

# PHÁT TRIỂN ỨNG DỤNG DI ĐỘNG THU THẬP DỮ LIỆU TỪ NGƯỜI DÙNG PHỤC VỤ GIÁM SÁT MÔI TRƯỜNG

NGUYỄN MAI DUNG

Khoa Công nghệ thông tin, Trường Đại học Mở- Địa chất

## Tóm tắt:

Ở Việt Nam, tình trạng ô nhiễm môi trường diễn biến hết sức phức tạp, chất lượng môi trường nhiều nơi suy giảm mạnh. Đối với các chương trình giám sát môi trường, dữ liệu và các quan sát thu thập được đều có giá trị và bắt buộc với nhiều dự án để cung cấp thông tin cho các cơ quan quản lý. Do vậy việc có một công cụ hỗ trợ trong việc thu thập và chia sẻ dữ liệu môi trường là vô cùng cần thiết. Hơn nữa, cùng với sự phát triển của khoa học và công nghệ, Citizen Science đang nổi lên như một hướng mới cho việc xây dựng các nguồn dữ liệu cung ứng cộng đồng. Nó giúp việc thu thập dữ liệu trở nên thuận tiện, ít tốn kém và tương đối nhanh chóng. Bài báo trình bày về việc xây dựng phần mềm trên thiết bị di động với các chức năng phục vụ mục đích thu thập và chia sẻ dữ liệu giám sát môi trường. Trong đó, nghiên cứu tập trung chủ yếu vào hai loại dữ liệu chính là chất lượng không khí và chất lượng nước. Một ứng dụng đa nền tảng Android và iOS được xây dựng trên Flutter với các chức năng chính như cho phép người sử dụng tạo các báo cáo ô nhiễm môi trường, hiển thị thông tin khu vực ô nhiễm, gửi thông tin cảnh báo ô nhiễm, v.v. Sản phẩm phần mềm của nghiên cứu phần nào đó giúp cho việc phát hiện, tiếp nhận và cảnh báo các vấn đề liên quan đến môi trường trở nên nhanh chóng, chính xác và tiện lợi hơn.

Từ khóa: Dữ liệu giám sát môi trường, Ứng dụng di động đa nền tảng, Citizen Science

## 1. Giới thiệu

Mục tiêu của Citizen Science là hỗ trợ nghiên cứu khoa học bằng cách đóng góp dữ liệu thu thập qua nhiều hình thức. Sự gia tăng của các thành tựu của Citizen Science đã góp phần tích cực vào việc triển khai nghiên cứu các ứng dụng phân tích, quản lý các vấn đề liên quan đến môi trường, góp phần nâng cao hiểu biết về hiện trạng môi trường đồng thời thúc đẩy sự tham gia của cộng đồng trong việc

bảo vệ môi trường. Tuy nhiên, việc sử dụng dữ liệu thu thập này còn nhiều hạn chế do thiếu sự chuẩn xác trong các quy trình thu thập dữ liệu. Ngoài ra, chất lượng dữ liệu thường không được xác định và dữ liệu thường phân tán và không có cấu trúc. Sự phát triển của cuộc cách mạng công nghiệp 4.0 thúc đẩy việc sử dụng dữ liệu được thu thập thông qua phát triển hệ thống tích hợp các công cụ và phương pháp để tạo điều kiện thuận lợi cho việc thu thập, truy cập và xác nhận dữ liệu. Qua phân

tích vai trò của Citizen Science trong việc giám sát môi trường và phân tích tiềm năng của công nghệ thông tin trong việc tận dụng những lợi ích của việc sử dụng dữ liệu do người dân thu thập. Nghiên cứu đề xuất giải pháp sử dụng một nguồn cung ứng dữ liệu mở nhằm đa dạng hóa các loại dữ liệu được thu thập, cụ thể sẽ tập trung vào hai loại dữ liệu chính, phục vụ cho việc giám sát chất lượng nước và chất lượng không khí. Sản phẩm của nghiên cứu là một hệ thống ứng dụng trên thiết bị di động cho phép công chúng chia sẻ các thông tin của họ về tình trạng của môi trường được mô tả.

### 2. Nghiên cứu liên quan

Một số ứng dụng liên quan đến môi trường phổ biến tại Việt Nam có thể kể đến như: ứng dụng HealthyAIR, được phát triển bởi Viện Môi trường và Tài nguyên - Đại học Quốc gia TP Hồ Chí Minh. Ứng dụng cho phép mọi người biết mức độ ô nhiễm ở hiện tại và dự báo ô nhiễm không khí trong 12 – 24 giờ tới. HealthyAIR có thể đo được nhiều chất gây ô nhiễm không khí, hiển thị chất lượng không khí bằng chỉ số AQI. Đồng thời, đưa ra các khuyến cáo, cảnh báo về chất lượng không khí [1]. Ứng dụng Envisoft, được nghiên cứu và phát triển bởi Trung tâm Quan trắc môi trường miền Bắc thuộc Tổng Cục Môi trường – Bộ Tài nguyên và Môi trường. Một số chức năng chính của Envisoft bao gồm: hiển thị chỉ số AQI theo giờ và ngày hiện tại; cung cấp biểu đồ diễn biến giá trị AQI; đưa ra các khuyến nghị về sức khỏe; lựa chọn các trạm quan trắc theo khu vực; xếp hạng chất lượng không khí theo các trạm [2]. Envisoft và HealthyAIR chỉ tập trung vào các chỉ số ô nhiễm không khí và nguồn dữ liệu được lấy chủ yếu từ các trạm quan trắc. Ngoài ra, công ty cổ phần EKGIS cũng cung cấp một nền tảng

Govone nhằm hiện đại hóa công tác kiểm tra quản lý và giám sát môi trường [3]. Govone cho phép người sử dụng giám sát số liệu trực tuyến; cung cấp chức năng trên di động cho phép thực hiện kiểm tra môi trường qua các thông tin, hình ảnh, vị trí; cho phép người dân gửi các phản ánh môi trường bằng thiết bị di động, v.v. Qua việc tìm hiểu này, có thể thấy tại Việt Nam hiện nay chưa có một ứng dụng liên quan đến môi trường nào tập trung phục vụ công tác thu thập dữ liệu và tận dụng được nhiều nguồn cung cấp khác nhau đặc biệt là tự nguồn cung ứng cộng đồng.

### 3. Cơ sở lý thuyết

#### 3.1. Công nghệ sử dụng

Hệ thống bao gồm hai thành phần chính và được xây dựng theo kiến trúc máy chủ - máy khách. Về phía máy chủ, nó sẽ được xây dựng dựa trên NodeJS - một nền tảng vận hành trên V8 JavaScript runtime của Chrome. NodeJS được xây dựng và phát triển từ năm 2009, bảo trợ bởi công ty Joyent, trụ sở tại California, Hoa Kỳ [4]. Để có thể tạo ra các API cũng như phát triển các chức năng cho phía máy chủ, tác giả đã sử dụng ExpressJS. ExpressJS hỗ trợ các phương thức HTTP và middleware tạo ra một API vô cùng mạnh mẽ và dễ sử dụng [5]. Về một phía máy khách (phần ứng dụng), Flutter Framework được lựa chọn. Đây là một bộ phát triển phần mềm giao diện người dùng mã nguồn mở do Google phát triển, được sử dụng để tạo các ứng dụng đa nền tảng từ một cơ sở mã duy nhất cho Android, iOS, Linux, macOS, Windows và web. Flutter sử dụng Dart, một ngôn ngữ lập trình hướng đối tượng hiện đại. Dart sử dụng những trình biên dịch khác nhau để biên dịch ra mã máy tương ứng [6]. Một số ưu điểm của Flutter có thể kể đến như hiệu suất ứng dụng rất cao, giao tiếp gần như trực tiếp với hệ

thống, ngôn ngữ kiểu tĩnh nhưng với cú pháp hiện đại, có thể chạy được giả lập mobile ngay trên web, tiện cho việc phát triển [7]. MongoDB được lựa chọn làm nền tảng lưu trữ chính cho hệ thống. MongoDB là một cơ sở dữ liệu NoSQL hướng tài liệu đa nền tảng và mã nguồn mở hàng đầu hiện nay. Nó cung cấp một cơ sở dữ liệu đa nền tảng, hiệu suất cao và khả năng mở rộng dễ dàng [8]. Thay vì sử dụng các bảng và hàng như trong cơ sở dữ liệu quan hệ truyền thống, MongoDB sử dụng các bộ sưu tập (Collections) và tài liệu (Documents).

### 3.2. Phương pháp thu thập dữ liệu

Về cơ bản các nguồn thu thập dữ liệu của hệ thống hoạt động theo mô tả như hình Hình 1. Dữ liệu của hệ thống được cung cấp qua ba nguồn thu thập chính:

1) Nguồn dữ liệu từ người dùng chia sẻ thông qua ứng dụng.

2) Nguồn dữ liệu về môi trường được thu thập từ các website.

3) Nguồn dữ liệu được trích xuất từ các nền tảng mạng xã hội.

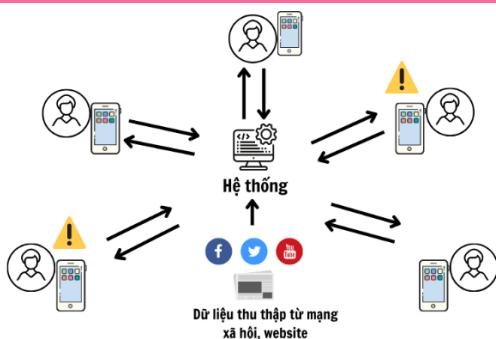
Cụ thể, khi người sử dụng truy cập vào ứng dụng, họ sẽ tham gia vào mạng lưới chia sẻ dữ liệu môi trường. Họ có thể thêm các địa điểm mà họ quan tâm về tình trạng môi trường hoặc chia sẻ các thông tin, dữ liệu về ô nhiễm môi trường tại một khu vực bất kỳ mà họ quan sát được thông qua biểu mẫu ứng dụng cung cấp (Hình 3). Các thông tin về vị trí được người dùng cấp phép quyền truy cập trên thiết bị di động thông qua ứng dụng, kết nối với Google Map API và được gửi tới hệ thống và lưu ở các trường *lat* và *lng* trong tài liệu Pollution của hệ thống được thiết kế trên cơ sở dữ liệu MongoDB (Bảng 1).

Bảng 1: Một số trường chính trong collection Pollution

Trường	Kiểu dữ liệu	Ý nghĩa
<i>_id</i>	objectId	Định danh thông tin chia sẻ
Images	Object	Hình ảnh về dữ liệu chia sẻ
qualityScore	double	Điểm đánh giá chất lượng
<i>lat</i>	double	Vĩ độ vị trí người dùng
<i>lng</i>	double	Kinh độ vị trí người dùng

Ngoài ra, hệ thống cũng tự động thu thập dữ liệu liên quan đến môi trường từ các website hoặc các nền tảng mạng xã hội như Facebook, Twitter, Youtube. Dữ liệu trên các nền tảng này sẽ chủ yếu được trích xuất thông qua các API mà từng nền tảng cung cấp. Thông tin ô nhiễm trong quá trình chia sẻ đến người dùng khác sẽ được kiểm duyệt bởi quản trị hệ thống. Người quản trị có thể cấu hình hệ thống phê duyệt tự động hoặc thủ công. Theo

mặc định, hệ thống sẽ đặt chế độ phê duyệt thủ công để đảm bảo tính chính xác của các báo cáo ô nhiễm nhằm nâng cao chất lượng thông tin cho hệ thống. Đây cũng có thể coi như một hình thức tham gia vào Citizen Science, khi nó cho phép người dân sử dụng ứng dụng để đóng góp vào việc thu thập và quản lý môi trường tại địa phương.



Hình 1: Mô hình thu thập dữ liệu của hệ thống

### 3.3. Tạo cảnh báo ô nhiễm từ hệ thống

Một chức năng quan trọng của hệ thống là việc hiển thị các cảnh báo về tình trạng ô nhiễm môi trường được gửi từ các người sử dụng tham gia hệ thống. Hệ thống sẽ gửi báo cáo của người dùng đến quản trị hệ thống hoặc quản trị khu vực. Bước tiếp theo, người quản trị có nhiệm vụ kiểm tra và xác minh các thông tin được gửi từ báo cáo của người dùng trên. Lúc này sẽ có hai luồng sự kiện xảy ra: một là quản trị viên từ chối thông tin nhận được, hai là xác nhận thông tin. Nếu trong trường hợp quản trị viên xác nhận thông tin. Hệ thống sẽ gửi yêu cầu đến Firebase Cloud Messaging để có thể gửi các thông báo đến những người dùng trong khu vực liên quan đến thông tin ô nhiễm. Các thông tin ô nhiễm cũng sẽ có thể xem chi tiết thông qua các chức năng của ứng dụng. Đặc biệt việc tích hợp dữ liệu bản đồ thông qua Google Map API giúp cho việc trực quan hóa các khu vực ô nhiễm được cụ thể và chính xác hơn.

## 4. Phân tích và xây dựng hệ thống

### 4.1. Chức năng của hệ thống

Hệ thống có ba kiểu người dùng: quản trị viên, quản trị khu vực, người dùng thông thường. Với mỗi kiểu người dùng, hệ thống sẽ cung cấp các chức năng cụ thể như sau:

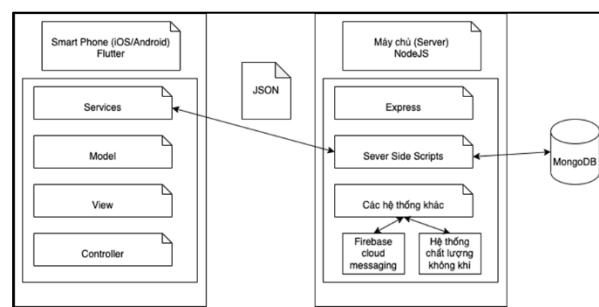
- Chức năng của quản trị viên: Quản lý người dùng, Quản lý tất cả ô nhiễm, Thông báo khẩn cấp, Xem các thông kê.

- Chức năng quản trị khu vực: Quản lý ô nhiễm trong khu vực, Thông báo khẩn cấp, Xem các thông kê.

- Chức năng người dùng thông thường: Tạo báo cáo ô nhiễm, Xem thông tin ô nhiễm, Nhận thông tin dự báo chất lượng không khí, Xem thông tin chất lượng không khí, Xem tin tức.

### 4.2. Kiến trúc của hệ thống

Kiến trúc của hệ thống bao gồm ba thành phần chính. Thành phần thứ nhất là ứng dụng được cài đặt trên thiết bị di động thông minh hệ điều hành Android/iOS. Ứng dụng được xây dựng trên nền tảng Flutter. Thành phần thứ hai là máy chủ được xây dựng bằng NodeJS. Máy chủ có nhiệm vụ kết nối đến cơ sở dữ liệu MongoDB - đây cũng chính là thành phần thứ ba trên hệ thống. Ứng dụng trên thiết bị di động sẽ giao tiếp với thành phần máy chủ thông qua các RESTful API. Hình 2 dưới đây thể hiện các thành phần trong kiến trúc hệ thống.



Hình 2: Kiến trúc hệ thống

### 4.3. Phân tích cơ sở dữ liệu

Cơ sở dữ liệu của hệ thống bao gồm các thành phần như sau:

(1) Users: bộ sưu tập này được sử dụng để lưu trữ các thông tin của người dùng trong hệ thống như tên người dùng, email, vai trò, vị trí của người dùng, v.v.

(2) Alerts: bộ sưu tập này được sử dụng để lưu thông tin các cảnh báo của hệ thống.

(3) Notifications: bộ sưu tập này được sử dụng để lưu trữ thông tin về các thông báo trên hệ thống.

(4) NotificationAlerts: bộ sưu tập này được sử dụng để lưu trữ thông tin về các thông báo cảnh báo trên hệ thống.

(5) Fcmtokens: bộ sưu tập này nhằm lưu trữ thông tin Token của nền tảng Firebase Cloud Message.

(6) Pollution: bộ sưu tập này dùng để lưu trữ thông tin báo cáo ô nhiễm.

(7) Tokens: bộ sưu tập này nhằm lưu trữ thông tin mã xác thực của người dùng trên hệ thống.

### 5. Kết quả và đánh giá

Hệ thống nghiên cứu được cài đặt thành công trên đa nền tảng ứng dụng phổ biến như Android và iOS. Hình 3 dưới đây mô tả chức năng tạo báo cáo ô nhiễm môi trường (thu thập dữ liệu từ người dùng) của hệ thống. Khi người dùng chọn chức năng tạo báo cáo ô nhiễm, màn hình ứng dụng sẽ hiển thị biểu mẫu nhập thông tin ô nhiễm như địa chỉ, loại ô nhiễm, mức độ ô nhiễm, mô tả, hình ảnh minh họa. Đây chính là chức năng giúp hệ thống có thể thu thập được các dữ liệu về môi trường do người sử dụng cung cấp.



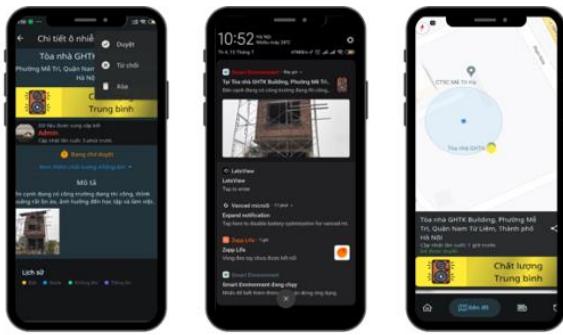
Hình 3: Chức năng tạo báo cáo ô nhiễm môi trường trên hệ thống

Đối với quy trình tạo cảnh báo từ hệ thống (Hình 4), sau khi người dùng tạo báo cáo. Hệ thống sẽ gửi báo cáo đến Quản trị viên hoặc Quản trị khu vực. Người quản trị chọn hành động như “Duyệt” hay “Từ chối” để xác nhận thông tin báo cáo. Khi duyệt báo cáo, ứng dụng gửi thông tin lên server thông qua API. Tại server sẽ thay đổi trạng thái báo cáo đó thành “Đã duyệt” và lưu vào cơ sở dữ liệu. Đồng thời, khi trạng thái là “Đã duyệt”, phía server sẽ lấy tất cả người dùng đang ở vị trí gần với báo cáo để gửi thông báo.

Ngoài ra, một số chức năng khác của hệ thống như tạo tài khoản, xem chỉ số chất lượng không khí, xem các tin tức ô nhiễm theo dữ liệu thu thập từ mạng xã hội và website, xem các thông tin khu vực ô nhiễm dưới dạng bản đồ, v.v. (Hình 5).



Hình 4: Quy trình tạo cảnh báo của hệ thống



Hình 5: Một số giao diện khác của ứng dụng

### 6. Kết luận

Trong nghiên cứu này, nhóm tác giả đã trình bày về phương pháp xây dựng ứng dụng thu thập dữ liệu giám sát môi trường do người dùng chia sẻ. Hệ thống được xây dựng dựa trên những công nghệ hiện đại như Flutter, NodeJS, MongoDB. Qua quá trình thử nghiệm trên nhiều thiết bị, có thể thấy các chức năng hoạt động ổn định. Quy trình tạo báo cáo ô nhiễm môi trường, giúp người sử dụng có thể dễ dàng cung cấp các thông tin về tình trạng môi trường trên hệ thống. Từ đó hình thành một mạng lưới người dùng chia sẻ dữ liệu như một hình thức tham gia vào Citizen Science giúp cho nguồn cung ứng dữ liệu của hệ thống trở nên phong phú và đa dạng hơn. Việc hiển thị các cảnh báo đến người dùng cũng được diễn ra dưới dạng thời gian thực. Tính năng chạy ngầm trên thiết bị cũng được tối ưu để không chiếm quá nhiều tài nguyên của thiết bị. Có thể thấy, ứng dụng đã tạo ra những kết quả ban đầu trong việc thu thập dữ liệu và giám sát các vấn đề liên quan đến môi trường.○

\* Bài báo là một phần kết quả nghiên cứu của đề tài cấp Bộ Giáo dục và Đào tạo, mã số B2022-MDA-01.

### Tài liệu tham khảo

[1]. P. A. -. T. THẢO, "RA MẮT APP CẢNH BÁO CHẤT LƯỢNG KHÔNG KHÍ TÍCH HỢP AI," vnuhcm, 2022. [Online]. Available: [https://vnuhcm.edu.vn/su-kien\\_33356864/ra-mat-app-canhanh-bao-chat-luong-khong-khi-tich-hop-ai/343232326864.html](https://vnuhcm.edu.vn/su-kien_33356864/ra-mat-app-canhanh-bao-chat-luong-khong-khi-tich-hop-ai/343232326864.html).

[2]. T. Minh, "Cài đặt ứng dụng Envisoft trên di động để biết thông tin ô nhiễm không khí," Báo tài nguyên môi trường, 2020. [Online]. Available: <https://baotainguyenmoitruong.vn/cai-dat-ung-dung-envisoft-tren-di-dong-de-biet-thong-tin-o-nhiem-khong-khi-299736.html>.

[3]. Govone, "Hệ thống cơ sở dữ liệu GIS quan trắc môi trường phục vụ chính quyền điện tử và đô thị thông minh," Govone, [Online]. Available: <https://govone.vn/moitruong/>.

[4]. "Niithanoi," 10 2021. [Online]. Available: <https://niithanoi.edu.vn/nodejs-la-gi-tong-hop-day-du-ve-nodejs-ban-can-biet.html>.

[5]. TopDev, "Expressjs là gì? Tất tần tật về Express.js," TopDev, [Online]. Available: <https://topdev.vn/blog/express-js-la-gi/>.

[6]. M. Napoli, in *Introducing Flutter and Getting Started*, 2019, pp. 1-23.

[7]. fsoft, "Flutter cho người mới bắt đầu. Giới thiệu. Ưu và Nhược điểm," 2021. [Online]. Available: <https://fsoft-academy.edu.vn/tin-tuc-su-kien/flutter-cho-nguoimoi-bat-dau-gioi-thieu-uu-va-nhuoc-diem/>.

[8]. T. H. LÊ, "MongoDB là gì?," 2022, [Online]. Available: <https://www.bkns.vn/mongodb-la-gi.html>.○

## Summary

### Development of a mobile app to collect users' data for environmental monitoring

*Nguyen Mai Dung*

*Hanoi University of Mining and Geology*

In Vietnam, the environmental pollution situation is very complicated; the quality of the environment has declined sharply in many places. The data and observations collected in environmental monitoring programs are valuable and required for many projects to provide information to regulatory agencies. Therefore, having a tool to support collecting and sharing environmental data is essential. Moreover, along with the development of science and technology, Citizen Science is emerging as a new direction for building community-sourcing data sources. It makes data collection convenient, inexpensive and relatively fast. This paper presents software development on mobile devices with functions to collect and share environmental monitoring data. In particular, the study focuses mainly on two types of data: air quality and water quality. A cross-platform Android and iOS application built on Flutter with key functions such as allowing users to create environmental pollution reports, display contaminated area information, and send pollution warning information, etc. The software product of the research partly helps to detect, receive and warn of environmental-related problems more quickly, accurately and conveniently. O

Keywords: Environmental monitoring data, Cross-platform mobile application, Citizen Science.