****

**HỌC VIỆN NGÂN HÀNG**

**KHOA HỆ THỐNG THÔNG TIN QUẢN LÝ**

****

**BÁO CÁO BÀI TẬP LỚN**

**MÔN KHO DỮ LIỆU VÀ KINH DOANH THÔNG MINH**

**ĐỀ TÀI: ỨNG DỤNG DATA WAREHOUSE VÀ BI TRONG VIỆC TỐI**

**ƯU HOÁ CHIẾN LƯỢC KINH DOANH CỦA NETFLIX**

|  |  |
| --- | --- |
| **Giảng viên hướng dẫn** | : T.S Đinh Trọng Hiếu |
| **Lớp** | : 241IS24A03 |
| **Nhóm** | : 12 |

**Hà Nội – 2024**

****

**HỌC VIỆN NGÂN HÀNG**

**KHOA HỆ THỐNG THÔNG TIN QUẢN LÝ**

****

**BÁO CÁO BÀI TẬP LỚN**

**MÔN KHO DỮ LIỆU VÀ KINH DOANH THÔNG MINH**

**ĐỀ TÀI: ỨNG DỤNG DATA WAREHOUSE VÀ BI TRONG VIỆC TỐI**

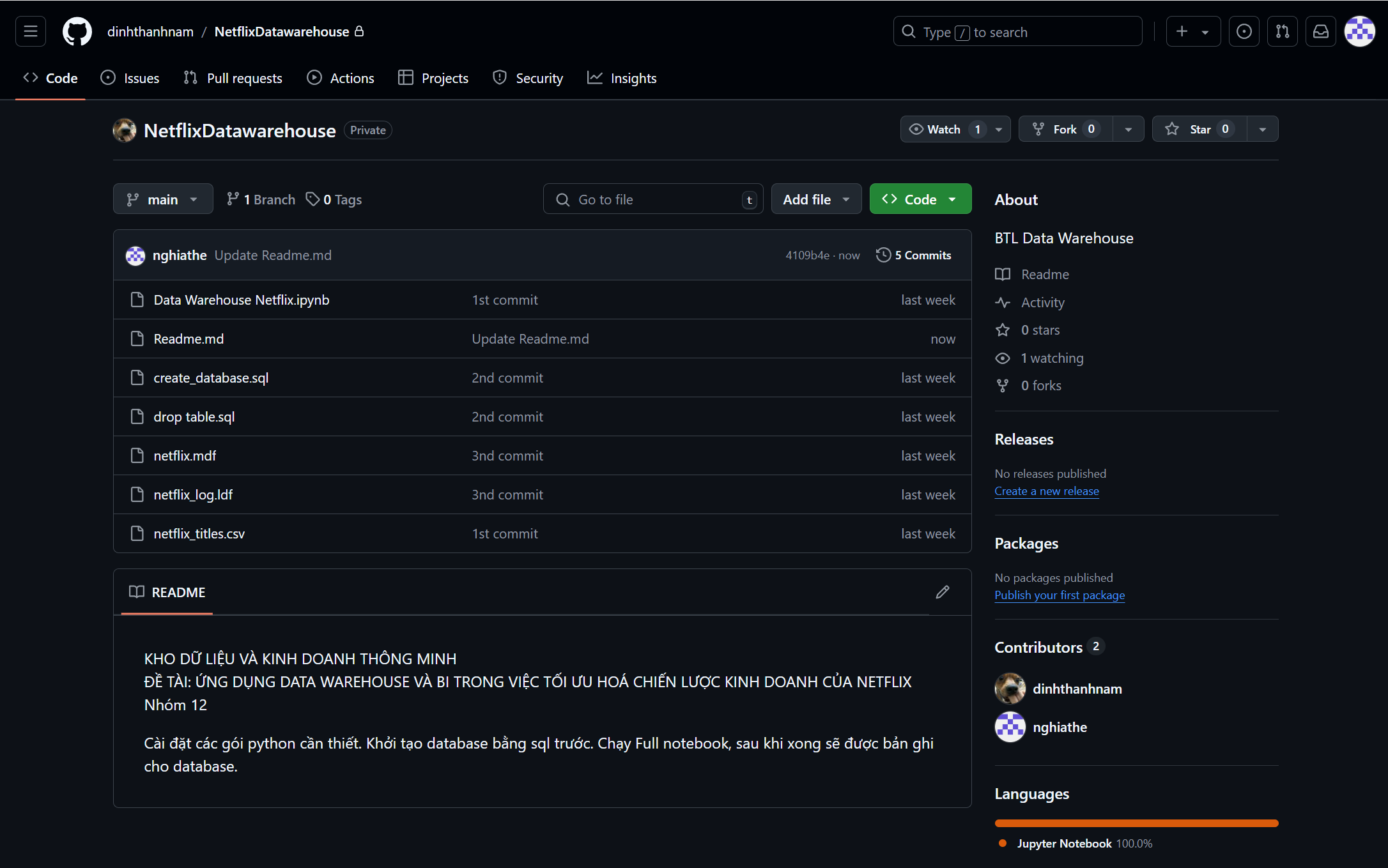
**ƯU HOÁ CHIẾN LƯỢC KINH DOANH CỦA NETFLIX**

|  |  |  |  |
| --- | --- | --- | --- |
| **Thành viên** | **Mã sinh viên** | **Tỉ lệ đóng góp** | **Xác nhận** |
| Đinh Thành Nam  (Nhóm trưởng) | 24A4042599 | 20% |  |
| Nguyễn Duy Hưng | 24A4042596 | 21.5% |  |
| Nguyễn Thế Nghĩa | 24A4042602 | 21.5% |  |
| Trần Quốc Việt | 24A4042618 | 18.5% |  |
| Trương Nhật Dũng | 24A4040013 | 18.5% |  |

**Hà Nội – 2024**

**HOẠT ĐỘNG NHÓM**

**1. Theo dõi tiến độ trên Github**

****

**2. Hoạt động nhóm**

Nhóm hoạt động thường xuyên qua kênh discord để trao đổi và phát triển bài tập:

# LỜI CẢM ƠN

Lời đầu tiên, nhóm 12 xin gửi lời cảm ơn đến Học viện Ngân hàng và khoa Công nghệ thông tin và kinh tế số đã tạo điều kiện cho chúng em và các bạn sinh viên có cơ hội được học tập trong một môi trường năng động, sáng tạo, tạo điều kiện nỗ lực và phát huy hết tiềm năng của mình. Nhóm xin gửi lời cảm ơn đặc biệt đến thầy Đinh Trọng Hiếu đã dành sự tâm huyết và nhiệt tình giảng dạy, giúp chúng em học được một môn học có ứng dụng cao trong ngành học và nghiệp vụ sau này, đáp ứng nhu cầu thực tiễn của các nhà tuyển dụng về kiến thức, kỹ năng cần có.

Trong học phần này, chúng em được tiếp xúc với các công cụ tạo dashboard, báo cáo, học về cơ bản của thiết kế kho dữ liệu, và được thực hiện một dự án thực tế thiết kế dashboard cho doanh nghiệp cụ thể. Đây là một trải nghiệm thú vị và là một kĩ năng quan trọng cho chúng em sau này.

Chúng em xin chân thành cảm ơn!

Nhóm sinh viên thực hiện

Nhóm 12

# LỜI CAM ĐOAN

Nhóm 12 xin cam đoan báo cáo này là sản phẩm nghiên cứu và thực hành của nhóm. Bài báo cáo đảm bảo tính liêm chính trong học tập, không đạo văn, gian lận, bịa đặt. Các thông tin tham khảo được trích dẫn nguồn đầy đủ và minh bạch.

Nhóm 12 xin chịu toàn bộ trách nhiệm nếu bài báo cáo vi phạm các điều trên.

Hà Nội, ngày 14 tháng 12 năm 2024

Đại diện nhóm 12

# MỤC LỤC

# DANH MỤC BẢNG, HÌNH VẼ

# LỜI MỞ ĐẦU

Trong thời đại cách mạng công nghệ 4.0, việc quản lý và khai thác dữ liệu trở thành một yếu tố quan trọng đối với sự phát triển bền vững của các doanh nghiệp. Netflix, một trong những công ty hàng đầu trong ngành công nghiệp giải trí trực tuyến, đã chứng tỏ sức mạnh của dữ liệu trong việc tối ưu hóa hoạt động kinh doanh và nâng cao trải nghiệm người dùng.

Dựa trên nhu cầu không ngừng tăng về việc phân tích dữ liệu, việc xây dựng một kho dữ liệu toàn diện, hiệu quả để quản lý, lưu trữ và truy xuất dữ liệu là vô cùng cần thiết. Kho dữ liệu này không chỉ giúp Netflix tối ưu hóa quá trình khuyến nghị phim cho người dùng, mà còn hỗ trợ việc phân tích hành vi người tiêu dùng, dự đoán xu hướng thị trường và ra quyết định chiến lược.

Trong bài tập lớn này, chúng em sẽ tập trung vào việc thiết kế và xây dựng một hệ thống kho dữ liệu cho Netflix. Hệ thống này sẽ được tối ưu hóa cho các mục tiêu như cải thiện hiệu quả truy xuất thông tin, hỗ trợ phân tích kinh doanh và cung cấp các công cụ báo cáo thông minh. Qua đó, bài tập không chỉ giúp người học củng cố kiến thức về kho dữ liệu, mà còn mang lại cái nhìn thực tế về ứng dụng công nghệ này trong ngành giải trí số.

# CHƯƠNG 1. TỔNG QUAN

## 1.1. Tổng quan về đảm bảo chất lượng phần mềm

### 1.1.1. Định nghĩa chất lượng phần mềm

- Định nghĩa theo IEEE (1991):

+ Định nghĩa 1: Chất lượng phần mềm là một mức độ mà một hệ thống, thành phần hệ thống hay tiến trình đáp ứng được yêu cầu đã được đặc tả.

+ Định nghĩa 2: Chất lượng phần mềm là mức độ mà một hệ thống, thành phần hệ thống hay tiến trình đáp ứng được yêu cầu và sự mong đợi của khách hàng hay người sử dụng.

- Định nghĩa theo Pressman: Chất lượng phần mềm là sự phù hợp của các yêu cầu cụ thể về hiệu năng và chức năng, các tiêu chuẩn phát triển phần mềm được ghi lại rõ ràng bằng tài liệu với các đặc tính ngầm định của tất cả các phần mềm được phát triển chuyên nghiệp.

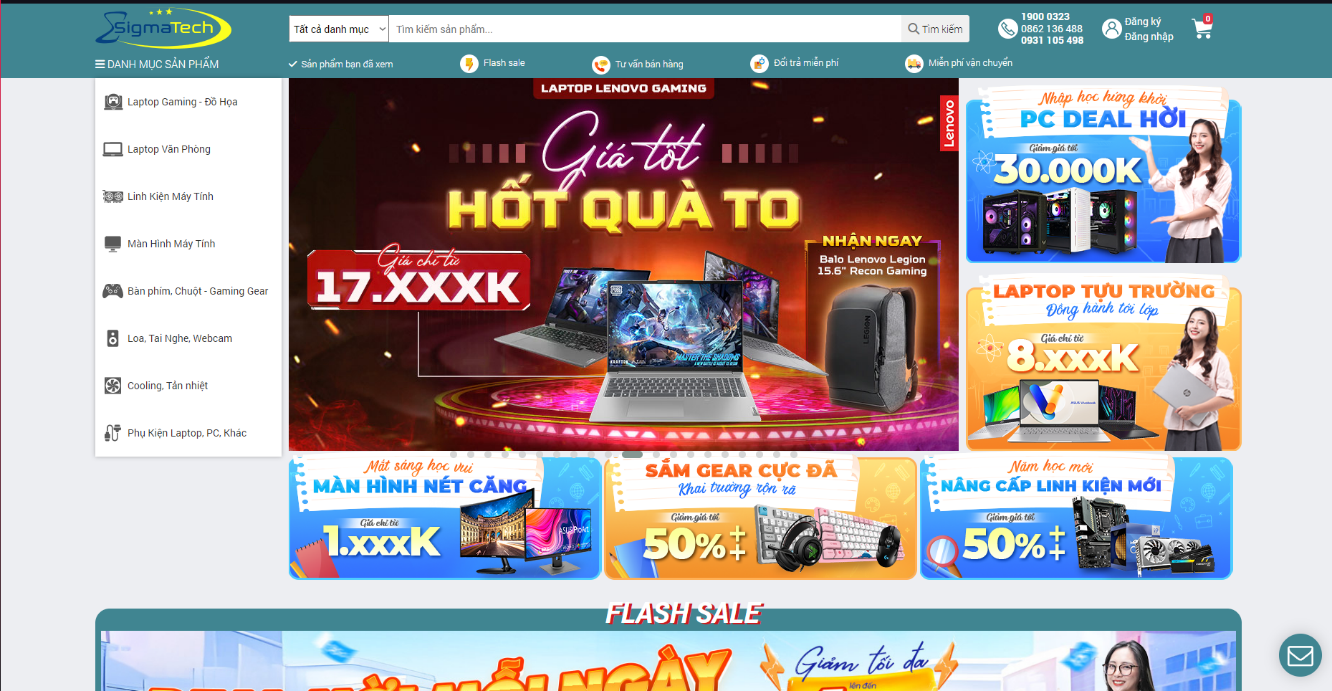
### 1.1.2. Định nghĩa về đảm bảo chất lượng phần mềm

Định nghĩa theo Daniel Galin: Đảm bảo chất lượng phần mềm là một tập hợp các hành động cần thiết được lên kế hoạch một cách hệ thống để cung cấp đầy đủ niềm tin rằng quá trình phát triển phần mềm phù hợp để thành lập các yêu cầu chức năng kỹ thuật cũng như các yêu cầu quản lý theo lịch trình và hoạt động trong giới hạn ngân sách.

## 1.2. Tổng quan về đề tài

### 1.2.1. Phát biểu bài toán

Trong thời đại công nghệ số, các website bán hàng điện tử đóng vai trò quan trọng trong việc cung cấp sản phẩm và dịch vụ đến người tiêu dùng một cách nhanh chóng và tiện lợi. SigmaTech, một website chuyên bán linh kiện điện tử phục vụ nhu cầu chơi game trên máy tính và laptop do chính nhóm tự thiết kế, cần đảm bảo rằng hệ thống của mình hoạt động ổn định, chính xác và đáp ứng tốt các yêu cầu của người dùng.



Ảnh 1: Website bán hàng điện tử SigmaTech

Vấn đề chính đặt ra là làm thế nào để phát hiện và khắc phục sớm các lỗi tiềm ẩn trong hệ thống, từ đó nâng cao trải nghiệm người dùng. Hệ thống SigmaTech bao gồm nhiều tính năng phức tạp như tìm kiếm và lọc sản phẩm, quản lý giỏ hàng, xử lý đơn hàng, thanh toán trực tuyến, quản lý tài khoản người dùng… Một lỗi nhỏ trong bất kỳ tính năng nào cũng có thể gây mất uy tín, ảnh hưởng đến doanh thu và lòng tin của khách hàng.

Do đó, mục tiêu của đề tài này là thực hiện kiểm thử và đảm bảo chất lượng cho website SigmaTech, tập trung vào việc phát hiện và sửa lỗi trong các chức năng cốt lõi. Thông qua việc áp dụng các chiến lược kiểm thử phù hợp, đề tài hướng đến việc đảm bảo hệ thống hoạt động chính xác, an toàn và thân thiện với người dùng, đồng thời đáp ứng được các tiêu chuẩn về chất lượng phần mềm. Kết quả cuối cùng là một báo cáo chi tiết về các lỗi được phát hiện, giải pháp khắc phục và đánh giá hiệu quả của quá trình kiểm thử.

### 1.2.2. Lý do chọn đề tài

Chúng tôi chọn đề tài này vì nhận thấy kiểm thử và đảm bảo chất lượng phần mềm là yếu tố then chốt trong việc phát triển và duy trì uy tín của các website bán hàng điện tử. Một trang web với tính năng không ổn định hoặc thiếu bảo mật có thể gây mất niềm tin nơi người dùng, ảnh hưởng trực tiếp đến doanh thu và hình ảnh thương hiệu.

Ngoài ra, lĩnh vực kiểm thử phần mềm đòi hỏi sự chính xác, sáng tạo và kỹ năng giải quyết vấn đề, rất phù hợp để áp dụng kiến thức học tập vào thực tiễn. Bằng cách thực hiện kiểm thử trên SigmaTech, chúng tôi không chỉ giải quyết bài toán kỹ thuật thực tế mà còn góp phần nâng cao trải nghiệm người dùng trong lĩnh vực thương mại điện tử.

* Đề tài này là cơ hội để nhóm nghiên cứu rèn luyện và ứng dụng các kỹ thuật kiểm thử hiện đại, đồng thời khám phá tầm quan trọng của chất lượng phần mềm đối với sự thành công của doanh nghiệp.

# CHƯƠNG 2.

## 2.1. Xây dựng kế hoạch đảm bảo chất lượng phần mềm

### 2.1.1. Kiểm soát tiến độ hoàn thành dự án

Để đảm bảo chất lượng phần mềm, tiến độ của từng giai đoạn trong dự án cần được kiểm soát chặt chẽ. Các hoạt động sẽ được chia nhỏ thành từng mốc cụ thể:

* Giai đoạn phân tích yêu cầu: Thu thập và xác định đầy đủ các yêu cầu chức năng và phi chức năng.
* Giai đoạn thiết kế: Đảm bảo tài liệu thiết kế đáp ứng yêu cầu và khả thi về mặt kỹ thuật.
* Giai đoạn phát triển: Kiểm tra chất lượng mã nguồn trong từng vòng lặp phát triển.
* Giai đoạn kiểm thử: Thực hiện các loại kiểm thử như kiểm thử chức năng, hiệu năng, bảo mật, ...

### 2.1.2. Kiểm soát tiến độ nội bộ và tác nhân tham gia bên ngoài

Dự án có sự tham gia của nhiều thành phần như nhóm phát triển nội bộ, nhóm kiểm thử và các bên liên quan từ bên ngoài. Việc phối hợp và quản lý tiến độ giữa các bên cần được thực hiện đồng bộ thông qua:

* Lịch họp định kỳ để cập nhật tiến độ và xử lý vấn đề phát sinh.
* Các báo cáo công việc hàng tuần và đánh giá hiệu quả công việc.
* Xác định trách nhiệm cụ thể cho từng thành viên và đội nhóm tham gia.

### 2.1.3. Kiểm soát lịch trình và tài nguyên dự án

Đảm bảo việc phân bổ thời gian, nhân lực và các tài nguyên khác hợp lý trong suốt quá trình thực hiện dự án:

* Ước tính thời gian và tài nguyên cần thiết cho từng hạng mục.
* Theo dõi việc sử dụng tài nguyên thông qua các công cụ quản lý.
* Đề xuất biện pháp điều chỉnh khi có sự chậm trễ hoặc thiếu hụt tài nguyên.

### 2.1.4. Đánh giá chất lượng của sản phẩm phần mềm

Đánh giá chất lượng sản phẩm sẽ dựa trên 4 các tiêu chí sau: chất lượng chức năng, hiệu suất, bảo mật và khả năng sử dụng.

## 2.2. Kiểm thử chức năng

## 2.3. Kiểm thử phi chức năng

### 2.3.1. Load Testing

- Kiểm thử tải (Load Testing): Kiểm tra khả năng của hệ thống khi có nhiều người dùng truy cập đồng thời.

- Phương pháp: Giả lập số lượng lớn người dùng hoặc các yêu cầu truy cập đồng thời.

- Kết quả mong đợi: Hệ thống đáp ứng tốt với x người dùng đồng thời.

### 2.3.2. Stress Testing

- Kiểm thử áp lực (Stress Testing): Đánh giá khả năng của hệ thống khi chịu tải cao hơn mức tối đa có thể chấp nhận được.

- Phương pháp: Tăng dần số lượng người dùng hoặc yêu cầu cho đến khi hệ thống gặp lỗi hoặc không thể đáp ứng.

- Kết quả mong đợi:

* Xác định "điểm vỡ" của hệ thống.
* Đánh giá khả năng hệ thống phục hồi sau khi gặp tải cao.

### 2.3.3. Usability Testing

- Kiểm thử khả năng sử dụng (Usability Testing): Đánh giá mức độ dễ sử dụng của hệ thống từ quan điểm của người dùng.

- Phương pháp: Thu thập phản hồi của người dùng về tính dễ hiểu, tiện lợi và thẩm mỹ của giao diện.

- Kết quả mong đợi: Giao diện dễ sử dụng, chức năng hoạt động mượt mà, người dùng không gặp vấn đề trong quá trình trải nghiệm.

### 2.3.4. Security Testing

- Kiểm thử bảo mật (Security Testing): Đảm bảo rằng hệ thống an toàn và bảo vệ các dữ liệu của người dùng.

- Phương pháp:

* Kiểm tra xác thực người dùng, bảo mật mật khẩu và mã hóa dữ liệu.
* Kiểm tra các lỗ hổng phổ biến như SQL Injection, XSS.
* Kiểm tra quyền truy cập của các vai trò người dùng khác nhau (Admin, Client).

- Kết quả mong đợi: Phát hiện các lỗ hổng và khắc phục trước khi triển khai website.

### 2.3.5. Portability Testing

- Kiểm thử tính di động (Portability Testing): Đánh giá khả năng của hệ thống để di chuyển, triển khai và hoạt động trên nhiều môi trường khác nhau.

- Phương pháp:

* Kiểm thử phần mềm trên các trình duyệt (Chrome, Firefox, Edge, Safari) và hệ điều hành khác nhau (Windows, macOS, Android, iOS).
* Kiểm tra khả năng thích ứng của giao diện trên các độ phân giải màn hình khác nhau (Desktop, Tablet, Mobile).

- Kết quả mong đợi: Phần mềm hoạt động ổn định và không gặp lỗi giao diện hoặc tính năng trên các nền tảng khác nhau.

## 2.4. Đánh giá chất lượng phần mềm

### 2.4.1. Tính đúng

- Phần mềm được coi là đúng khi thực hiện chính xác các chức năng đã được đặc tả và đáp ứng nhu cầu của khách hàng.

- Ví dụ: Hệ thống quản lý dữ liệu không thể xử lý khi file cơ sở dữ liệu rỗng hoặc vượt quá giới hạn số bản ghi là các hệ thống vi phạm tính đúng.

### 2.4.2. Tính kiểm thử được

- Phần mềm có tính kiểm thử được nếu có thể dễ dàng thiết lập các kịch bản kiểm thử và xác minh các chức năng hoạt động chính xác.

- Ví dụ: Hệ thống đặt hàng trực tuyến cho phép kiểm thử tính năng "thêm sản phẩm vào giỏ hàng" bằng cách nhập số lượng, chọn sản phẩm và xác nhận xem giỏ hàng có cập nhật đúng hay không.

### 2.4.3. Tính an toàn

- Phần mềm cần đảm bảo dữ liệu và chức năng được bảo mật, mã hóa để tránh bị sao chép, sửa đổi hoặc phá hoại.

- Ví dụ: Cơ chế xác thực đa lớp (multi-factor authentication) để ngăn truy cập trái phép.

### 2.4.4. Tính toàn vẹn

- Phần mềm đảm bảo tính toàn vẹn khi có khả năng ngăn chặn xâm nhập trái phép và có thể phục hồi dữ liệu khi xảy ra sự cố.

- Ví dụ: Hệ thống tự động sao lưu dữ liệu và phục hồi khi gặp sự cố mất điện.

### 2.4.5. Tính tiêu chuẩn và chuẩn hóa

- Phần mềm cần tuân thủ các tiêu chuẩn kỹ thuật được công nhận và hỗ trợ chuyển đổi dữ liệu giữa các hệ thống khác nhau.

- Ví dụ: Phần mềm xuất dữ liệu theo định dạng CSV để tương thích với các công cụ khác.

### 2.4.6. Tính độc lập

- Phần mềm nên hoạt động độc lập với thiết bị, cấu trúc dữ liệu và nội dung đầu vào để tăng tính linh hoạt.

- Ví dụ: Phần mềm quản lý bán hàng có thể chạy trên nhiều hệ điều hành như Windows, Linux và macOS.

### 2.4.7. Tuân thủ các tiêu chuẩn kỹ thuật

Việc áp dụng hoặc không áp dụng các tiêu chuẩn kỹ thuật không quyết định hoàn toàn chất lượng phần mềm. Tuy nhiên:

* Đối với phần mềm quan trọng như y tế hoặc tài chính, việc tuân thủ tiêu chuẩn quốc tế là bắt buộc (ví dụ: FDA hoặc ISO).
* Chất lượng phần mềm được đánh giá qua khả năng đáp ứng yêu cầu thực tế và trải nghiệm người dùng hơn là tuân thủ nghiêm ngặt các quy chuẩn.

🡪 Điều quan trọng là đảm bảo phần mềm hoạt động ổn định, đáp ứng mục tiêu của khách hàng và người dùng cuối.

# KẾT LUẬN

Hệ thống báo cáo và dashboard mà nhóm xây dựng đã đạt được nhiều kết quả tích cực, cụ thể:

* **Hiệu quả hỗ trợ quản lý:** Các báo cáo trực quan và dễ hiểu, hỗ trợ các BI Users như Content Team, Finance Team, và Original Programming Team đưa ra các quyết định chiến lược nhanh chóng và hiệu quả hơn.
* **Ý nghĩa đối với doanh nghiệp:** Hệ thống giúp Netflix hiểu rõ hơn về thói quen người dùng, tối ưu danh mục nội dung, và cải thiện ROI từ các khoản đầu tư sản xuất nội dung. Điều này không chỉ nâng cao hiệu quả hoạt động mà còn cải thiện năng lực cạnh tranh.
* **Bài học thu được:** Nhóm đã nắm bắt quy trình ETL, cách xây dựng kho dữ liệu và tầm quan trọng của việc lựa chọn nền tảng BI phù hợp như Power BI.

**Khuyến nghị**

1. **Nâng cao chất lượng dữ liệu:**
   * Áp dụng các biện pháp làm sạch dữ liệu tự động để giảm lỗi.
   * Xây dựng cơ chế kiểm tra và đồng bộ dữ liệu từ các nguồn đầu vào khác nhau.
2. **Khai thác giá trị dữ liệu hiệu quả hơn:**
   * Kết hợp AI và machine learning để dự đoán xu hướng thị trường và tối ưu thuật toán gợi ý nội dung.
   * Sử dụng các báo cáo chi tiết hơn để theo dõi hiệu quả của từng nội dung tại từng khu vực.
3. **Tăng cường an toàn và bảo mật:**
   * Sử dụng mã hóa dữ liệu và quản lý quyền truy cập chặt chẽ.
   * Đảm bảo tuân thủ các quy định bảo mật quốc tế, như GDPR.
   * Cập nhật thường xuyên các biện pháp phòng chống rủi ro an ninh mạng.

# TÀI LIỆU THAM KHẢO

Marketing, P. (không ngày tháng). Được truy lục từ https://pamarketing.vn/mo-hinh-kinh-doanh-cua-netflix-p1/

*Netflix 10+ Year Stock Data [2002-2020]*. (2024). Được truy lục từ Kaggle: https://www.kaggle.com/datasets/aayushmishra1512/netflix-stock-data

*Netflix data with IMDB scores*. (2024). Được truy lục từ Kaggle: https://www.kaggle.com/datasets/sarahjeeeze/imdbfile