



UEH500

UEH  
UNIVERSITY

# XÂY DỰNG KHO DỮ LIỆU CHO TRƯỜNG ĐẠI HỌC TỪ ĐÓ ĐỀ XUẤT CÁC CHIẾN LƯỢC GIÚP TĂNG THỨ HẠNG TRÊN BẢNG XẾP HẠNG U-MULTIRANK

Công trình dự thi  
**BIT34**

Thái Sơn

Mỹ Anh

Yến My

Phương Duyên

Quang Tiến

# BỐI CẢNH

- ✓ **Thế giới:** Quá trình toàn cầu hóa và hội nhập quốc tế đang diễn ra rất mạnh mẽ.
- ✓ **Việt Nam:** Bối cảnh giáo dục ĐH Việt Nam hội nhập ngày càng sâu rộng với hệ thống giáo dục ĐH trong khu vực và trên thế giới. Kết quả xếp hạng cao mang đến danh tiếng và uy tín cho trường đại học, quốc gia và thu hút nhân tài, đầu tư và hợp tác

# LÍ DO

- ✓ Dữ liệu cực kỳ quan trọng với chúng ta
- ✓ Bùng nổ dữ liệu dẫn đến bài toán lưu trữ và phân tích
- ✓ Các trường đại học cần có giải pháp lưu trữ tốt dữ liệu và khai thác được câu chuyện mà chúng mang lại

# MỤC TIÊU

- ✓ Phân tích dữ liệu theo nhiều góc nhìn
- ✓ Đưa ra những giải pháp chiến lược
- ✓ Hỗ trợ quá trình ra quyết định
- ✓ Cải thiện vị thế của trường Đại học trên những BXH (cụ thể là BXH U - Multirank)



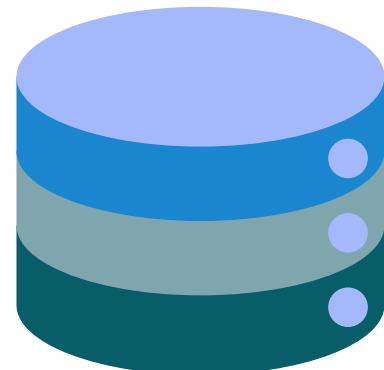
# Kiến trúc kho dữ liệu cho Bảng xếp hạng U-Multirank

## Tổng quan về BXH U-Multirank

Mục tiêu chính của dự án là cung cấp một công cụ minh bạch và cách tiếp cận đa chiều để đánh giá và so sánh hiệu suất của các cơ sở giáo dục đại học và nghiên cứu trên toàn cầu.

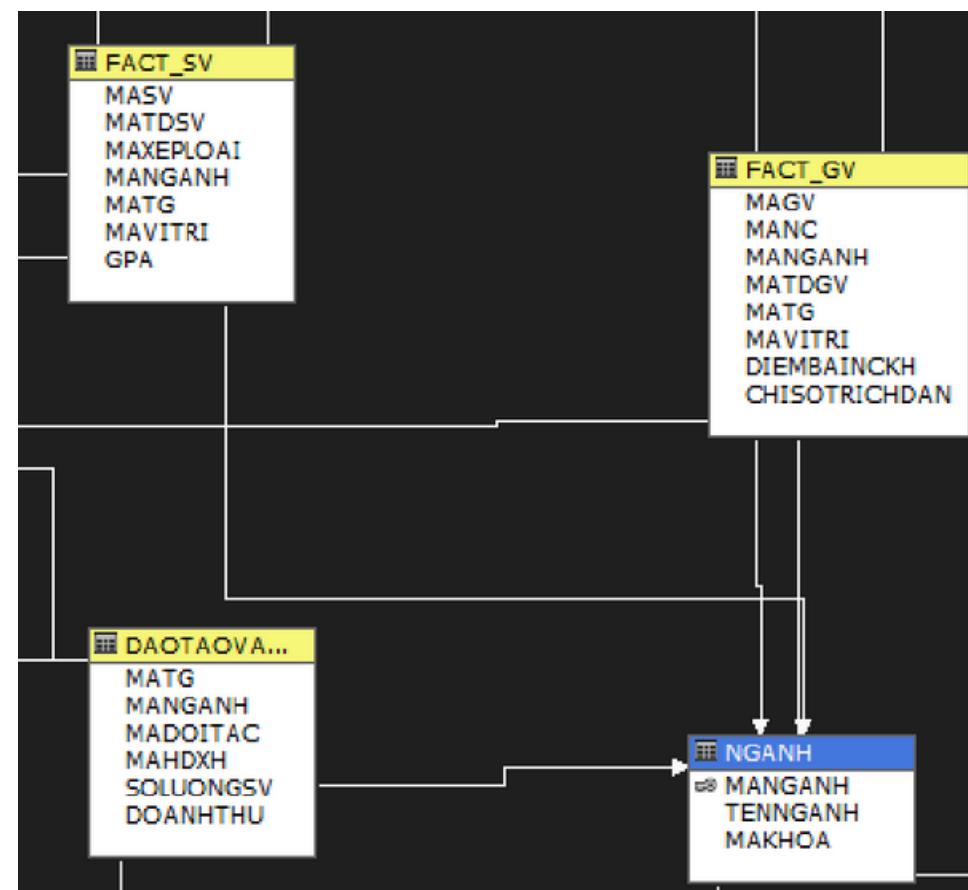
U-Multirank so sánh kết quả hoạt động của các trường đại học theo năm khía cạnh, bao gồm:

- 01 Giảng dạy và học tập
- 02 Nghiên cứu
- 03 Chuyển giao tri thức
- 04 Định hướng quốc tế
- 05 Liên kết nghiên cứu khu vực



# Kiến trúc kho dữ liệu cho Bảng xếp hạng U-Multirank

**1** Hướng chủ đề  
(Subject-oriented)



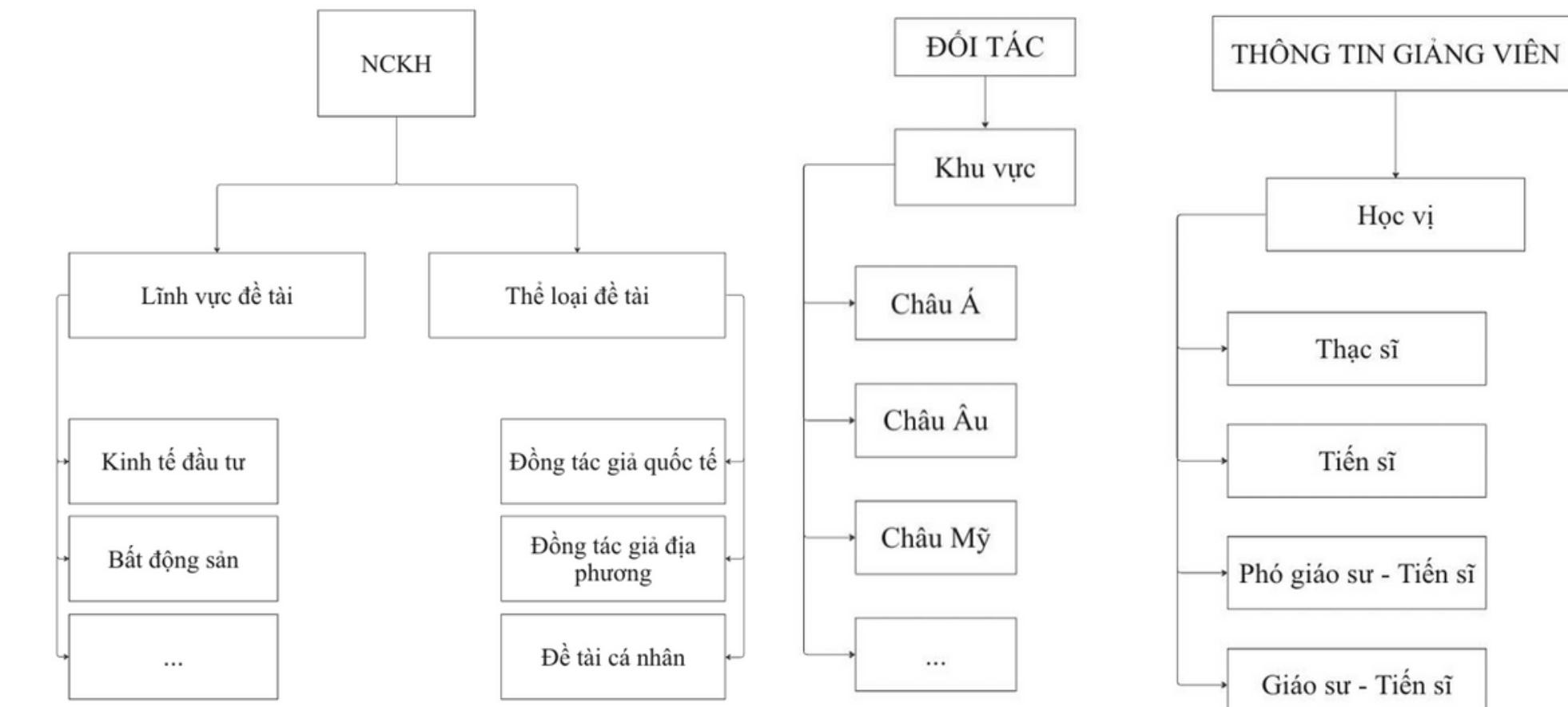
**2** Tính tích hợp  
(Integrated data)

**3** Dữ liệu gắn thời gian  
(Time-variant data)

## Tính chất kho dữ liệu

**4** Dữ liệu không biến động  
(Non Volatile Data)

**5** Độ mịn và chi tiết dữ liệu  
(Data granularity)

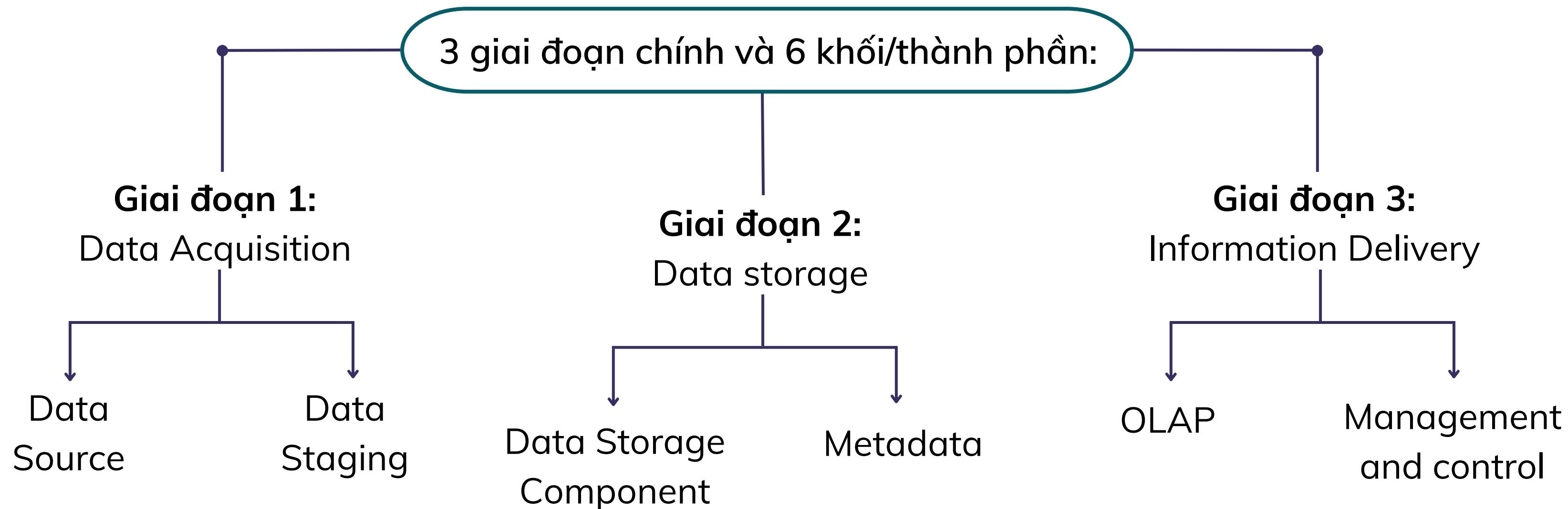




# Kiến trúc kho dữ liệu cho Bảng xếp hạng U- Multirank



## Kiến trúc kho dữ liệu



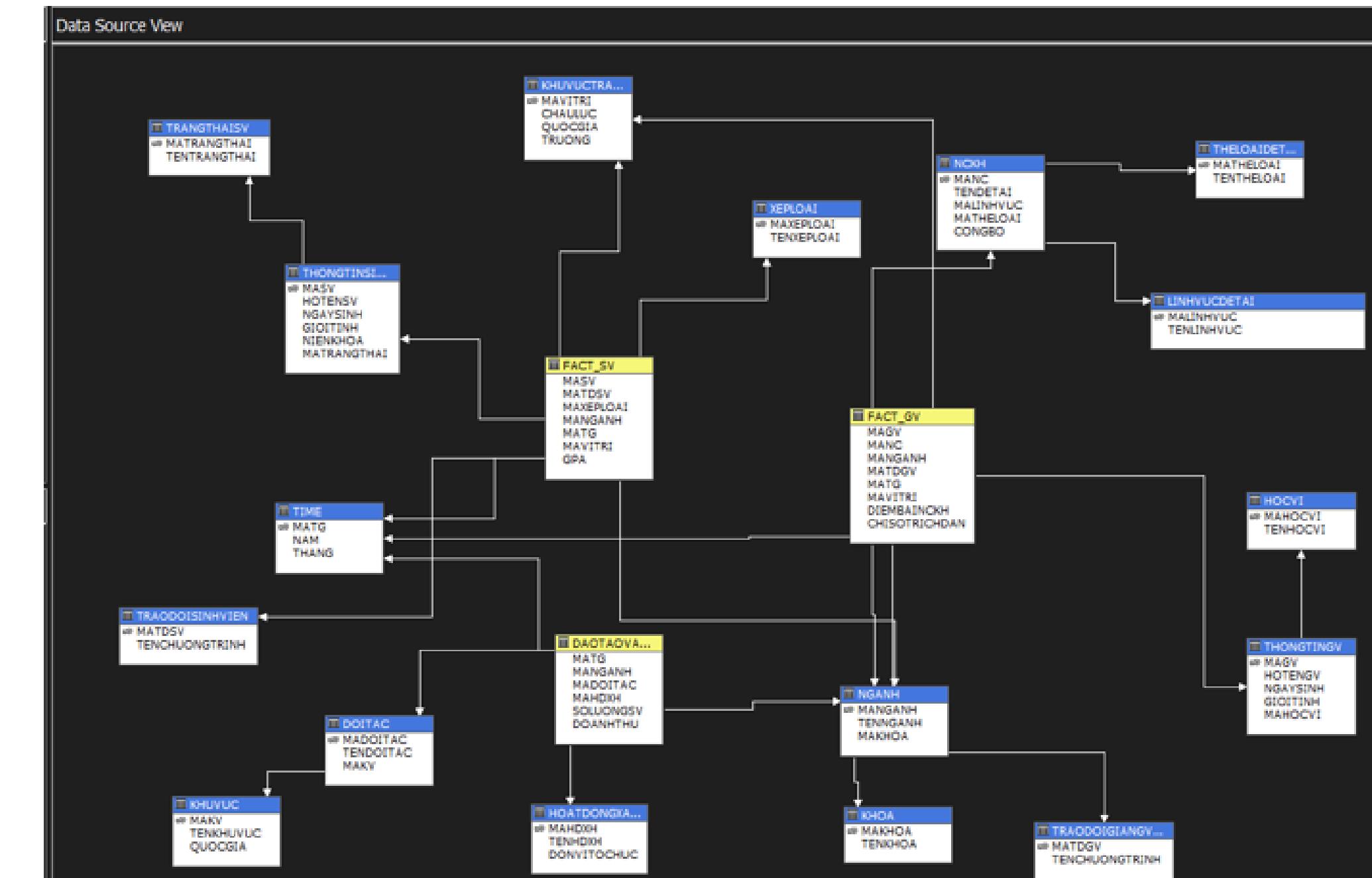


# Kiến trúc kho dữ liệu cho Bảng xếp hạng U-Multirank



- Lược đồ hình chòm sao - Galaxy Schema
- Lược đồ Galaxy được xem như là bước lặp và kết hợp từ 2 dạng lược đồ quan hệ Hình sao (Star) và Bông tuyết (Snowflake)

## Thiết kế CSDL cho kho dữ liệu





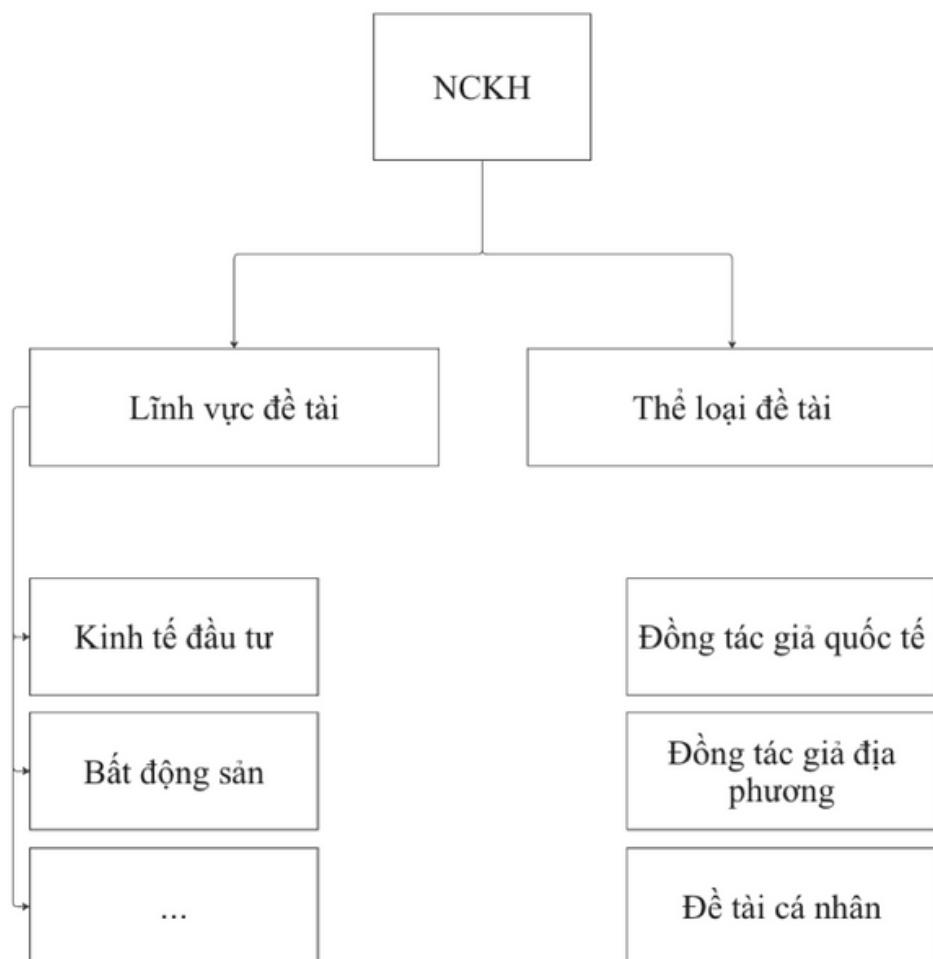
# Kiến trúc kho dữ liệu cho Bảng xếp hạng U- Multirank



Tại sao là Kho dữ liệu mà không phải là  
Cơ sở dữ liệu ?

✓ Sự phân cấp

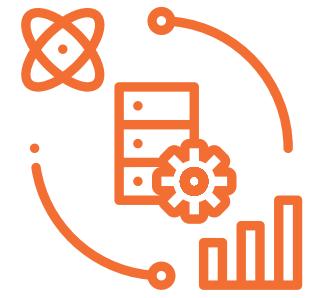
✓ Kho dữ liệu phục vụ cho việc phân  
tích và đề xuất chiến lược



Tiêu chí	CSDL	Kho DL
<b>Mục đích</b>	Ghi lại và lưu trữ	Phân tích DL
<b>Phương pháp xử lý</b>	Xử lý giao dịch trực tuyến (OLTP)	Xử lý phân tích trực tuyến (OLAP).
<b>Nhiệm vụ</b>	Giúp thực hiện các hoạt động cơ bản cho doanh nghiệp	Cho phép phân tích tình hình doanh nghiệp và ra quyết định CL
<b>Hướng</b>	Hướng ứng dụng	Hướng chủ đề
<b>Phạm vi</b>	Trong một ứng dụng	Tập hợp nhiều ứng dụng
<b>Cập nhật DL</b>	Thời gian thực	Định kỳ
<b>Thao tác</b>	Đọc, xóa, cập nhật	Chỉ đọc



# Kiến trúc kho dữ liệu cho Bảng xếp hạng U-Multirank



Tại sao là Kho dữ liệu mà không phải là Cơ sở dữ liệu ?

- ✓ Hiệu quả trong quá trình truy vấn

## MDX

```
SELECT [Measures].[FACT SV Count]* {[TIME].[THANG].&[3],[TIME].[THANG].&[9]} ON COLUMNS,  
FILTER(NONEMPTY  
(  
{[NGANH].[TENKHOA].children*[NGANH].[TENNGANH].children}  
*{[THONGTINSINHVIEN].[MATRANGTHAI].&[DTN]}  
*[XEPMLOAI].[MAXEPMLOAI].&[XLG] )  
, [Measures].[FACT SV Count] >= 10)  
ON ROWS  
FROM [UNISERSITY DATAWAREHOUSE]  
WHERE [TIME].[NAM].&[2020]
```

			FACT SV Count	FACT SV Count
Kinh doanh quốc tế - Marketing	Logistics và Quản lý Chuỗi cung ứng	DTN	XLG	3
Kinh doanh quốc tế - Marketing	Ngoại thương	DTN	XLG	5
Luật	Luật kinh tế	DTN	XLG	(null)
Quản trị	Quản trị	DTN	XLG	7
Quản trị	Quản trị kinh doanh	DTN	XLG	15
Tài chính	Tài Chính	DTN	XLG	5
				9
				10
				10
				8
				(null)
				12

## SQL

```
SELECT TENKHOA, TENNGANH , TRANGTHAI , XEPLOAI , COUNT( TENNGANH) AS SO_LUONG_TOT_NGHIEP  
FROM KHOA K JOIN NGANH N ON K.MAKHOA = N.MAKHOA  
JOIN THONGTINSINHVIEN SV ON N.MANGANH = SV.MANGANH  
WHERE THOIGIAN = '2020/3' or THOIGIAN = '2020/9'  
GROUP BY TENNGANH , TRANGTHAI , XEPLOAI ,TENKHOA HAVING COUNT( TENNGANH) >= 10
```

TENKHOA	TENNGANH	TRANGTHAI	XEPLOAI	SO_LUONG_TOT_NGHIEP
1 Kinh doanh quốc tế - Marketing	Logistics và Quản lý Chuỗi cung ứng	DTN	XLG	15
2 Luật	Luật kinh tế	DTN	XLG	10
3 Kinh doanh quốc tế - Marketing	Ngoại thương	DTN	XLG	15
4 Quản trị	Quản trị	DTN	XLG	15
5 Quản trị	Quản trị kinh doanh	DTN	XLG	15
6 Tài chính	Tài chính	DTN	XLG	17



# DEMO TRUY VẤN



○ Gồm 4 câu truy vấn:

## Truy vấn 1

Truy vấn số lượng **Giảng Viên** theo từng **ngành** có **học vị** là Thạc sĩ đã từng tham gia chương trình **trao đổi** quốc tế theo từng **Châu lục** và có bài **NCKH** được công bố trong **năm** 2022, 2023

## Truy vấn 2

Truy vấn số lượng **Sinh Viên** đã tốt nghiệp **loại** Giỏi hoặc Xuất Sắc theo từng **Khoa** đã từng tham gia chương trình **trao đổi** ở các **châu lục** trong **năm** 2022

## Truy vấn 3

Truy vấn số lượng **HĐXH** đã được tổ chức vào **năm** 2022, 2023 theo từng **Khoa** được phối hợp với **đối tác** ngoài Châu Á.

## Truy vấn 4

Truy vấn những **Giảng viên** có **học vị** là tiến sĩ có bài **NCKH** theo từng lĩnh vực thuộc từng **khoa** có tổng số tỷ lệ trích dẫn  $\geq 60$  của **năm** 2021 và năm 2022



# KẾT LUẬN VÀ ĐỀ XUẤT

THANK YOU  
VERY MUCH!



## ƯU ĐIỂM

- ✓ Tính cấp thiết
- ✓ Khả năng áp dụng vào thực tiễn
- ✓ Đóng góp và đề xuất tri thức



## NHƯỢC ĐIỂM

- ✓ Sự khác biệt về dữ liệu so với thực tế
- ✓ Thách thức về nguồn lực để triển khai
- ✓ Sự phụ thuộc nhất định vào 1 Bảng xếp hạng



## ĐÓNG GÓP & HƯỚNG PHÁT TRIỂN

- ✓ Giải quyết vấn đề thực tiễn
- ✓ Đề xuất mô hình quản lý giáo dục ĐH hiệu quả
- ✓ Thúc đẩy nghiên cứu trong lĩnh vực GDĐH và nhiều lĩnh vực khác