

ASSIGNMENT 1
MÔN HỌC: CÁC HỆ CƠ SỞ DỮ LIỆU
MÃ MÔN HỌC: DBI202x_2.1-A_VN

1. Phân tích các yêu cầu chức năng, yêu cầu dữ liệu của trang báo điện tử

1.1. Yêu cầu chức năng

- Cho phép người đọc xem và đọc các bài báo mới được cập nhật lên tại trang chủ, có thể xem được tóm tắt và toàn bộ nội dung chi tiết của bài báo. Các bài báo mới nhất được cập nhật sẽ dựa trên ngày đăng bài.
- Cho phép người đọc lọc bài báo dựa trên các tiêu chí như ngày đăng, chủ đề, tên bài báo, ...
- Người đọc có thể viết bình luận ở cuối các bài báo để bày tỏ ý kiến của bản thân, không bắt buộc người đọc khi đọc bài phải để lại bình luận.
- Phóng viên có thể viết một hoặc nhiều bài báo.
- Biên tập viên quản lý các bài báo, thực hiện phê duyệt các yêu cầu đăng báo của phóng viên.

1.2. Yêu cầu dữ liệu

- Dữ liệu bao gồm 3 loại người dùng: Người đọc, Phóng viên và Biên tập viên.
- Thông tin của Người đọc: Mật khẩu, Họ và tên, Giới tính, Email, Ngày tham gia báo điện tử.
- Thông tin của Biên tập viên: Mật khẩu, Họ và tên, Giới tính, Email.
- Thông tin của Phóng viên: Mật khẩu, Họ và tên, Giới tính, Email, Bút danh.
- Các bài báo sẽ bao gồm: Tên bài báo, Tóm tắt bài báo, Nội dung chi tiết, Ngày viết báo, Ngày đăng, Ngày chỉnh sửa bài báo, Tác giả, Người duyệt bài, Trạng thái (Not Approved, Approved), Số lượng truy cập.
- Ngoài ra, mỗi bài báo sẽ nằm trong một chủ đề nhỏ, mỗi chủ đề nhỏ thuộc một chủ đề lớn, và mỗi chủ đề đều có tên.
- Các bình luận sẽ gồm: Nội dung bình luận và ngày đăng bình luận. Nếu người đọc đọc bài báo nhưng không để lại bình luận thì có thể để trống nội dung bình luận và ngày bình luận.

2. Mô hình ERD và thiết kế lược đồ cơ sở dữ liệu quan hệ của trang báo điện tử

2.1. Chi tiết mô hình ERD

2.1.1. Các thực thể, thuộc tính và khoá chính của chúng

- *Thực thể Viewer:*

Thực thể Viewer (Người đọc) là một trong ba loại thực thể người dùng. Thực thể người đọc bao gồm các thuộc tính:

- Viewer_ID: Mã số xác định của người đọc, mã số này duy nhất (Là khoá chính của thực thể Viewer).
- Password: Mật khẩu đăng nhập vào trang báo điện tử của người đọc.
- Full_Name: Họ và tên đầy đủ của người đọc.
- Gender: Giới tính của người đọc.
- Email: Email của người đọc.
- Joining_Date: Ngày người đọc mở tài khoản tham gia trang báo điện tử.

- *Thực thể Reporter:*

Thực thể Reporter là thực thể Phóng viên, bao gồm các thông tin sau:

- Reporter_ID: Mã số xác định của phóng viên, mã số này duy nhất (Là khoá chính của thực thể Reporter).
 - Password: Mật khẩu đăng nhập vào trang báo điện tử của phóng viên.
 - Full_Name: Họ và tên đầy đủ của phóng viên.
 - Email: Email của phóng viên.
 - Gender: Giới tính của phóng viên.
 - Pseudonym: Bút danh của phóng viên.
- *Thực thể Editor:*

Thực thể Editor là thực thể Biên tập viên, bao gồm các thông tin sau:

- Editor_ID: Mã số xác định của biên tập viên, mã số này duy nhất (Là khoá chính của thực thể Editor).
- Password: Mật khẩu đăng nhập vào trang báo điện tử của biên tập viên.
- Full_Name: Họ và tên đầy đủ của biên tập viên.
- Gender: Giới tính của biên tập viên.
- Email: Email của biên tập viên.

- *Thực thể Article:*

Thực thể này nêu rõ các thuộc tính chi tiết cho mỗi bài báo trên trang báo điện tử, cụ thể:

- Article_ID: Mã định danh cho mỗi bài báo (Là khoá chính của thực thể Article).
- Article_Name: Tên bài báo.
- Summary: Tóm tắt nội dung của bài báo.
- Full_Content: Nội dung chi tiết của bài báo (Bao gồm hình ảnh).
- Creating_Date: Ngày giờ khởi tạo bài báo.
- Posting_Date: Ngày giờ bài báo được đăng chính thức lên trang web báo điện tử, là ngày bài báo được Biên tập viên phê duyệt. Chỉ hiển thị khi trạng thái là “Approved”, nếu trạng thái là “Not Approved”, thông tin này có thể để trống.
- Updating_Date: Ngày giờ lần cuối cùng bài báo được chỉnh sửa.
- Status: Trạng thái phê duyệt của bài báo (Not Approved, Approved).
- Access_Number: Số lượng người đọc truy cập vào bài báo.

- *Thực thể Category:*

Một bài báo thuộc một chủ đề nào đó. Một chủ đề bao gồm các thuộc tính sau:

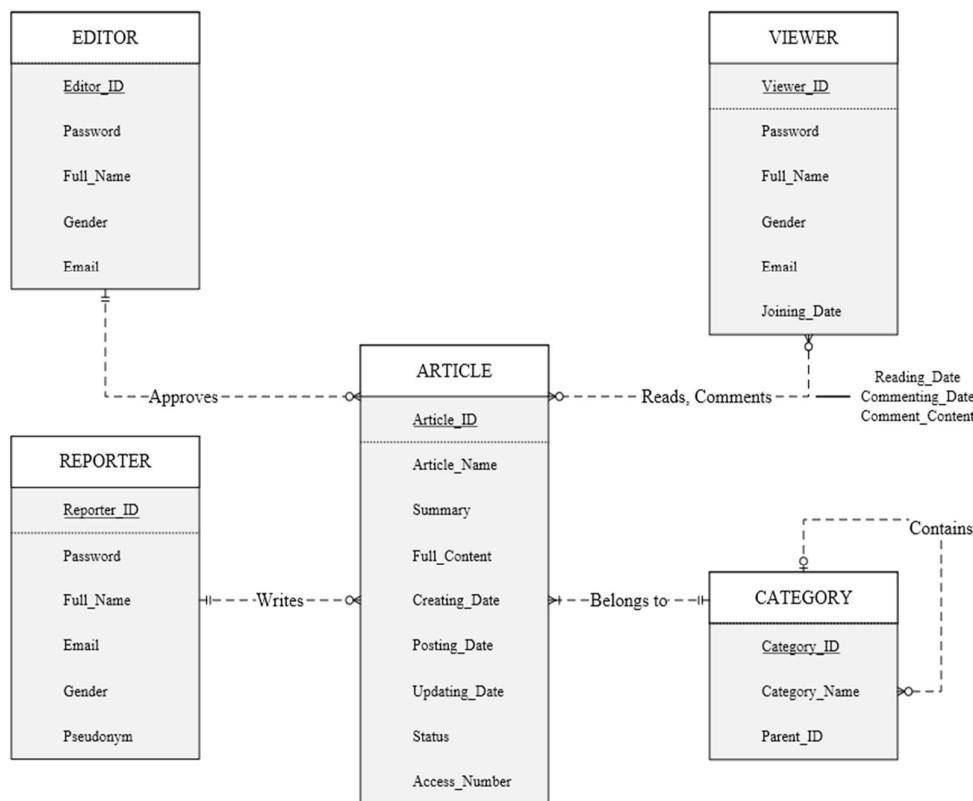
- Category_ID: Mã số xác định của chủ đề, mã số này là duy nhất (Là khoá chính của thực thể Category).
- Category_Name: Tên của chủ đề.
- Parent_ID: Mã số xác định của chủ đề lớn.

2.1.2. Các quan hệ và cardinality (lượng số) trong ERD

- Quan hệ giữa Viewer và Article (many-to-many): Một người đọc có thể đọc và bình luận nhiều bài báo, và một bài báo cũng được đọc và bình luận bởi nhiều người đọc khác nhau.

Cả Viewer và Article đều không bắt buộc. Quan hệ này có chứa các thuộc tính Ngày đọc bài báo, Ngày để lại bình luận, Nội dung bình luận.

- Quan hệ giữa Editor và Article (one-to-many): Một biên tập viên có thể duyệt nhiều bài báo, nhưng một bài báo chỉ được duyệt bởi một biên tập viên. Article không bắt buộc.
- Quan hệ giữa Reporter và Article (one-to-many): Một phóng viên có thể viết nhiều bài báo, nhưng một bài báo chỉ được viết bởi một phóng viên. Article không bắt buộc.
- Quan hệ Article và Category (one-to-many): Một chủ đề có thể chứa nhiều bài báo, nhưng một bài báo chỉ thuộc duy nhất một chủ đề.
- Quan hệ tự tham chiếu trong Category (one-to-many): Một chủ đề lớn có thể chứa nhiều chủ đề nhỏ, nhưng một chủ đề nhỏ chỉ nằm duy nhất trong một chủ đề lớn. Chủ đề (lớn và nhỏ) không bắt buộc.



Hình 2.1. Mô hình quan hệ thực thể ERD của trang báo điện tử

2.2. Chuyển đổi ERD sang lược đồ cơ sở dữ liệu quan hệ

2.2.1. Quy tắc loại thực thể

- Các thực thể Viewer, Editor, Reporter, Article và Category đều được chuyển thành các bảng tương ứng gồm chi tiết khoá chính và các cột như sau:
 - Bảng Viewer: Khoá chính là Viewer_ID, các cột bao gồm Password, Full_Name, Gender, Email, Joining_Date.
 - Bảng Editor: Khoá chính là Editor_ID, các cột bao gồm Password, Full_Name, Gender, Email.

- Bảng Reporter: Khoá chính là Reporter_ID, các cột bao gồm Password, Full_Name, Gender, Email, Pseudonym.
- Bảng Article: Khoá chính là Article_ID, các cột bao gồm Article_Name, Summary, Full_Content, Creating_Date, Posting_Date, Updating_Date, Status, Access_Number.
- Bảng Category: Khoá chính là Category_ID, cột Category_Name, Parent_ID.

2.2.2. Quy tắc quan hệ one-to-many (1-M)

- Quan hệ 1-M “Writes” giữa Reporter và Article: Author trở thành khoá ngoại trong bảng Article. Author tương ứng với Report_ID.
- Quan hệ 1-M “Approves” giữa Editor và Article: Approver trở thành khoá ngoại trong bảng Article. Approver tương ứng với Editor_ID.
- Quan hệ 1-M “Belongs to” giữa Article và Category: Category_ID trở thành khoá ngoại trong bảng Article.

2.2.3. Quy tắc quan hệ many-to-many (M-N)

- Quan hệ M-N “Reads, Comments” giữa Viewer và Article được chuyển thành bảng Viewer_Article với khoá chính là kết hợp 2 cột Viewer_ID và Article_ID (Từng cột riêng biệt sẽ là khoá ngoại của bảng). Đồng thời, các thuộc tính của quan hệ này sẽ trở thành các cột trong bảng Viewer_Article.

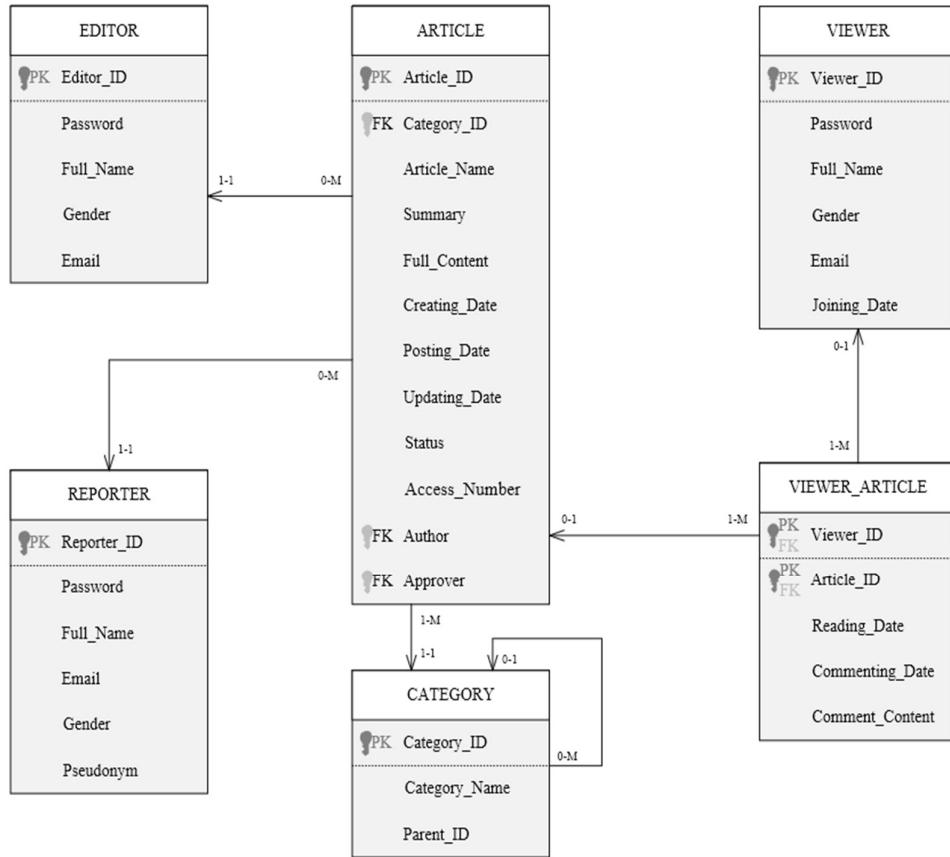
2.2.4. Quy tắc quan hệ tự tham chiếu

- Quan hệ tự tham chiếu “Contains” trong bảng Category: Không thay đổi cấu trúc của bảng Category.

2.2.5. Các ràng buộc trong bảng

- Bảng Viewer:
 - Viewer_ID: Ràng buộc PRIMARY KEY.
 - Email: Ràng buộc UNIQUE và NOT NULL, vì Email không được trùng với Email đã tồn tại. Cũng có thể coi đây là một khoá ứng viên để xác định người đọc.
 - Password, Full_Name: Ràng buộc NOT NULL cho các thông tin bắt buộc cho người đọc.
 - Gender: Ràng buộc NOT NULL và CHECK giá trị nhập vào chỉ cho phép “Male” hoặc “Female”.
 - Joining_Date: Ràng buộc NOT NULL và DEFAULT lấy giá trị ngày hiện tại theo hệ thống.
- Bảng Editor:
 - Editor_ID: Ràng buộc PRIMARY KEY.
 - Email: Ràng buộc UNIQUE và NOT NULL, vì Email không được trùng với Email đã tồn tại. Cũng có thể coi đây là một khoá ứng viên để xác định biên tập viên.
 - Password, Full_Name: Ràng buộc NOT NULL cho các thông tin bắt buộc cho biên tập viên.
 - Gender: Ràng buộc NOT NULL và CHECK giá trị nhập vào chỉ cho phép “Male” hoặc “Female”.

- Bảng Reporter:
 - Reporter_ID: Ràng buộc PRIMARY KEY.
 - Email: Ràng buộc UNIQUE và NOT NULL, vì Email không được trùng với Email đã tồn tại. Cũng có thể coi đây là một khoá ứng viên để xác định người viết báo (Phóng viên).
 - Password, Full_Name: Ràng buộc NOT NULL cho các thông tin bắt buộc cho phóng viên.
 - Gender: Ràng buộc NOT NULL và CHECK giá trị nhập vào chỉ cho phép “Male” hoặc “Female”.
 - Pseudonym: Không cần ràng buộc, có thể NULL.
- Bảng Article:
 - Article_ID: Ràng buộc PRIMARY KEY.
 - Category_ID: Ràng buộc FOREIGN KEY, tham chiếu từ bảng Category.
 - Article_Name: Ràng buộc UNIQUE và NOT NULL, tránh trường hợp trùng tên giữa các bài báo. Cũng có thể coi Article_Name là một khoá ứng viên để xác định bài báo.
 - Summary, Full_Content, Creating_Date, Access_Number: Ràng buộc NOT NULL cho các thông tin bắt buộc của mỗi bài báo.
 - Author: Ràng buộc FOREIGN KEY, tham chiếu từ bảng Reporter.
 - Approver: Ràng buộc FOREIGN KEY, tham chiếu từ bảng Editor.
 - Posting_Date: Không cần ràng buộc, có thể NULL.
 - Updating_Date: Không cần ràng buộc, có thể NULL.
 - Status: Ràng buộc NOT NULL và CHECK giá trị nhập vào chỉ cho phép “Approved” hoặc “Not Approved”.
- Bảng Category:
 - Category_ID: Ràng buộc PRIMARY KEY.
 - Category_Name: Ràng buộc NOT NULL cho tên chủ đề.
 - Parent_ID: Không cần ràng buộc, có thể NULL.
- Bảng Viewer_Article:
 - Viewer_ID, Article_ID: Ràng buộc PRIMARY KEY.
 - Viewer_ID: Ràng buộc FOREIGN KEY.
 - Article_ID: Ràng buộc FOREIGN KEY.
 - Reading_Date: Ràng buộc NOT NULL, vì khi người đọc truy cập vào trang báo thì ngày đọc sẽ được ghi lại.
 - Commenting_Date và Comment_Content: Không cần ràng buộc, có thể NULL.



Hình 2.2. Lược đồ cơ sở dữ liệu quan hệ chuyên từ ERD của trang báo điện tử

3. Phụ thuộc hàm và chuẩn hoá các bảng dữ liệu chuẩn BCNF

3.1. Phụ thuộc hàm

- Bảng Viewer:
 - Viewer_ID → {Password, Full_Name, Email, Gender, Joining_Date}.
- Bảng Editor:
 - Editor_ID → {Password, Full_Name, Email, Gender}.
- Bảng Reporter:
 - Reporter_ID → {Editor_ID, Password, Full_Name, Gender, Email, Pseudonym}.
- Bảng Article:
 - Article_ID → {Category_ID, Article_Name, Summary, Full_Content, Creating_Date, Posting_Date, Updating_Date, Status, Access_Number, Author, Approver}.
- Bảng Category: Category_ID → {Category_Name, Parent_ID}.
- Bảng Viewer_Article: {Viewer_ID, Article_ID} → {Reading_Date, Commenting_Date, Comment_Content}.

3.2. Chuẩn hoá các bảng dữ liệu theo chuẩn BCNF

- Đối với 1NF: Các bảng dữ liệu đảm bảo mỗi một ô trong bảng chỉ có giá trị đơn nhất. Các bảng không tồn tại các thuộc tính tính toán và mỗi dòng đều duy nhất. Vì vậy, dữ liệu thỏa điều kiện của 1NF.

- Đối với 2NF: Bảng Viewer_Article có khoá chính là kết hợp giữa 2 cột Viewer_ID và Article_ID, tuy nhiên, tất cả các cột không phải là khoá đều phụ thuộc đầy đủ vào khoá chính. Đồng thời, tất cả các bảng còn lại đều có khoá chính chỉ gồm một thuộc tính. Do đó, dữ liệu thoả điều kiện của 2NF.
- Đối với 3NF: Thoả mãn điều kiện 3NF vì không tồn tại phụ thuộc bắc cầu giữa các bảng dữ liệu.
- Đối với BCNF: Mỗi một định thức trong các phụ thuộc hàm đều là khoá chính hoặc là khoá ứng viên. Do đó, không tồn tại sự phụ thuộc của thuộc tính là khoá vào thuộc tính không phải là khoá. Điều kiện BCNF được thoả mãn.