

Môn học: DBI202x – Các hệ cơ sở dữ liệu

Bài tập số: 1

Tên bài tập: Thiết kế cơ sở dữ liệu

1.	<i>Yêu cầu</i>	1
2.	<i>Xây dựng mô hình ERD</i>	1
3.	<i>Lược đồ cơ sở dữ liệu</i>	3
4.	<i>Tổng kết</i>	6

1. Yêu cầu

CSDL nhằm phục vụ một trang báo điện tử; cụ thể, lưu dữ liệu của các bài báo, người dùng và các tương tác của người dùng với bài báo

Về dữ liệu bài báo, cần có số định danh bài báo, tên bài báo, tóm tắt nội dung, nội dung đầy đủ, ngày đăng, người đăng, ngày duyệt, người duyệt, chủ đề, trạng thái của bài đăng (chờ duyệt, đã duyệt, đã ẩn, đã xóa) tài nguyên được sử dụng trong bài báo (ảnh hoặc video hoặc audio).

Cụ thể, một bài báo có thể sử dụng nhiều tài nguyên, một tài nguyên có thể sử dụng trong nhiều bài báo. Dữ liệu về tài nguyên chứa số định danh tài nguyên, loại tài nguyên, đường dẫn thư mục chứa tài nguyên. Ngoài ra, vì bài báo trải qua nhiều giai đoạn để được thể hiện nên phải lưu lịch sử trạng thái của bài báo.

Về dữ liệu người dùng, cần có số định danh người dùng, tên đăng nhập, password, thời gian hiệu lực bắt đầu của người dùng, ngày hết hiệu lực của người dùng (nếu còn đang active thì coi như ngày hết hiệu lực là vô hạn), các quyền của người dùng.

Cụ thể, các quyền của người dùng bao gồm *xem, yêu cầu đăng, phê duyệt yêu cầu đăng, quản lý người dùng, quản lý người viết*. Một người dùng có thể có nhiều quyền, ít nhất phải có quyền xem. Các quyền của người dùng có thể thay đổi theo thời gian (ban đầu được khởi tạo, sau đó có thể thêm hoặc bớt bởi người dùng có quyền “quản lý người dùng”); do đó, phải lưu lịch sử hiệu lực của các quyền của người dùng.

Về tương tác liên quan bài báo, cần có số định danh tương tác, bài báo được tương tác, người dùng thực hiện tương tác, thời điểm thực hiện tương tác, loại tương tác, đặc tính của tương tác.

Cụ thể các loại tương tác bao gồm: **truy cập bài báo**, *tìm kiếm, gửi yêu cầu đăng bài, duyệt yêu cầu đăng bài, từ chối yêu cầu đăng bài*. Tùy vào loại tương tác, cần lưu trữ các đặc tính khác nhau, chẳng hạn đối với tương tác truy cập cần chứa dữ liệu điểm truy cập trước đó (từ trang chủ hay từ search engine), đối với tương tác tìm kiếm thì lưu dữ liệu về keyword cần tìm kiếm, hay đối với tương tác từ chối yêu cầu sẽ lưu dữ liệu về lý do từ chối,

...

2. Xây dựng mô hình ERD

Từ mô tả yêu cầu về dữ liệu của bài báo, cần có thực thể và mối quan hệ sau:

- Thực thể **Articles**: chứa dữ liệu về bài báo với các thuộc tính sau: articleID (số định danh bài báo), articleName (tên bài báo), articleSum(tóm tắt nội dung), articleContent (nội dung đầy đủ), datePost (ngày đăng), dateApprove (ngày duyệt), articleApprove (người duyệt), articleAuthor(người đăng), articleCat (chủ đề), status (trạng thái)
- Thực thể **Resources**: chứa dữ liệu về các hình ảnh/video/audio được sử dụng trong bài báo với các thuộc tính: resourceID (số định danh của đa phương tiện), resourceCat (loại đa phương tiện), resourceFolder (đường dẫn thư mục chứa đa phương tiện)
- Thực thể **ArticleHist**: là thực thể yếu, phụ thuộc định danh vào thực thể Articles, chứa lịch sử trạng thái của bài báo, với các thuộc tính: version (số thứ tự trạng thái), status (trạng thái), statusDate (ngày được sang trạng thái)
- Quan hệ **Contains**: giữa thực thể Articles và Resources, cardinality: many – many; optional – mandatory (một bài báo có thể không sử dụng đa phương tiện, nhưng 1 đa phương tiện có trong database phải có ít nhất được sử dụng bởi 1 bài báo).

- Quan hệ **records ArticleStatus Changes**: giữa thực thể Articles và ArticleHist, đây là quan hệ định danh.

Từ mô tả về dữ liệu người dùng, cần có các thực thể và mối quan hệ sau:

- Thực thể **Users**: chứa dữ liệu người dùng, chứa các thuộc tính: userID (số định danh người dùng), userName (tên đăng nhập), password (mật khẩu), effDate (ngày hiệu lực bắt đầu của người dùng), inEffDate (ngày hết hiệu lực)
- Thực thể **Roles**: chứa dữ liệu liên quan đến các quyền có thể có của người dùng, chứa các thuộc tính: roleID (số định danh quyền), roleName (tên quyền), effDate (ngày hiệu lực bắt đầu), inEffDate (ngày kết thúc hiệu lực)
- Quan hệ **Is Granted**: giữa thực thể Users và Roles, cardinality: many – many, mandatory – optional; ngoài ra quan hệ này còn có các thuộc tính effDate (ngày hiệu lực bắt đầu của quyền đối với 1 người dùng), inEffDate (ngày hiệu lực kết thúc)

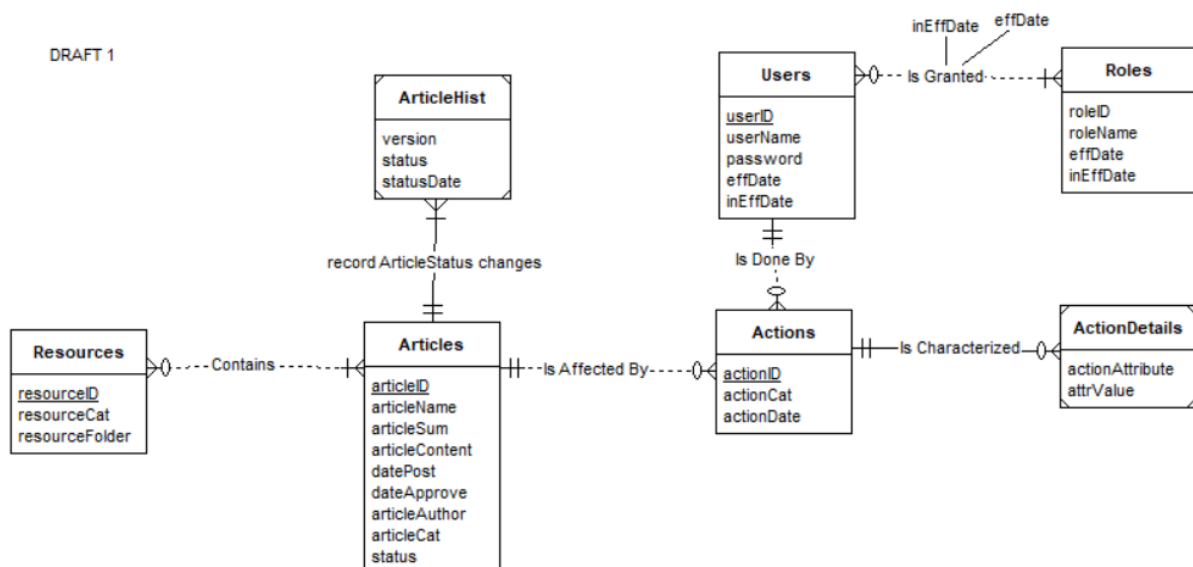
Từ mô tả về tương tác, cần có các thực thể và mối quan hệ sau:

- Thực thể **Actions**: chứa dữ liệu về tương tác đối với các bài báo, chứa các thuộc tính: actionID (số định danh tương tác), actionCat (loại tương tác), actionDate (ngày thực hiện)
- Thực thể **ActionDetails**: là thực thể yếu, phụ thuộc định danh vào thực thể Actions, chứa dữ liệu chi tiết về các tương tác, chứa các thuộc tính: actionAttribute (tên thuộc tính liên quan đến tương tác), attrValue (giá trị)
- Quan hệ **Is Characterized**: là quan hệ định danh giữa thực thể Actions và ActionDetails.

Ngoài ra bổ sung các quan hệ:

- Quan hệ **is Affected**: là quan hệ giữa thực thể Articles và Actions, cardinality: one – many, mandatory – optional (một bài báo có thể có hoặc không tương tác, nhưng 1 tương tác chỉ thuộc về 1 bài báo duy nhất)
- Quan hệ **is Done By**: là quan hệ giữa thực thể Actions và Users, cardinality: one – many, mandatory – optional (một tương tác được thực hiện bởi 1 người dùng duy nhất, và người dùng có thể không hoặc có nhiều tương tác)

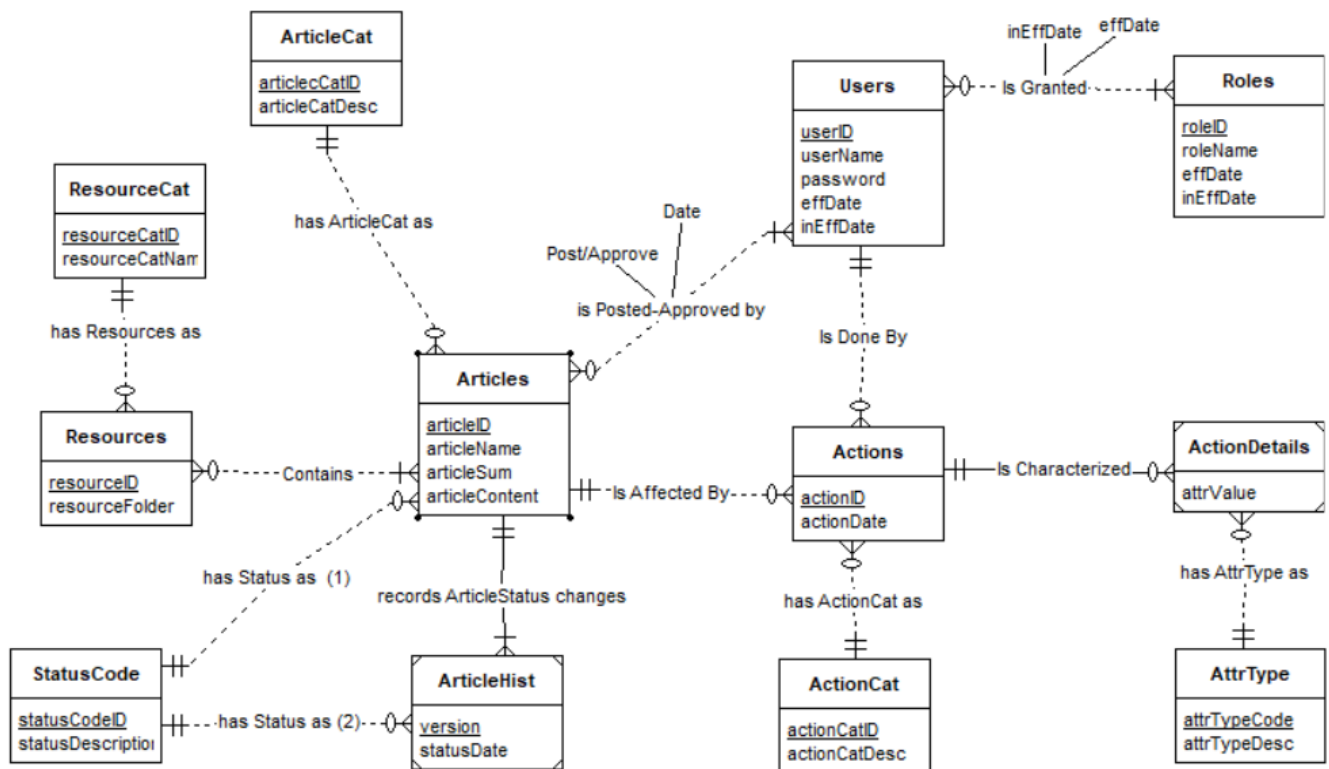
➔ Sơ lược ERD như bên dưới:



Một số biến đổi và bổ sung:

- Thực thể Articles chứa một số thuộc tính như articleAuthor, articleApprove liên quan đến thực thể Users -> bỏ các thuộc tính datePost, dateApprove, articleAuthor, articleApprove; thêm quan hệ Is Posted-Approved By (có thuộc tính PostApprove – cờ báo là người dùng thực hiện đăng bài hay duyệt bài, Date – ngày đăng hoặc ngày duyệt bài)
- Bổ sung các thực thể articleCat, resourceCat, statusCode, actionCat, actionAttribute và các quan hệ cần thiết để tránh lỗi dị thường sửa đổi khi chuyển đổi ở các bước tiếp theo

→ Kết luận ERD:



3. Lược đồ cơ sở dữ liệu

Từ ERD được thiết kế ở phần 2,

- Áp dụng quy tắc loại thực thể, CSDL sẽ có các bảng và cột như sau:

ResourceCat		
PK	<u>resourceCatID</u>	int
	resourceCatName	varchar

ArticleCat		
PK	<u>articleCatID</u>	int
	articleCatDesc	varchar

Users		
PK	<u>userID</u>	int
	userName	varchar
	password	varchar
	effDate	date
	inEffDate	date

Roles		
PK	<u>roleID</u>	int
	roleName	varchar
	effDate	date
	inEffDate	date

Resources		
PK	<u>resourceID</u>	int
	resourceFolder	varchar

Articles		
PK	<u>articleID</u>	int
	articleName	varchar
	articleSum	varchar
	articleContent	varchar

Actions		
PK	<u>actionID</u>	int
	actionDate	varchar
	articleSum	varchar
	articleContent	varchar

ActionDetails		
	attrValue	varchar

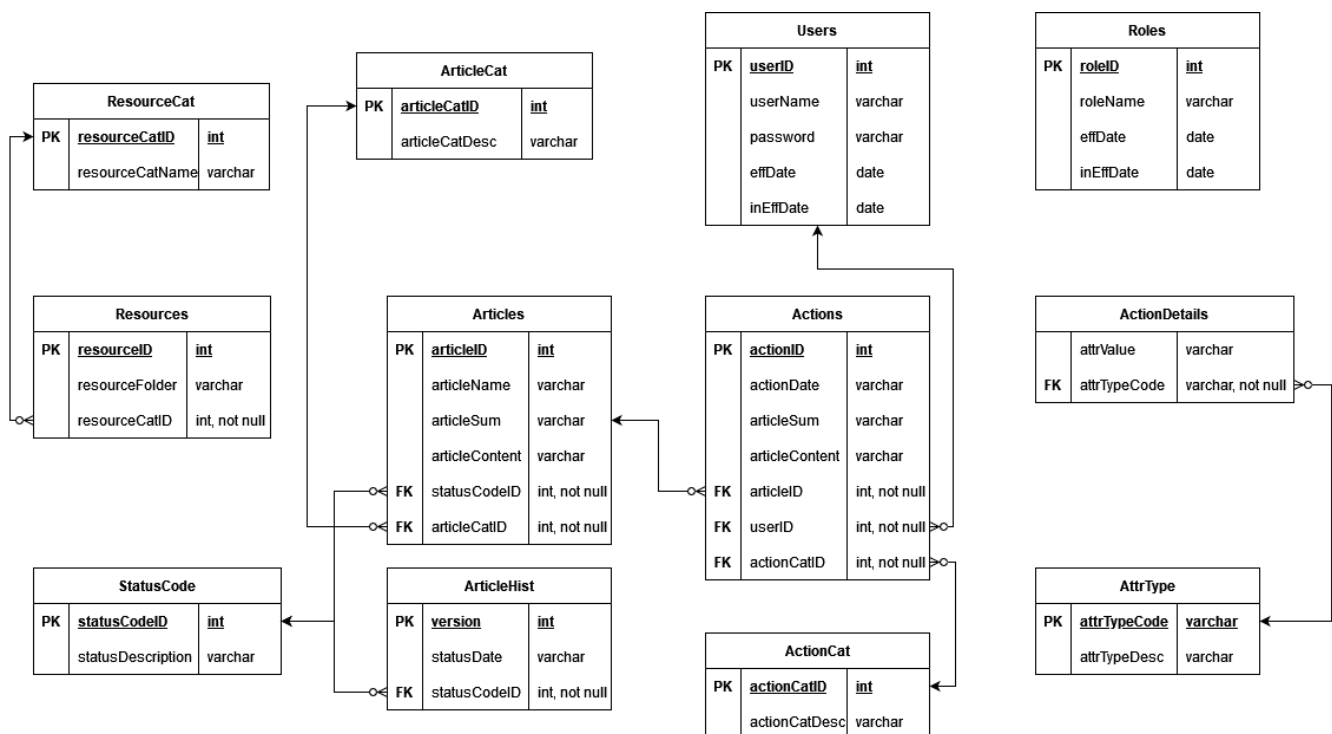
StatusCode		
PK	<u>statusCodeID</u>	int
	statusDescription	varchar

ArticleHist		
PK	<u>version</u>	int
	statusDate	varchar

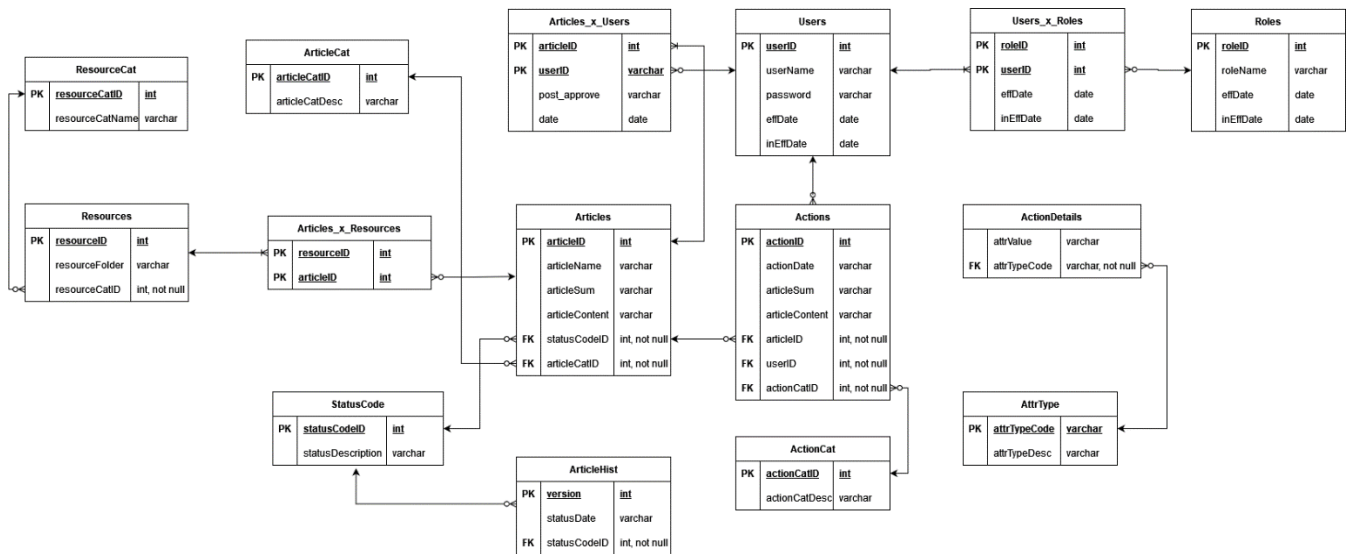
ActionCat		
PK	<u>actionCatID</u>	int
	actionCatDesc	varchar

AttrType		
PK	<u>attrTypeCode</u>	varchar
	attrTypeDesc	varchar

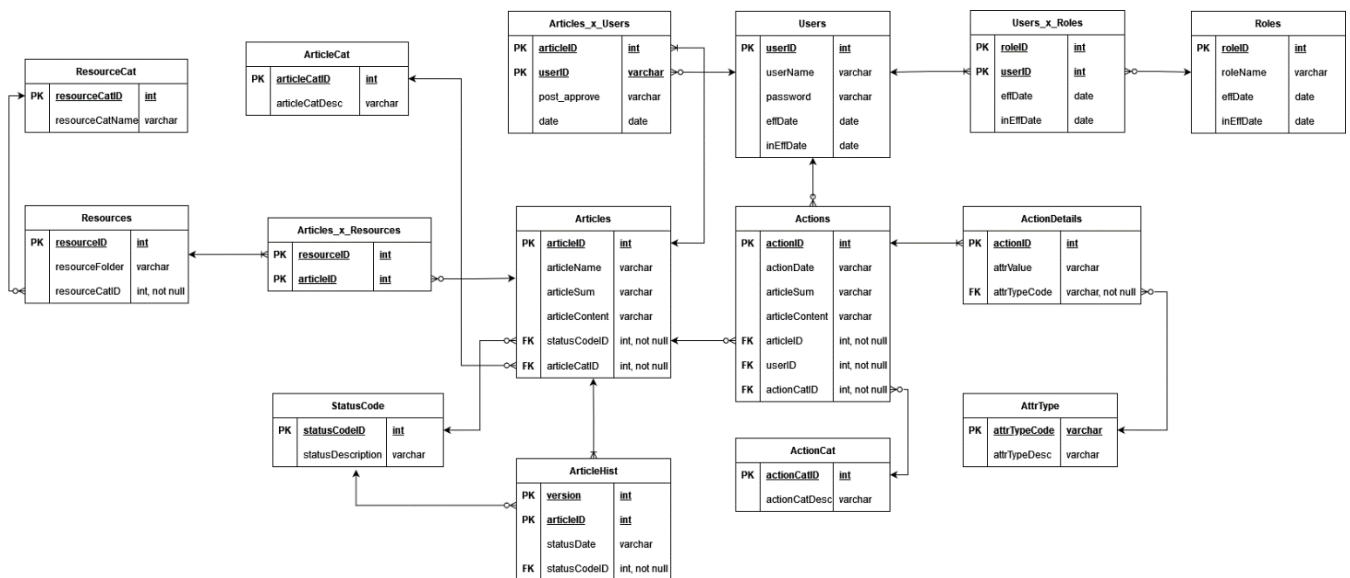
- Áp dụng quy tắc quan hệ one-to-many: bổ sung khoá ngoại vào các bảng con và ràng buộc not null cho các cột khoá ngoại này



- Áp dụng quy tắc quan hệ many-to-many: tạo các bảng với khoá chính kết hợp



- Áp dụng quy tắc quan hệ định danh: thêm khoá chính cho các bảng được chuyển đổi từ các loại thực thể yếu



- Xác định các phụ thuộc hàm (FD) và biến đổi về dạng chuẩn: chia các bảng thành 3 nhóm
 - o Nhóm 1: các bảng được biến đổi từ các thực thể bình thường; khoá chính của mỗi bảng xác định các cột khác trong bảng => thoả chuẩn BCNF (không cần liệt kê FD)
 - o Nhóm 2: các bảng được biến đổi từ các quan hệ many-many, cụ thể:
 - Articles_x_Resources: chỉ có 2 cột và khoá chính => thoả BCNF
 - Article_x_Users: khoá chính là articleID, userID; tuy nhiên FD là **articleID, post_approve** → **userID, date** => cần đặt khoá chính kết hợp articleID, post_approve, chuyển userID thành khoá ngoại
 - Users_x_Roles: khoá chính kết hợp roleID, userID; FD là **roleID, userID** -> **effDate, inEffDate** => thoả BCNF
 - o Nhóm 3: các bảng biến đổi từ các thực thể yếu, phụ thuộc định danh, cụ thể

- ArticleHist: khoá chính kết hợp version, articleID; FD là version, articleID -> statusDate, statusCodeID => thoả BCNF
- ActionDetails: khoá chính actionID, FD là actionID, attrTypeCode -> attrValue => cần đặt khoá chính kết hợp actionID, attrTypeCode

4. Tổng kết

Lược đồ CSDL cuối cùng

