện toán đám mây (Cloud) | * Amazon Web Services (AWS) * Salesforce.com * MS Azure, SQL Azure, Azure Mobile, Signal R, Notification Hub, Azure Service Bus * Google Cloud Platform (GCP) * Heroku * VMWare |
| IoT và thiết bị thông minh  (Internets of things) | * Smart Devices * Smart Camera, Drone * IoT Devices Testing * Smart Agriculture * Asset Tracking * Vehicle Tracking |
| Kiểm thử phần mềm | * Test Automation * Performance and Traffic Test * API Test * Security Test * Cloud Test * Big Data Test |
| Trí tuệ nhân tạo (AI) | * Machine Learning * Deep Learning * Reinforcement Learning * Image Recognition * Pattern Recognition NLP |
| Mã nguồn mở | * PHP, Ruby, Python, Perl * Symfony, Laravel, Django * Magento, Drupal |
| Phần mếm nhúng | * FPGA: FIFO, PISO, MUX, Hash, DDR RAM Handling, VHDL & Verilog, Xilinx ISE * OS: VxWorks, Linux/Real-time Linux, Android, ucLinux, ttylinux, gentoo * MCU: ARM, x86, MIPS, PIC, AVR32 * Interface: UBS, Ethernet, LCD, CF/MMC/SD * Firmware & Device Drivers: Keil C, IAR, Altium, Buildroot, Openembedded, Openwrt * Wireless: GSM/GPRS, Zigbee, BLE, Beacon, RFID, WIFI, 3G, MQTT |
| Devops | * CI/CD/DevOps * Jenkins, Docker, Ansible * Git, Maven, Chef, Bamboo |
| Dữ liệu lớn – Phân tích dữ liệu  (Big Data & Analytics) | * Machine Learning * Deep Learning * Reinforcement Learning * Image Recognition * Pattern Recognition NLP |
| Java | * 2EE : EJB, Servlets, JSP, JNDI, RMI/IIOP, JTA/JTS, JMS, JavaMail, JSF, JMS, JCA connectors, JDBC, Hibernate, JPA, Spring, Guice, Struts, Liferay portal * Web Desktop: HTML/ XHTML, CSS, JavaScript, AJAX, J2SE, Swing, GWT, GWT-Platform, Sencha GXT, Eclipse RCP, Macromedia Flex * Application Server: Apache Tomcat, JBoss AS, GlassFish * Web Services & SOA: WSDL, SOAP, REST, XSD, WS-I, UDDI, JAX-RPC, AXIS, XML RPC, WS-Security * Business Process & Workflow: BPMN2, JBoss, jBPM5 * Tools: * Eclipse, Maven, Jenkins, SonarQube, SoapUI, PowerBuilder * Atlassian: JIRA, Confluence, Fisheye, Crucible, Bambo * SVN, Git, VersionOne |
| BlockChain | * Hyperledger * Ethereum * Smart Contract * Stellar * Dapp * Solidity * Golang |
| Ứng dụng điện thoại (Mobile) | * NFC, Bluetooth, Beacon * e-Pub3, pdf, mp3, mp4 * 3D, OpenGL, OGRE 3D * Offline Map * WebGL * HTML5, JQuery, Sencha Touch * Camera * SIP, RTP, SDP * Facebook, Twitter * M2M * Push Notification * Mobile DRM * Xamarin, PhoneGap, Titanium QR Code * VoIP, MjSIP, PjSIP * iCloud, Dropbox, SkyDrive * BYOD, Mobile Security * Google Map, Bing Map * OpenStreet Map * AR * Voice Recognition and Control * SMS * Soft Phone, Call Control * Linux kernel * Smart TV |
| Viễn thông (Telecom Network) | * Enterprise * Communications * Carrier Communications * Wireless Network * Networking EMS, NMS * OSS/BSS |

###### Bảng 1: Các hoạt động chuyên ngành công ty TMA Solutions

### 1.2.2 Môi trường làm việc

Là doanh nghiệp CNTT có 24 năm kinh nghiệm, làm việc với rất nhiều khách hàng hàng đầu thế giới, TMA tự hào đạt các chuẩn Quốc tế cao cấp như CMMI-L5, TL 9000, ISO 9001:2000, ISO 27001:2013.

Để có được những thành công trên, TMA chủ trương tuyển dụng những lập trình viên tài năng, có khả năng giao tiếp bằng tiếng Anh và có kỹ năng tốt. Ứng viên sau khi gia nhập TMA sẽ được hướng dẫn cụ thể về công ty, các quy định, vị trí công việc. Theo đó, ngay trong quá trình thử việc, nhân viên sẽ được tham gia các khóa đào tạo về kỹ năng làm việc, giao tiếp để bổ sung đầy đủ các yêu cầu cần thiết khi làm việc tại TMA.

Trong công việc, TMA khuyến khích nhân viên bày tỏ quan điểm của mình để đóng góp cho sự phát triển bền vững của tập thể. Bên cạnh đó công ty luôn ghi nhận thành tích của các cá nhân và có chính sách khen thưởng kịp thời.

Tại TMA, môi trường làm việc tiện nghi, thoải mái không chỉ nằm ở cơ sở vật chất mà còn ở các chính sách hỗ trợ về tài chính và và chăm sóc sức khỏe cho nhân viên và gia đình. Tiêu biểu là chương trình bảo hiểm chăm sóc sức khỏe toàn diện, chính sách cho vay tiền không lãi suất, chính sách hỗ trợ tài chính khẩn cấp cho các trường hợp khó khăn hoặc mắc bệnh hiểm nghèo…

# CHƯƠNG 2: NỘI DUNG THỰC TẬP

## 2.1 Kế hoạch công việc được giao

|  |  |  |  |
| --- | --- | --- | --- |
| **Tuần** | **Thời gian** | **Nội dung**  **công việc được giao** | **Kết quả** |
| 1 | 23/05/2022  29/05/2022  (10 buổi) | Tìm hiểu ngôn ngữ lập trình Scala  Tìm hiểu Spark apache | Hoàn thành |
| 2 | 30/05/2022  05/06/2022  (10 buổi) | Thực hành số 1:  + Sử dụng Spark apache phân tích và chuyển đổi dữ liệu với ngôn ngữ Scala. | Hoàn thành |
| 3 | 06/06/2022  12/06/2022  (10 buổi) | Tìm hiểu một số dịch vụ của Azure cho việc phân tích dữ liệu:  + Azure Storage Account (Az Blob Storage, Delta Lake Storage Gen2, Az Key Vault). | Hoàn thành |
| 4 | 13/06/2022  19/06/2022  (10 buổi) | Tìm hiểu một số dịch vụ của Azure cho việc phân tích dữ liệu:  + Azure Data Factory(Azure Data Pipeline, Azure Data Pipeline Trigger) | Hoàn thành |
| 5 | 20/06/2022  26/06/2022  (10 buổi) | Tìm hiểu một số dịch vụ của Azure cho việc phân tích dữ liệu  + Azure Databrick (Workspace, Notebook, Secret scope) | Hoàn thành |
| 6 | 27/06/2022  03/07/2022  (10 buổi) | Thực hành số 2:  + Tạo một *Azure Storage Account* thuộc *Data Lake Storage gen2*, sau đó upload .parquet file vào (container/lv1/).  + Sử dụng *Data factory*, tạo một pipeline di chuyển dữ liệu từ *Azure Storage Account* ở trên tới một *Azure Storage Account* khác. | Hoàn thành |
| 7 | 04/07/2022  10/07/2022  (10 buổi) | Thực hành số 3:  + Trong *Data brick*, sử dụng *Spark* với ngôn ngữ Scala để di chuyển dữ liệu từ /lv1/\*.parquet to /lv2.  (Sử dụng Az Key Vault để bảo mật thông tin) | Hoàn thành |
| **Tuần** | **Thời gian** | **Nội dung**  **công việc được giao** | **Kết quả** |
| 8 | 11/07/2022  17/07/2022  (10 buổi) | Thực hành số 4:  + Xây dựng .jar file, sử dụng ngôn ngữ Scala, thêm vào Az Pipeline để thay đổi và di chuyển dữ liệu từ /lv1/\*.parquet to /lv2 | Hoàn thành |

###### Bảng 2: Kế hoạch phân công công việc

## 2.2 Phương pháp thực hiện

### 2.2.1 Tìm hiểu lý thuyết và công nghệ liên quan

#### 2.2.1.1 Scala

Ngôn ngữ Scala (viết tắt cho scalable language) là một ngôn ngữ được thiết kế để có khả năng mở rộng tùy theo nhu cầu của người sử dụng. Ngôn ngữ Scala được sử dụng trong nhiều lĩnh vực, quy mô, từ những đoạn script nhỏ cho đến những hệ thống lớn.

Scala chạy trên nền máy ảo Java và nó tương thích hoàn toàn với Java. Ngoài ra nó cũng có thể chạy trên .NET, tuy nhiên chưa ổn định.

Scala được xây dựng với các tính năng của lập trình hàm và lập trình hướng đối tượng. Sự kết hợp của lập trình hàm và lập trình hướng đối tượng được thể hiện trong nhiều khía cạnh, hai kiểu lâp trình này bổ sung sức mạnh cho nhau đặc biệt khi hệ thống được mở rộng. Lập trình hàm có thể xây dựng nhanh chóng và dễ dàng những tính năng từ những phần nhỏ trong khi lập trình hướng đối tượng thích hợp để cấu trúc một hệ thống lớn.

Đặc điểm của Scala:

* Cơ động linh hoạt về cú pháp
* Sở hữu vòng lặp for mạnh mẽ
* Được thiết kế để phù hợp với cả lập trình hàm cũng như lập trình hướng đối tượng
* Hệ thống kiểu dữ liệu phong phú và có khả năng mở rộng (định nghĩa thêm method cho kiểu có sẵn)

Ưu điểm của Scala:

* Giúp các lập trình viên dễ dàng phát triển, [xây dựng web](https://wiki.tino.org/tai-lieu-hoc-thiet-ke-web-cho-nguoi-moi-bat-dau/)
* Ngôn ngữ lập trình Scala ngắn gọn, súc tích nên người dùng dễ dàng đọc và hiểu được ngôn ngữ
* Giảm thiểu kích thước của lập trình
* Tránh được những sai sót, trục trặc có thể xảy ra tại thời điểm biên dịch
* Scala có thể nâng cấp, mở rộng cao và cung cấp cho các lập trình viên một loạt các ngôn ngữ máy độc đáo, dễ dàng thêm chúng vào thư viện. Nhờ vậy, hoạt động lập trình của bạn nhanh chóng và an toàn khi dựa vào thư viện hàm có sẵn.
* Hỗ trợ hầu hết các đặc tính OOP, FP, Highly Functional.
* Scala giúp cho quá trình lập trình diễn ra song song và tốt hơn, hiệu suất cao, cú pháp linh hoạt.

Hạn chế của Scala:

* Khả năng tương thích ngược còn kém: Khi bạn cập nhật lên phiên bản mới nhất của Scala, bạn nên cẩn thận trong việc thay đổi tên cho các Package, Class, Method, Function. Bởi vì, ngôn ngữ Scala sau mỗi lần cập nhật Version mới sẽ khó nhận diện các yếu tố thuộc Version cũ, bắt buộc bạn phải đổi tên phù hợp.

#### 2.2.1.2 Apache Spark

Apache Spark là một framework mã nguồn mở tính toán cụm, được phát triển sơ khởi vào năm 2009 bởi AMPLab. Sau này, Spark đã được trao cho Apache Software Foundation vào năm 2013 và được phát triển cho đến nay.

Tốc độ xử lý của Spark có được do việc tính toán được thực hiện cùng lúc trên nhiều máy khác nhau. Đồng thời việc tính toán được thực hiện ở bộ nhớ trong (in-memories) hay thực hiện hoàn toàn trên RAM.

Ưu điểm của Apache Spark:

* Đơn giản và dễ sử dụng
* Khả năng, tốc độ phân tích thời gian thực ấn tượng
* Được hỗ trợ bởi các thư viện cấp cao
* Khả năng tương thích cao và hỗ trợ nhiều ngôn ngữ lập trình

Hạn chế của Apache Spark:

* Tiềm ẩn nguy cơ mất quyền kiểm soát tài khoản cho Microsoft
* Tăng sự cạnh tranh giữa các nguồn lực phát triển
* Chạy theo kinh doanh số lượng
* Sự suy giảm trong lợi nhuận trước mắt và doanh thu
* Tăng tốc toàn cầu khóa và cạnh tranh thị trường
* Tăng chi phí lưu trữ và tích hợp

#### 2.2.1.3 Microsoft Azure

Microsoft Azure là nền tảng tính toán đám mây được xây dựng bởi Microsoft dành cho xây dựng, kiểm thử, triển khai và quản lý các ứng dụng và dịch vụ thông qua mạng lưới trung tâm dữ liệu được quản lý bởi Microsoft. Nó cung cấp các phần mềm, nền tảng, và hệ thống cơ sở hạ tầng như các dịch vụ hỗ trợ nhiều ngôn ngữ lập trình, framework, công cụ khác nhau.

Ưu điểm của Microsoft Azure:

* Thời gian triển khai nhanh hơn và dễ sử dụng cho khách hàng
* Nguồn lực phát triển dồi dào
* Nguồn thu nhập định kỳ
* Cải thiện TCO và margin để phân biệt các IP
* Cơ hội để phân tích cơ sở khách hàng của Microsoft
* Hạ thấp chi phí vòng đời ứng dụng

Hạn chế của Microsoft Azure:

* Tiềm ẩn nguy cơ mất quyền kiểm soát tài khoản cho Microsoft
* Tăng sự cạnh tranh giữa các nguồn lực phát triển
* Chạy theo kinh doanh số lượng
* Sự suy giảm trong lợi nhuận trước mắt và doanh thu
* Tăng tốc toàn cầu khóa và cạnh tranh thị trường
* Tăng chi phí lưu trữ và tích hợp

#### 2.2.1.4 IntelliJ

Phần mềm IntelliJ IDEA, sản phẩm nổi tiếng của JetBrains đã nhận được rất nhiều giải thưởng. Phần mềm được thiết kế để cải tiến năng suất cho các nhà phát triển. IntelliJ IDEA cung cấp trình soạn thảo thông minh, trình phân tích mã và tập hợp mạnh mẽ của refactorings hỗ trợ một loạt các ngôn ngữ lập trình, các khuôn khổ và công nghệ, và đã sẵn sàng để sử dụng.

Ưu điểm của IntelliJ:

* Cho phép các nhà phát triển tập trung phát triển và quản lý tất cả các tác vụ thông thường.
* Cho phép viết, gỡ lỗi, tái cấu trúc, kiểm tra và tìm hiểu mã của bạn mà không bị ảnh hưởng.
* Xử lý liền mạch cơ sở mã hỗn hợp của Java, Ruby, Groovy, Python và Scala.
* Tự động duy trì chất lượng mã.
* Theo dõi và sửa lỗi trên tất cả các cấp độ – từ các câu đến kiến trúc tổng thể.
* Tạo mã “sạch”, nhanh chóng thực hiện mã trong thời gian ngắn nhất.
* Được thiết kế để làm việc trên tất cả các quy mô – từ cá nhân đến doanh nghiệp.
* Hỗ trợ tất cả các ngôn ngữ, công nghệ và framework chính.
* Làm việc với các hệ thống điều khiển phiên bản phổ biến và TeamCity, server tích hợp liên tục.

### 2.2.2 Các bước thực hiện dự án thực tế

* Đọc tài liệu tìm hiểu về Scala và Apache Spark
* Thực hiện cài đặt và thực hiện bài tập về Scala dòng lệnh bằng Command Prompt
* Sử dụng IntelliJ để lập trình Scala dưới dạng giao diện người dùng; tiến hàng đọc file, xử lý dữ liệu bằng Scala
* Tải tệp “*.parquet*” lên Azute Storage Account thuộc Data Lake Storage gen2
* Sử dụng Data factory, tạo một pipeline di chuyển dữ liệu từ Azure Storage Account ở trên tới một Azure Storage Account khác
* Trong Data brick, sử dụng Spark với ngôn ngữ Scala để di chuyển dữ liệu từ /lv1/\*.parquet to /lv2
* Sử dụng Az Key Vault để bảo mật thông tin trên
* Xây dựng .jar file, sử dụng ngôn ngữ Scala, thêm vào Az Pipeline để thay đổi và di chuyển dữ liệu từ /lv1/\*.parquet to /lv2

# CHƯƠNG 3: KẾT QUẢ

## 3.1 Kiến thức lý thuyết được củng cố

Lập trình hướng đối tượng áp dụng vào ngôn ngữ Scala

## 3.2 Kinh nghiệm thực hành học hỏi được

* Lập trình Scala dựa trên Java
* Sử dụng Apache Spark
* Sử dụng công cụ lập trình IntelliJ
* Sử dụng bộ nhớ của đám mây khổng lồ của Microsoft Azure

## 3.3 Kinh nghiệm thực tiễn được tích lũy

* Làm quen với môi trường làm việc thực tế trong công ty
* Nâng cao khả năng làm việc nhóm
* Tác phong làm việc được cải thiện, khả năng quản lý và sắp xếp thời gian tốt hơn
* Nâng cao tinh thần, trách nhiệm trong công việc
* Khả năng tìm toài, học hỏi công nghệ mới
* Phát triển kỹ năng lập trình.

## 3.4 Những đóng góp của bản thân vào dự án

* Những việc làm được:
* Hoà nhập được môi trường làm việc, kỹ năng giao tiếp và tham gia các khoá đào tạo tại công ty.
* Cải thiện kiến thức về Java và sử dụng các công cụ hỗ trợ để phát triển một số tính năng trong dự án.
* Làm quen với các thành phần trong dự án.
* Thực hiện các bài tập để hiểu rõ về dự án.
* Kết quả công việc đóng góp cho cơ quan nơi thực tập:
* Thông qua quá trình đào tạo tạo tiền đề để tham gia phát triển dự án thực tế.
* Tham gia theo dõi, bảo trì hệ thống mạng.
* Góp phần xây dựng một vài tính năng đơn giản trong hệ thống.
* Những việc chưa làm được
* Do công nghệ và phương pháp sử dụng tương đối mới, dự án có tính quy mô lớn và thời gian thực tập tương đối ngắn nên chưa thể nắm rõ các thành phần trong dự án.