

BÁO CÁO

THIẾT KẾ VÀ XÂY DỰNG CƠ SỞ DỮ LIỆU PHỤC VỤ CHO MỘT TRANG BÁO ĐIỆN TỬ

I. Phân tích chức năng và dữ liệu cần lưu trữ cho trang báo điện tử

1. Chức năng:

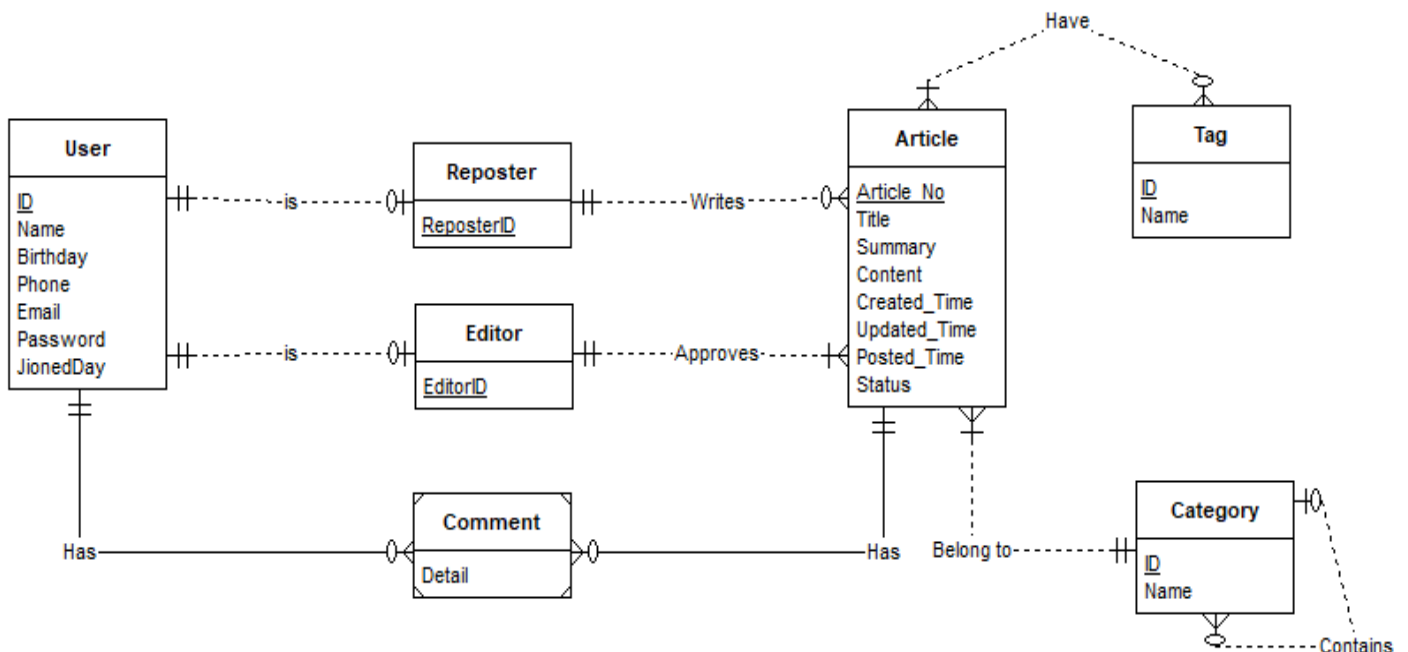
- Có 3 cấp độ người dùng: Người dùng là Độc Giả bình thường, người dùng là Phóng Viên và người dùng là Biên Tập Viên.
- Với các Độc Giả bình thường: có thể tạo tài khoản đăng nhập, liên kết qua email hoặc số điện thoại. Truy cập xem các tin nóng từ trang chủ, tin mới nhất, hoặc truy cập danh mục bất kỳ để xem nội dung mà mình muốn trong 1 lĩnh vực nhất định.
- Chức năng đăng nhập của người dùng dưới vai trò Độc Giả được dùng để các nhân hóa trải nghiệm như: có thể bình luận nêu ý kiến, bày tỏ cảm xúc với các bình luận từ người dùng khác, chỉnh sửa danh mục ưu tiên hiển thị, lưu đánh dấu các bài báo để lần sau đọc lại, nhận thông báo về bài báo mới qua email.
- Các Độc Giả cũng có thể gửi các bài ý kiến qua chuyên mục Ý Kiến, Góc Nhìn hoặc chuyên mục Tâm Sự.
- Đối với người dùng dưới vai trò Phóng Viên trực thuộc mỗi lĩnh vực nhất định sẽ viết bài báo gửi về cho tòa soạn.
- Các Biên Tập Viên trực thuộc mỗi lĩnh vực sẽ chịu trách nhiệm xét duyệt bài báo, nếu đạt yêu cầu sẽ được đăng, nếu không sẽ được yêu cầu chỉnh sửa hoặc xóa.

2. Dữ liệu:

- Nội dung bài báo: Tiêu đề, bản tóm tắt để hiển thị trên trang chủ, nội dung chi tiết khi được truy cập, chuyên mục, tag, ngày đăng bài, tác giả (phóng viên), người duyệt (biên tập viên).
- Dữ liệu người dùng: Họ tên, ngày sinh, số điện thoại, email, mật khẩu, ngày tham gia, số lượt bình luận.
- Các bài ý kiến của độc giả, bình luận.
- Dữ liệu về phóng viên: Tên, các bài đã đăng.

II. Mô hình ERD và giải thích

1. Mô hình ERD



2. Giải thích mô hình ERD

2.1. Các thực thể (7):

2.1.1. Thực thể Article (Bài báo):

Chứa các thông tin cần thiết để một bài báo hiển thị đầy đủ nội dung trên trang chủ trang web, nơi người dùng có thể truy cập và đọc chúng

- Article_No: Mã số báo
- Title: Tiêu đề bài báo
- Summary: Nội dung tóm tắt hiện trên trang chủ
- Content: Nội dung chi tiết khi độc giả nhấp vào link bài báo để xem
- Created_Time: Thời gian bài báo được tạo (chờ phê duyệt)
- Updated_Time: Thời gian cuối cùng bài báo được chỉnh sửa (áp dụng cho các trường hợp cần sửa đổi, hoặc các bản tin live cập nhật trực tiếp)
- PostedTime: Thời gian bài báo được đăng lên trang web (đã phê duyệt)
- Status: có 4 trạng thái: Pending (chờ duyệt), Editing (chờ sửa), Scheduled (được lên lịch), Posted (đã đăng)
- Visit: Số lượt truy cập của bài báo (Cần lưu trữ để đánh giá chất lượng thông tin, đưa lên trang nhất tin hot, hoặc đánh giá thị hiếu độc giả về việc quan tâm đến các tin tức nào)

2.1.2. Thực thể User (Người dùng):

Chứa thông tin người dùng, các người dùng này có thể là Biên Tập Viên, Phóng Viên. Nếu không có bất cứ liên kết nào với 2 thực thể trên, sẽ chỉ là Độc Giả bình thường và không có các tính năng nâng cao.

- ID: Mã số người dùng
- Name: Tên người dùng
- Birthday: Ngày sinh
- Phone: Số điện thoại
- Email: Địa chỉ thư điện tử
- Password: Mật khẩu đăng nhập
- JoinedDay: Ngày tham gia

2.1.3. Thực thể Reposter (Phóng viên tòa soạn):

Làm việc dưới sự quản lý trực tiếp của tòa soạn, công việc chính là thu tập tin tức, quay phim, chụp ảnh, công tác hiện trường, viết bài báo.

- ReposterID: Đây cũng chính là UserID khi đã được liên kết

2.1.4. Thực thể Editor (Biên tập viên tòa soạn):

Nhận bài viết của các Phóng Viên để sàng lọc, chỉnh sửa hình thức, ngôn từ, kiểm tra nguồn thông tin để xác định độ chính xác của bài viết trước khi xuất bản, nhằm đảm bảo uy tín cho tòa soạn cũng như các Phóng Viên.

Các Biên Tập Viên có thể yêu cầu các Phóng Viên sửa lại bài nếu chưa đạt yêu cầu, xóa bài viết, hoặc lên lịch đăng khi bài đã được duyệt.

- EditorID: Đây cũng chính là UserID khi đã được liên kết

2.1.5. Thực thể Category (Đề mục):

Các bài báo được phân loại theo đề mục để người dùng dễ truy cập, và cũng như để dễ quản lý dưới cấp độ cơ sở dữ liệu. VD: Thể Thao, Thời Sự, Khoa Học,

- ID: Mã số đề mục
- Name: Tên đề mục

2.1.6. Thực thể Tag (Thẻ):

Ngoài việc phân loại theo đề mục, các bài báo còn có thể được phân loại và liên kết theo thẻ về các chủ đề nóng, truy cập nhiều, ... để độc giả tiện truy cập.

VD: Lê Văn Luyện, U23 Việt Nam, Chiến tranh Nga-Ukraina,

- ID: Mã số thẻ
- Name: Tên thẻ

2.1.7. Thực thể yếu Comment (Bình luận):

Chứa thông tin về bình luận của Độc Giả trong các bài báo.

- Detail: Nội dung bình luận

2.2. Các mối quan hệ (9):

Giới hạn min của mỗi quan hệ có thể bắt buộc – mandatory (ít nhất 1 thực thể tham gia) hoặc không bắt buộc – Optional (có thể không có thực thể nào).

Có 3 mối quan hệ chính: One-One, One-Many và Many-Many.

Ta quy ước, nếu để trống thì giới hạn min chính là Mandatory (bắt buộc).

2.2.1. User (One) – Reposter (One-Optional)

User là thực thể chứa các thông tin người dùng, một Reposter tồn tại cần phải được liên kết với một và chỉ một User.

Nhưng một User không nhất định phải là Reposter.

2.2.2. Reposter (One) – Article (Many-Optional)

Một hoặc nhiều Bài Báo có thể được viết bởi một Phóng Viên.

Nhưng một Phóng Viên có thể chưa viết được Bài Báo nào (*chủ yếu nằm ở các trường hợp mới cộng tác*).

2.2.3. User (One) – Editor (One-Optional)

User là thực thể chứa các thông tin người dùng, một Editor tồn tại cần phải được liên kết với một và chỉ một User.

Nhưng một User không nhất định phải là Editor.

2.2.4. Editor (One) – Article (Many)

Một hoặc nhiều bài báo có thể được duyệt bởi 1 Biên tập viên.

Và 1 Biên tập viên phải có trách nhiệm duyệt ít nhất 1 bài báo.

2.2.5. User (One) – Comment (Many-Optional)

Một Độc Giả có thể để lại nhiều bình luận hoặc không.

Một bình luận được hình thành thì chắc chắn phải thuộc về 1 Độc Giả.

2.2.6. Article (One) – Comment (Many-Optional)

Một bài báo có thể có rất nhiều bình luận hoặc không.

Một bình luận được hình thành thì chắc chắn chỉ nằm trong 1 bài báo.

2.2.7. Category (One) – Article (Many)

Một chuyên mục có thể có 1 hoặc nhiều bài báo.

Nhưng 1 bài báo chỉ thuộc về 1 chuyên mục nhất định.

2.2.8. Category (One-Optional) – Category (Many-Optional) (Quan hệ tự tham chiếu)

Một đề mục lớn có thể chứa 0 hoặc nhiều đề mục nhỏ hơn.

Một đề mục nhỏ có thể được chứa bởi 0 hoặc 1 đề mục lớn hơn (*một đề mục tổng thì không thể được chứa bởi bất kỳ đề mục nào khác*).

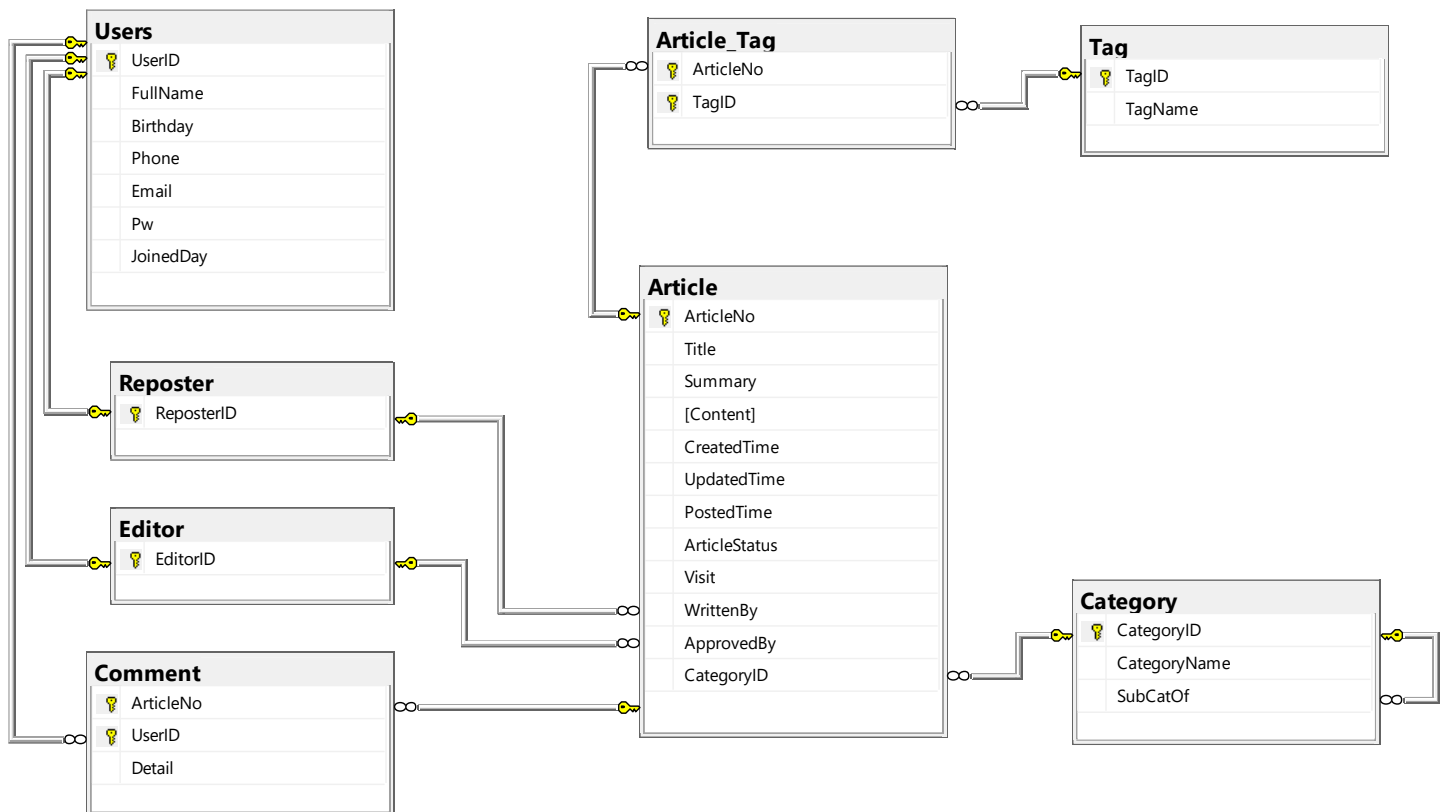
2.2.9. Article (Many) – Tag (Many-Optional)

Mỗi bài báo có thể được gắn 0 hoặc nhiều thẻ.

Mỗi thẻ khi được tạo ra phải được liên kết với ít nhất 1 bài báo.

III. Lược đồ cơ sở dữ liệu

3.1. Lược đồ



7 thực thể: User, Reposter, Editor, Article, Category, Tag, Comment được chuyển thành các bảng tương ứng, bao gồm các thuộc tính thành phần, và thêm thuộc tính là khóa ngoại đại diện cho các mối quan hệ.

Ngoài ra, tạo thêm 1 bảng mới: Article_Tag tương ứng với mối quan hệ Many-To-Many giữa bảng giữa Article và Tag.

3.2. Bảng và các ràng buộc trong bảng

3.2.1. User

- Khóa chính {UserID}: Xác định người dùng bất kỳ.
- Unique & Not Null {Email, Phone}: Ngoài UserID, email và phone cũng có thể được sử dụng để xác định người dùng bất kỳ vì mọi email và phone đăng ký của mỗi người đều khác nhau.
- Check {Phone}: Đảm bảo số phải đủ 10 số, không được ít hoặc nhiều hơn
- Not Null {FullName, Birthday, Phone, Pw, JoinedDay}: Dữ liệu bắt buộc tối thiểu cần có nên không thể để trống.

3.2.2. Reposter

- Khóa chính {ReposterID}: Đồng thời cũng là khóa ngoại tham chiếu từ bảng User, giúp xác định một User là Reposter. Khi đó User này sẽ được cấp toàn bộ chức năng mà một Reposter có thể có.
- Khóa ngoại & Not Null {ReposterID}

3.2.3. Editor

- Khóa chính {EditorID}: Đồng thời cũng là khóa ngoại tham chiếu từ bảng User, giúp xác định một User là Editor. Khi đó User này sẽ được cấp toàn bộ chức năng mà một Editor có thể có.
- Khóa ngoại & Not Null {EditorID}

3.2.4. Article

- Khóa chính {ArticleID}: Xác định bài báo bất kỳ.
- Unique & Not Null {Title}: Tiêu đề bài báo phải là duy nhất và không được trống, để tránh việc các bài báo trùng nhau và các phóng viên đưa tin về một nội dung nhiều lần.
- Not Null {Summary, Content, Created_Time, Updated_Time, Posted_Time, Status}: Đây là các thông tin cần thiết cho 1 bài báo và không được để trống.
- Khóa ngoại & Not Null {Written_By, Approved_By, CategoryID}: lần lượt được tham chiếu từ khóa chính của 3 bảng: Reposter (ReposterID), Editor (EditorID) và Category (CategoryID).
3 ràng buộc này chỉ ra: Người viết bài báo phải là Phóng Viên, chỉ có Biên Tập Viên mới có thể duyệt bài, và một bài báo phải thuộc về 1 đề mục nhất định.

3.2.5. Category

- Khóa chính {CategoryID}: Xác định đề mục
- Unique & Not Null {CategoryName}: Tên đề mục phải là duy nhất, không được trống và không được trùng
- Khóa ngoại {SubCATof}: được tham chiếu từ khóa chính của bảng Category, nhằm xác định một đề mục có phải là đề mục con của một đề mục khác hay không. Đây chính là quan hệ tự tham chiếu trong ERD.

3.2.6. Tag

- Khóa chính {TagID}: Xác định thẻ
- Unique & Not Null {TagName}: Giống như đề mục, tên thẻ phải là duy nhất, không được trống và không được trùng.

3.2.7. Comment

- Khóa chính {Article_No, UserID}: tập hợp của 2 trường, cũng đồng thời là khóa ngoại, giúp xác định được bình luận của User nào trong bài báo nào.
- Khóa ngoại & Not Null {Article_No}, {UserID}: được tham chiếu từ khóa chính của 2 bảng Article (Article_No) và User (UserID).
- Not Null {Detail}: một bình luận được hình thành và được duyệt qua thì không thể được bỏ trống.

3.2.8. Article_Tag

Bảng này được chuyển đổi từ mối quan hệ Many-to-Many giữa 2 thực thể Article và Tag. Giúp xác định bài báo được gắn các thẻ nào, và các thẻ liên kết bao gồm các bài báo nào.

- Khóa chính {Article_No, TagID}: tập hợp của 2 trường.
- Khóa ngoại & Not Null {Article_No}, {TagID}: được tham chiếu từ khóa chính của 2 bảng Article (Article_No) và Tag (TagID).

3.3. Phụ thuộc hàm và chuẩn hóa BCNF

Xét về chuẩn hóa BCNF, các bảng dưới đây đều đáp ứng các tiêu chuẩn chuẩn hóa:

- Không chứa thuộc tính suy ra, đa trị hoặc nhóm các thuộc tính lặp \Rightarrow Đáp ứng 1NF
- Không chứa các phụ thuộc hàm bộ phận (EX: Các thuộc tính thường phụ thuộc vào 1 phần của khóa chính, đối với các khóa chính có từ 2 tập con) \Rightarrow Đáp ứng 2NF
- Không chứa các phụ thuộc hàm bắc cầu (EX: Khóa chính X xác định Y, trong khi Y lại xác định Z) \Rightarrow Đáp ứng 3NF
- Khóa chính không phụ thuộc hàm vào các thuộc tính không khóa \Rightarrow Đáp ứng BCNF

3.3.1. User_Info

ID → {Name, Birthday, Phone, Email, Password, JoinedDay}

Email → ID

Bảng này gồm 2 khóa ứng viên là ID và Email, nhưng do ta chọn ID làm khóa chính nên lúc này Email sẽ trở thành khóa phụ.

3.3.2. Article

Article_No → {Title, Summary, Content, Created_Time, Updated_Time, Posted_Time, Status, Visit}

Title → Article_No

3.3.3. Category

ID → {Name, SubCATof}

Name → ID

3.3.4. Tag

ID → Name

Name → ID

3.3.5. Comment

{UserID, Article_No} → Detail