ONLINE NEWSROOM DATABASE DESIGN

XÂY DỰNG DATABASE TÒA SOẠN BÁO ĐIỆN TỬ

1. **Phân tích các yêu cầu chức năng và dữ liệu của trang báo điện tử mong muốn thực hiện.**

CSDL Trang báo điện tử được thiết kế để lưu trữ dữ liệu về nội dung về các bài đăng, người xem, phóng viên viết bài và biên tập viên phê duyệt.

**CSDL gồm các đối tượng chính:** Bài báo chưa được phê duyệt(RawContent), Nội dung đã được phê duyệt(ConfirmedContent) ,Phóng viên viết bài(ContentWriter), Quản lí phê duyệt(Manager)

**Liệt kê các thuộc tính trong các thực thể:**

**Content Writer:** Mã phóng viên (WriterID) **PK** , Tên phóng viên (WriterName), Ngày sinh (DateOfBirth), Liên hệ (WriterContact)

**RawContent:** Mã bài viết (ContentID)(**PK)**, Tên bài viết (ContentName), Chủ đề (ContentTopic), Tóm tắt nội dung (ContentSummary), Nội dung đầy đủ(ContentFull)

**Manager:** Mã quản lí phê duyệt (ManagerID) **(PK)** , Tên quản lí phê duyệt (ManagerName), Liên hệ (ManagerContact)

**ConfirmedContent** lưu những bài báo đã được Manager phê duyệt từ bài báo thô ( RawContent). ConfirmedContent là thực thể yếu sẽ lấy cặp khoá chính ContentID( của RawContent) và ManagerID (trong Manager) là khoá định danh. Gồm các thuộc tính:

Tên bài báo đã được quản lí xác nhận (ConfirmName), Nội dung bài báo đã được chỉnh sửa và xác nhận (ConfirmContent), Ngày xác nhận (ConfirmDate), Tình trạng xác nhận hay bác bỏ bài báo (ConfirmStatus)

**Các quan hệ giữa các thực thể:**

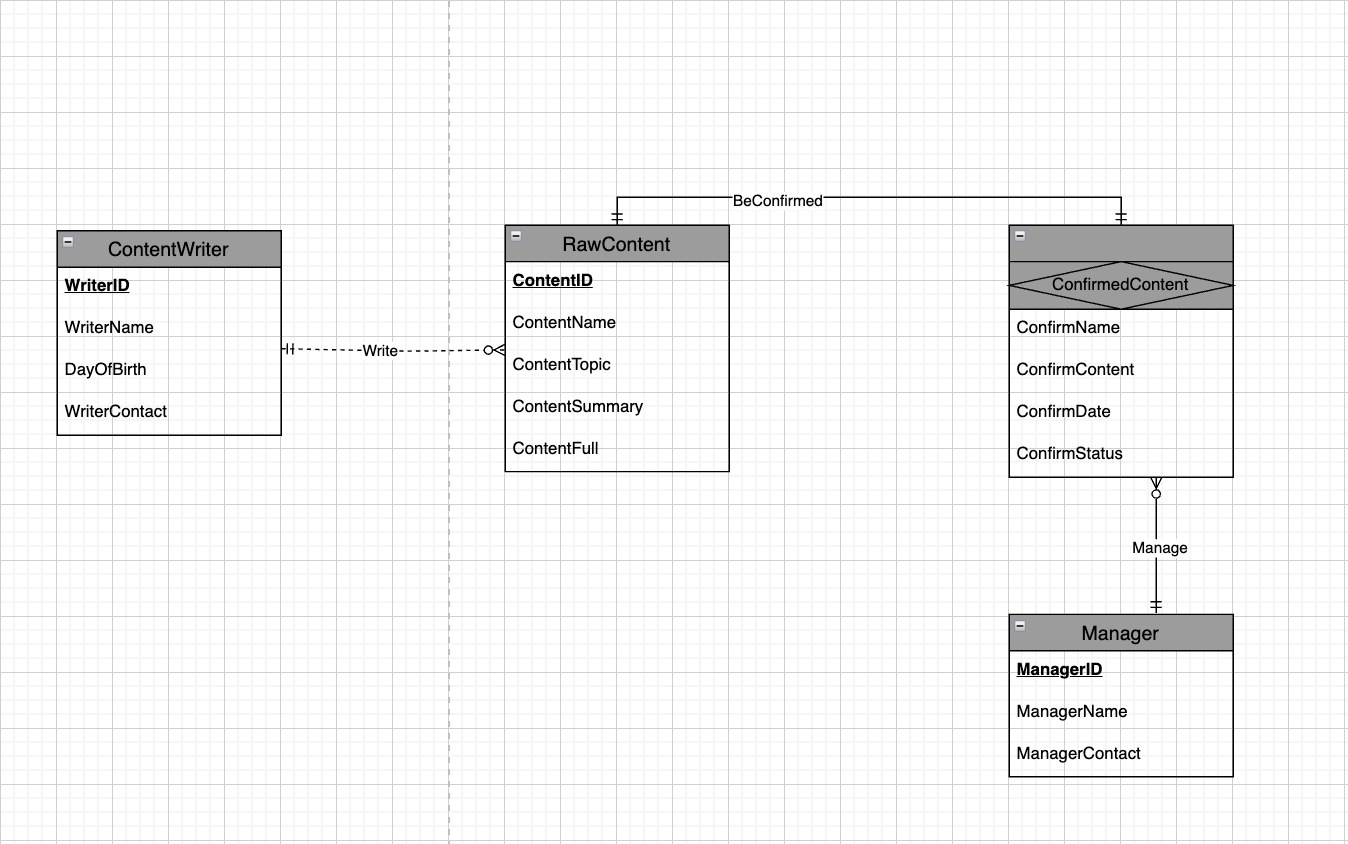
ContentWriter sẽ viết bài báo thô được lưu trữ và chờ phê duyệt được lưu trong bảng RawContent. Một ContentWriter viết được nhiều bài báo, một bài báo chỉ được một và chỉ một người viết.

RawContent sau khi được Manager kiểm duyệt sẽ được lưu trong ConfirmedContent. Một RawContent chỉ được một ConfirmedContent( Với trường hợp bài viết được thông qua hay không sẽ thể hiện ở cột ConfirmStatus, không ảnh hưởng tới khoá chính ContentID)**.** Một ConfirmedContent chỉ được kiểm duyệt từ một RawContent.

Manager lưu thông tin người kiểm duyệt bài báo ConfirmedContent. Một Manager được duyệt không hoặc nhiều CornfirmedContent. Một ConfirmedContent bắt buộc chỉ có một và chỉ một người chịu trách nhiệm duyệt.

1. **Xây dựng mô hình ERD thể hiện lại các yêu cầu trên**

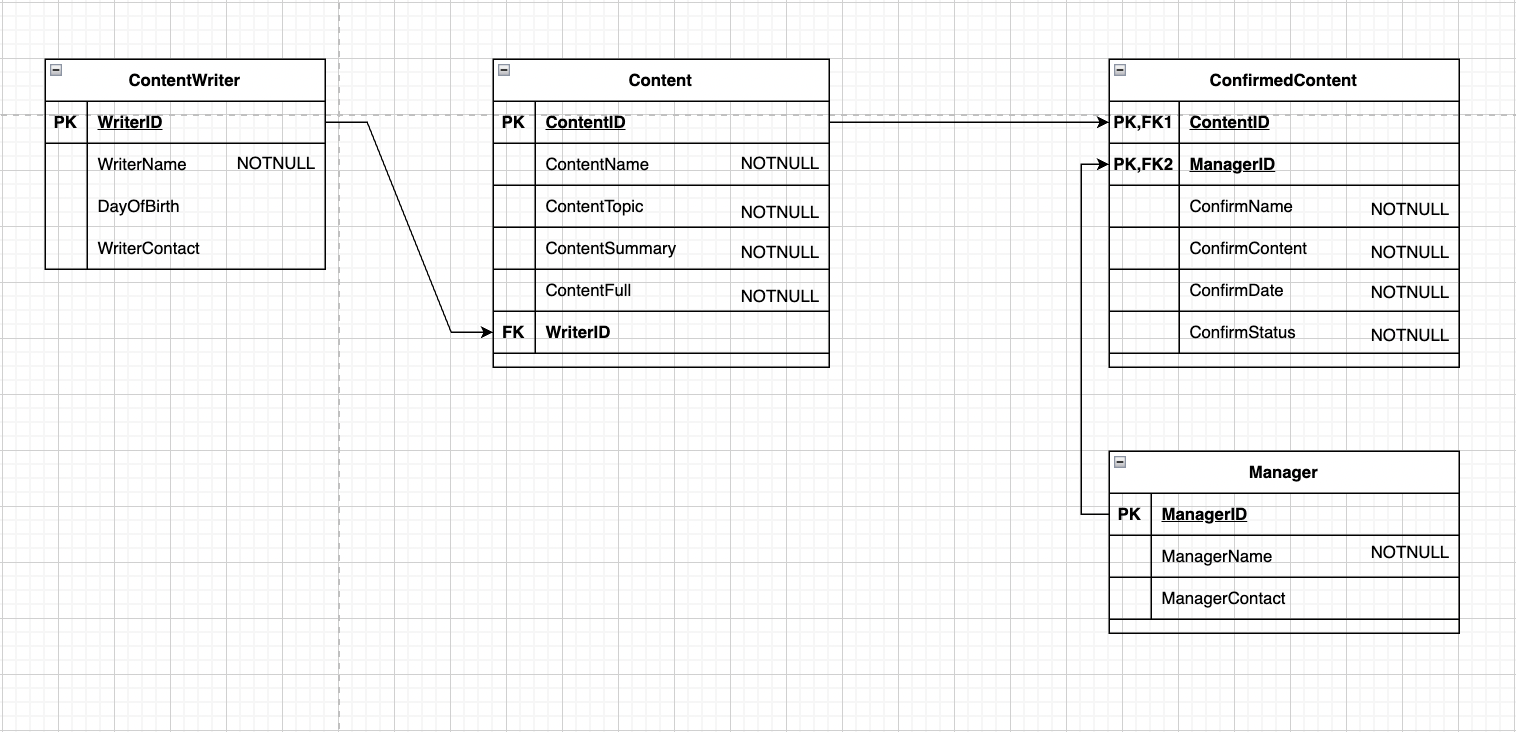
* Đầu tiên ta mô tả các thực thể mạnh bằng các khối vuông có tên thực thể nằm trên cùng. Trong các khối vuông lần lượt là khoá chính được in đậm gạch chân và các thuộc tính của thực thể.
* Thực thể yếu ConfirmedContent (không có khoá chính để định danh) sẽ phụ thuộc định danh bằng cặp khoá chính ContentID ( của bảng RawContent) và ManagerID ( của bảng Manager). Do đó thực thể này sẽ có kí hiệu gạch chéo các góc khối vuông để nhận biết.
* Sau khi vẽ được mô hình các thực thể, ta sẽ đi đặt tên cho quan hệ theo cách dễ hiểu. Sau đó ta sẽ xác định các đường quan hệ giữa các thực thể ( nét đứt là đường nối giữa các thực thể mạnh, nét liền thể hiện mối quan hệ định danh).
* Tiếp theo, ta dựa vào quan hệ giữa các thực thể mô tả ở trên để xác định Cardinality tối thiểu và tối đa giữa các thực thể. Sau đó mô tả lại bằng kí hiệu Crow’s Foot phù hợp.
* Sau khi thực hiện các bước trên ta có ERD hoàn thiện ở bên dưới.



1. **Lược đồ CSDL tương ứng**

* Ta áp dụng các quy tắc đã học để chuyển đổi từ ERD sang mô hình CSDL quan hệ.
* Theo quy tắc loại thực thể ( Entity Type), các thực thể mạnh sẽ trở thành các bảng, có các khoá chính được xác định là thuộc tính có gạch chân, các thuộc tính sẹ trở thành các cột trong bảng.
* Theo quy tắc 1- Many , biến các quan hệ 1-M thành khoá ngoại. Ta có ContentID ( của bảng ContentWriter) sẽ trở thành khoá ngoại của bảng Content. Tương tự, ManagerID là khoá ngoại FK của bảng ConfirmedContent.
* Áp dụng quy tắc phụ thuộc định danh, thực thể yếu ConfirmedContent trở thành bảng mang khoá chính kết hợp của hai bảng định danh.

Sau đó, ta xác định các ràng buộc NOTNULL cho các cột quan trọng không được để trống của các bảng.



1. **Kiểm tra chuẩn hoá CSDL về dạng 3NF**

3NF là dạng gần tiệm cận với BCNF yêu cầu mọi định thức phải là duy nhất trong một bảng. Để chứng minh CSDL đã được chuẩn hoá về dạng BCNF, ta xác định được trong một bảng, có chỉ có một định thức của phụ thuộc hàm.

Đầu tiên ta xác định các phụ thuộc hàm của mỗi bảng.

Ở bảng ContentWriter, ta thấy chỉ có một định thức duy nhất (WriterID) với các hàm phụ thuộc như sau:

WriterID →WriterName,DayOfBirth,WriterContact

Tương tự với các bảng tiếp theo:

ContentID→ContentName, ContentTopic,ContentSummary, ContentFull

ManagerID → ManagerID, ManagerContact

ContentID,ManagerID→ConfirmName, ConfirmContent, ConfirmDate, ConfirmStatus

Ta có thể thấy với các định thức ContentID,ManagerID, và cặp khoá chính kết hợp ContentID.ManagerID, không có trường hợp sai lệch phụ thuộc hàm xảy ra. Hay nói cách khác, không có trường hợp các giá trị của định thức trùng nhau mà các giá trị của hàm phụ thuộc khác nhau

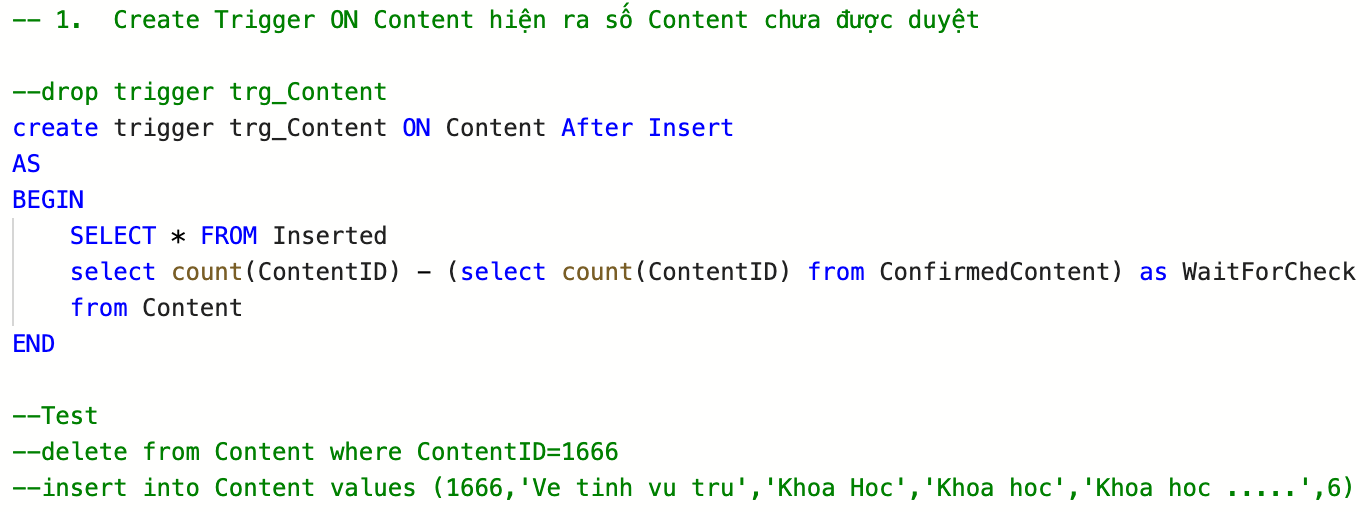
Do đó có thể nói mô hình CSDL quan hệ ở trên đã về dạng chuẩn hoá BCNF.

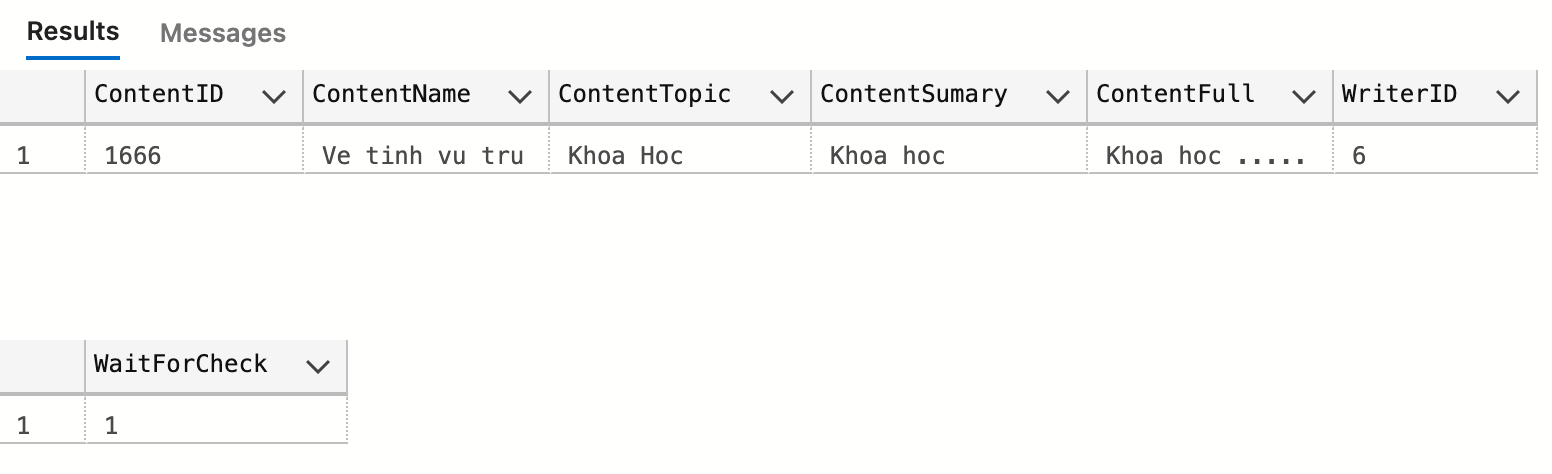
**1. Tạo 1 Trigger.**

Khi Insert một Record vào bảng Content thì:

- Hiện các Values đã được Inserted

- Hiện thông báo số lượng các Content đang chờ để được phê duyệt từ Manager.





**2. Tạo một Stored Procedure**

Kiểm tra Content đã được kiểm duyệt hay chưa ?

Truyền biến (@ContentID) cho Stored Procedure SP\_CheckContent

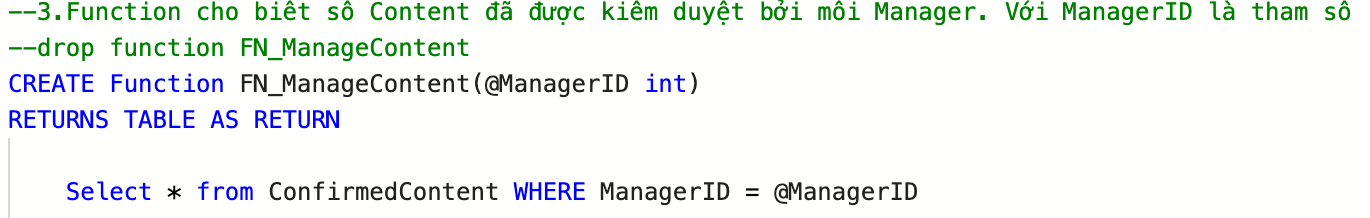
- Sử dụng IF Exist để kiểm tra ContentID có tồn tại? Nếu không trả về “Mã ContentID không tồn tại”

- Trả về ContentID từ 2 bảng Content LEFT JOIN ConfirmedContent



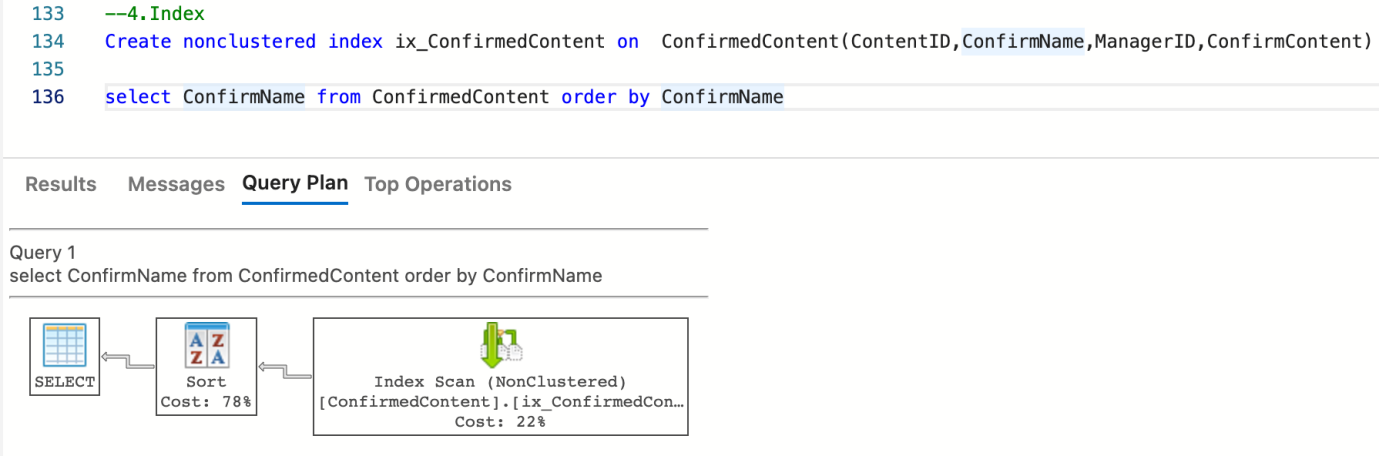
1. **Tạo một Function**

Function cho biết số Content đã được kiểm duyệt bởi mỗi Manager. Với ManagerID là tham số



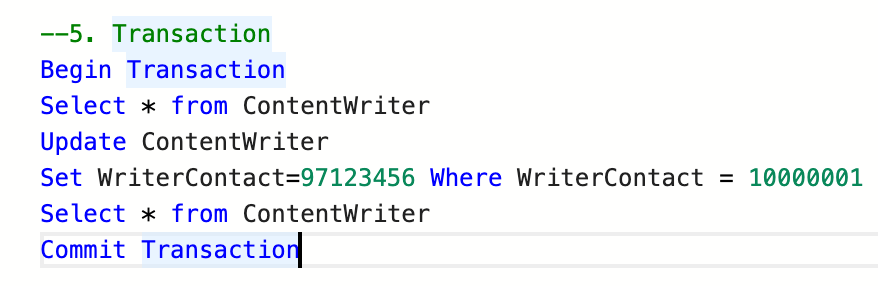
1. **Tạo một Index**

Tạo NonClusterded Index cho các cột như hình bên dưới bảng ConfirmedContent. Câu lệnh trên sẽ tạo một bảng Index độc lập với bảng dữ liệu,giúp truy vấn dữ liệu với trường không có Clustered Index nhanh hơn.



1. **Tạo một Transaction**

Transaction thực hiện Update Contact của Writer thoã điều kiện Where.



**Truy vấn trên CSDL**

**1. Truy vấn dữ liệu trên một bảng**

Tìm các bài viết liên quan tới chủ đề Thể Thao.



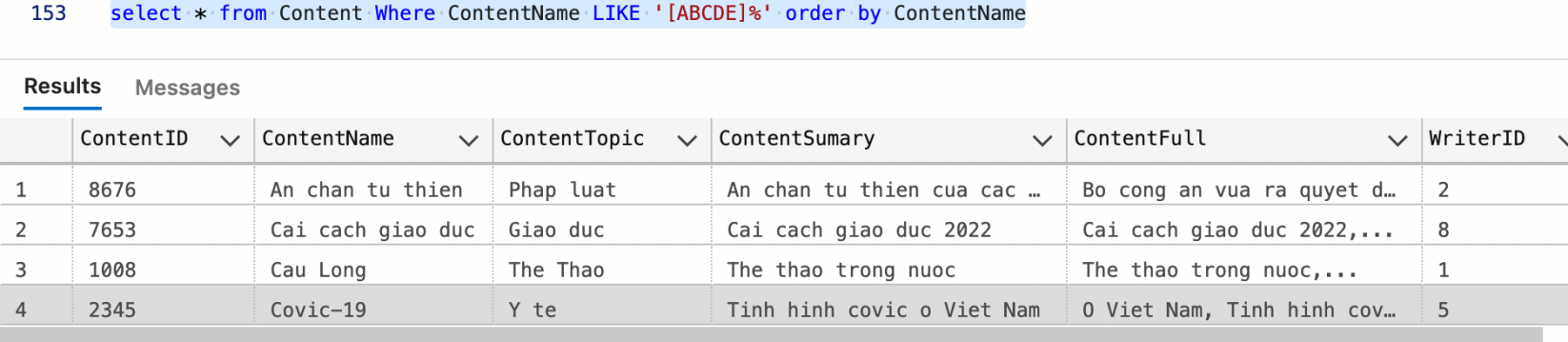
**2.- Truy vấn có sử dụng Order by**

Sắp xếp danh sách Manager trong bảng Manager theo tên.



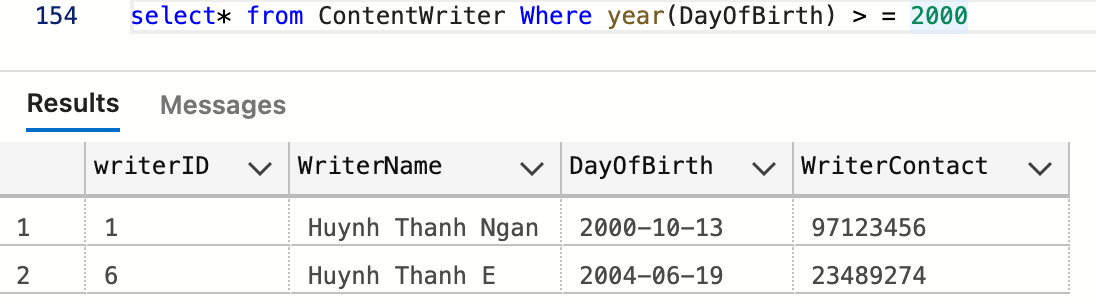
1. **Truy vấn sử dụng toán tử Like và các so sánh xâu ký tự.**

Tìm tên bài viết (ContentName) trong bảng Content bắt đầu bằng những chữ cái ABCDE



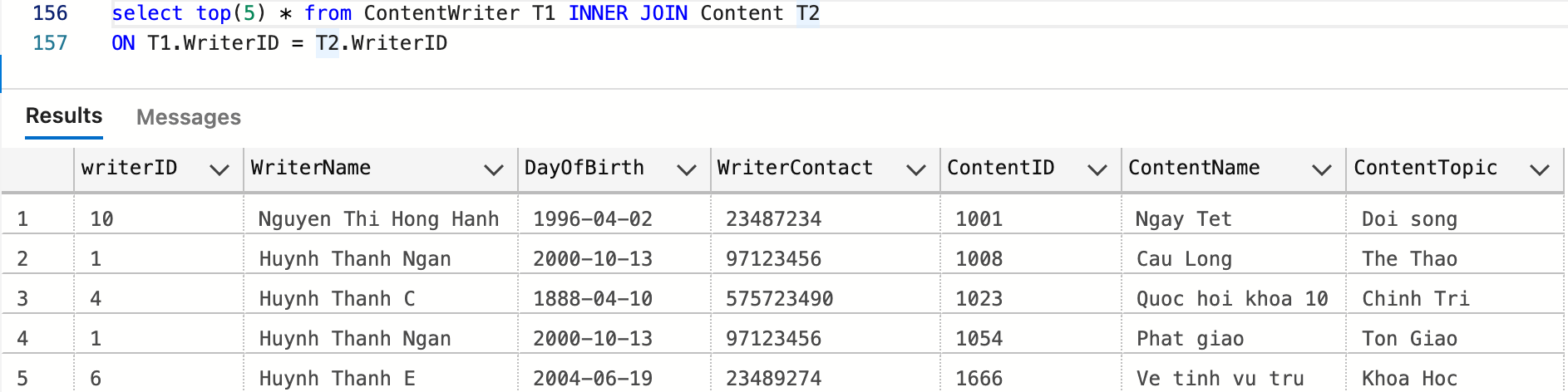
**4.Truy vấn liên quan tới điều kiện về thời gian**

Truy vấn thông tin người viết báo ( ContentWriter ) có năm sinh lớn hơn 2000



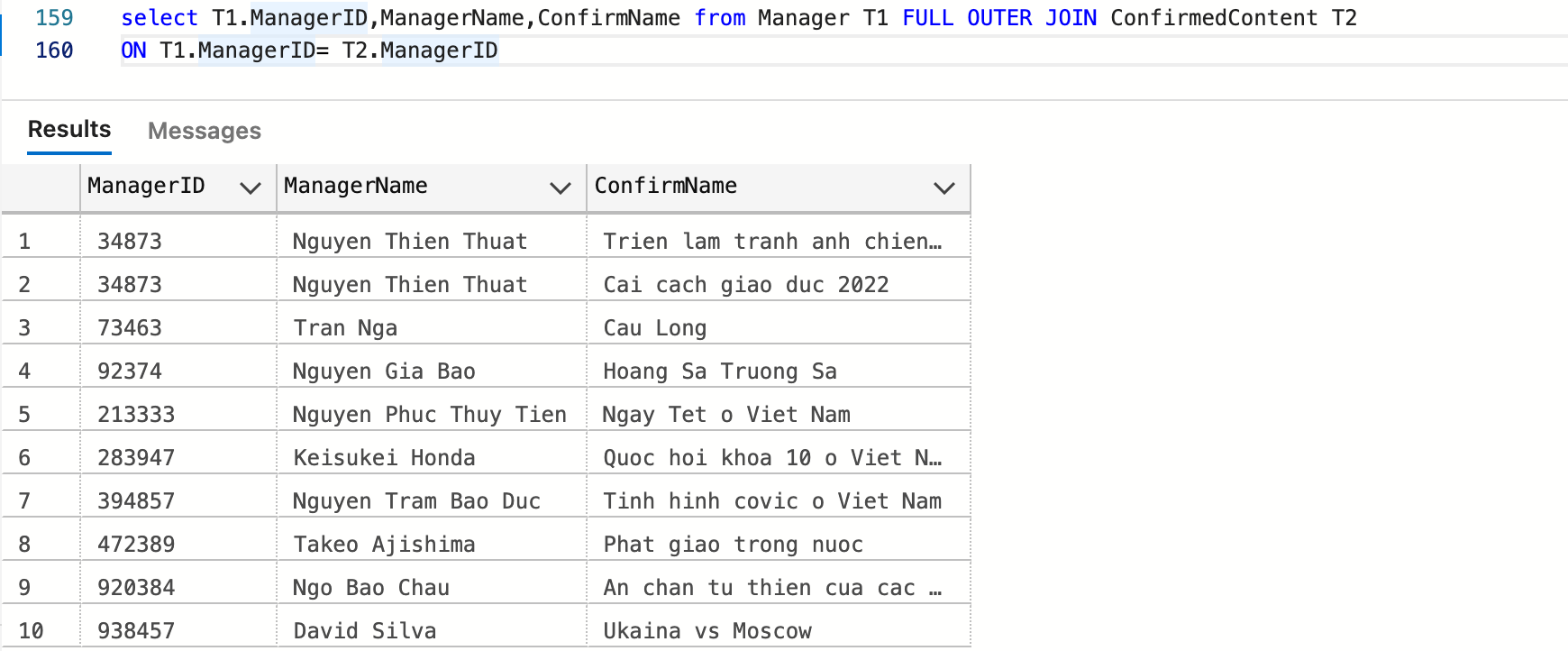
1. **Truy vấn dữ liệu từ nhiều bảng sử dụng Inner join**

Truy vấn top 5 thông tin của Người viết bài Và nội dung bài viết của họ

****

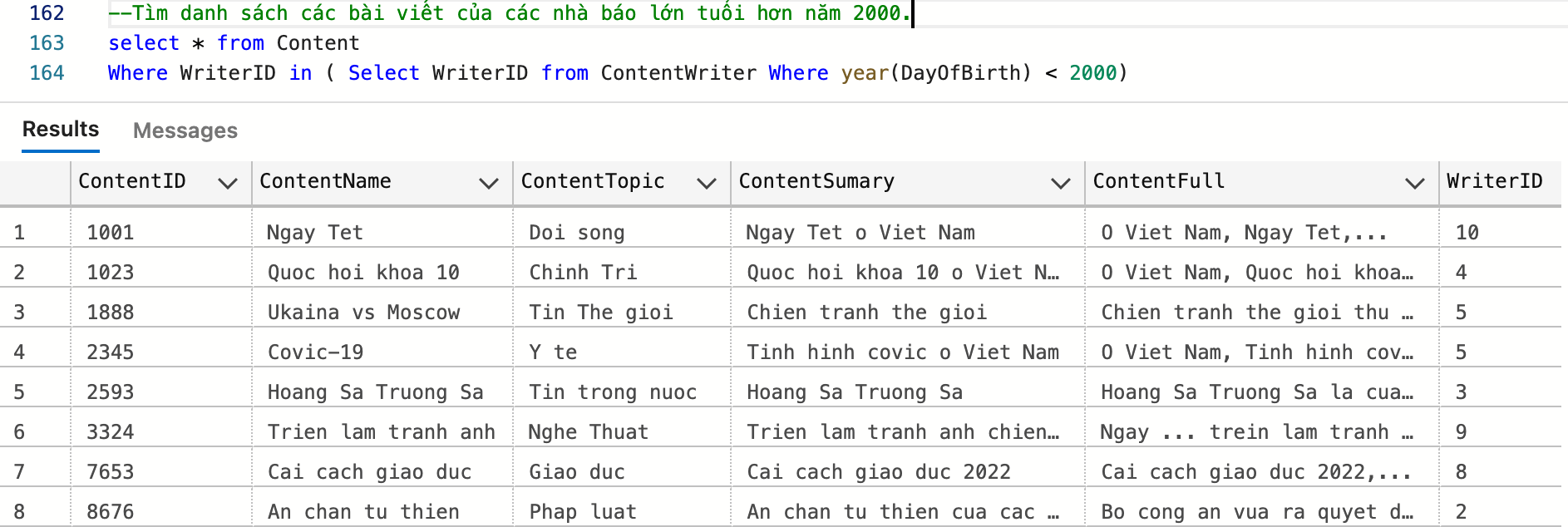
1. **Truy vấn sử dụng Full Outer Join**

Tìm danh sách Tên các Manager kiểm duyệt các ConfirmedContent. Sử dụng FULL OUTER JOIN giữa 2 bảng Manager FULL OUTER JOIN ConfirmedContent

****

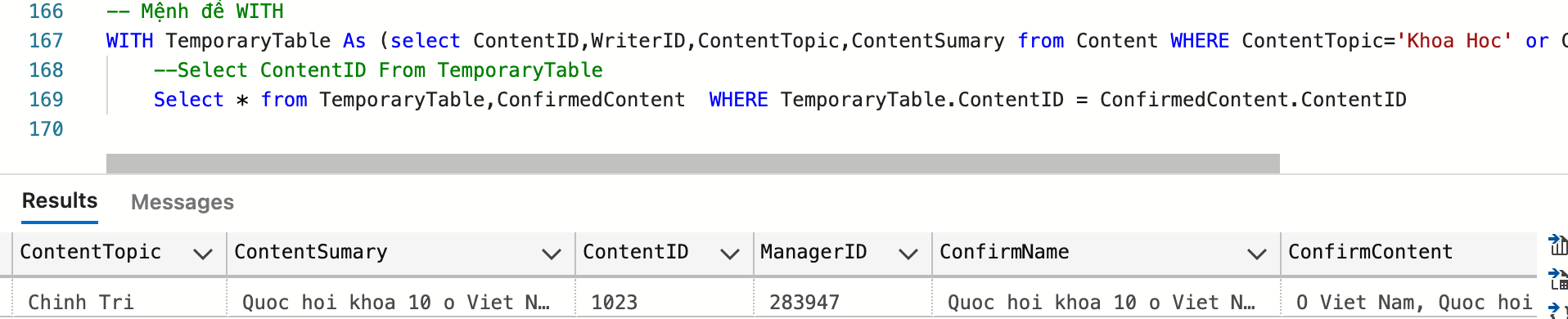
**7.Truy vấn sử dụng truy vấn con.**

Tìm danh sách các bài viết của các nhà báo lớn tuổi hơn năm 2000.



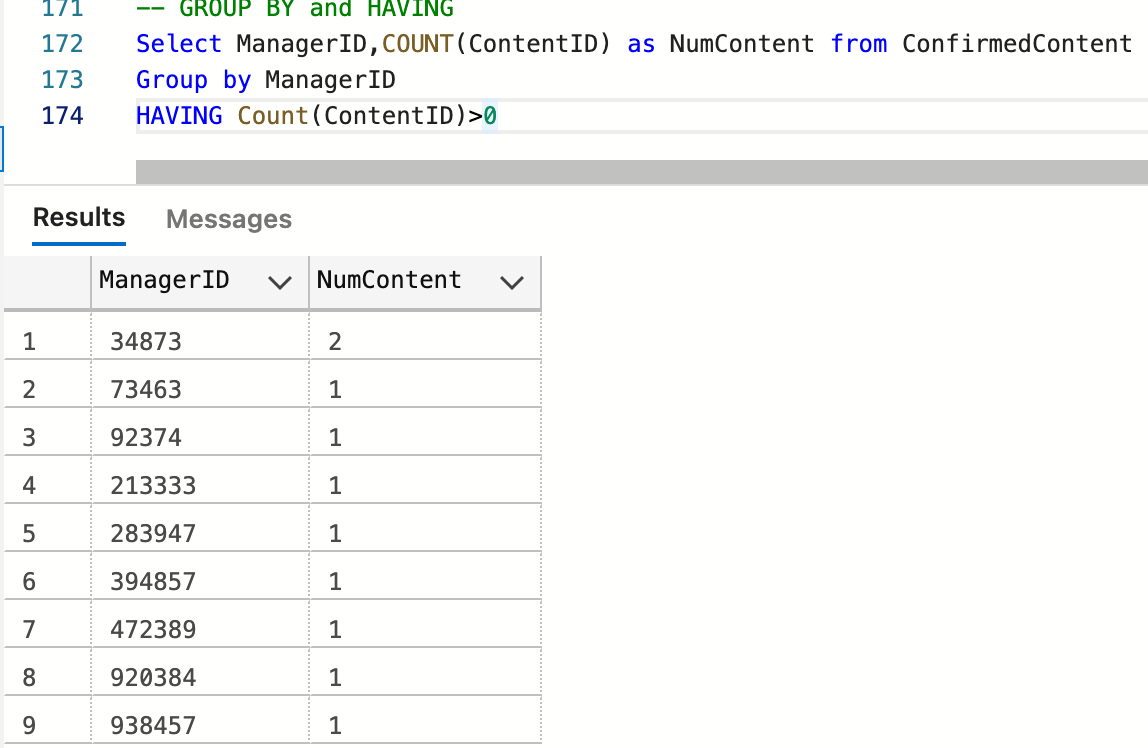
**8.Truy vấn sử dụng With.**

Truy vấn nội dung của bài viết liên quan tới chủ đề Chính trị và Khoa học lưu vào bảng phụ Temporary bằng mệnh đề WITH. Sau đó JOIN với bảng ConfirmedContent để tìm nội dung đã chỉnh sửa thuộc những Topic cần tìm. (Vì bảng ConfirmedContent không có cột Topic nên phải dùng JOIN để truy vấn)

****

**9.Truy vấn thống kê sử dụng Group by và Having**

Tìm số lượng bài viết đã được kiểm duyệt của từng Manager với điều kiện HAVING là Num(ContentID) > 0

****

1. **Truy vấn sử dụng function (hàm) đã viết trong bước trước.**

