

XÂY DỰNG VÀ KHAI THÁC KHO DỮ LIỆU CƯỚC KHÁCH HÀNG TẠI CÔNG TY THÔNG TIN VIỄN THÔNG ĐIỆN LỰC

Học viên: Phạm Thị Đào

Đơn vị công tác: Công ty Thông Tin Viễn Thông Điện Lực

Email: daocntt@gmail.com

GVHD: TS. Nguyễn Trí Thành

Đơn vị công tác: Trường Đại học Công Nghệ - ĐHQG – Hà Nội

Email: ntthanh@vnu.edu.vn

Từ khóa: *Kho dữ liệu, làm sạch dữ liệu, oracle warehouse builder*

1. GIỚI THIỆU BÀI TOÁN

Trình bày lý do chọn đề tài:

- Tại Việt Nam, việc ứng dụng kho dữ liệu trong lĩnh vực viễn thông còn nhiều hạn chế (chỉ có 2/9 doanh nghiệp xây dựng kho dữ liệu)
- Tại Công ty thông tin viễn thông điện lực (EVNTelecom) có rất nhiều nguồn dữ liệu khác nhau từ các hệ thống tác nghiệp: CSDL khách hàng miền Bắc, CSDL khách hàng miền Nam, CSDL cước khách hàng, CSDL công nợ, CSDL thanh khoản.
- Tại EVNTelecom có nhiều bất cập trong quản lý thông tin và trong quy trình báo cáo lãnh đạo

Nội dung bài toán đặt ra:

- Xây dựng kho dữ liệu tích hợp từ 5 nguồn dữ liệu nói trên
- Khai thác kho dữ liệu phục vụ báo cáo lãnh đạo Công ty

2. NỘI DUNG LUẬN VĂN

- Tổng quan về kho dữ liệu
- Nghiên cứu công cụ Oracle Warehouse Builder và ứng dụng xây dựng kho dữ liệu tại EVNTelecom
- Khảo sát thực trạng tại EVNTelecom và đề xuất giải pháp xây dựng kho dữ liệu
- Thực hiện xây dựng kho dữ liệu theo các bước: khai báo dữ liệu nguồn/đích, hồ sơ hóa dữ liệu, trích chọn làm sạch dữ liệu, sinh các luật dữ liệu, thiết kế luồng ánh xạ dữ liệu, tạo kho dữ liệu chủ đề
- Khai thác kho dữ liệu phục vụ báo cáo

3. KẾT LUẬN

Kết quả đạt được:

- Xây dựng kho dữ liệu tích hợp từ 5 nguồn dữ liệu khác nhau
- Tập trung quá trình làm sạch dữ liệu, đảm bảo chất lượng dữ liệu trong kho
- Xây dựng hệ thống báo cáo

Định hướng trong tương lai:

- Khai thác kho dữ liệu ở mức cao hơn: khai phá dữ liệu, phát hiện tri thức
- Tích hợp các nguồn dữ liệu đa dạng khác: doanh thu cước internet, số liệu quản lý thiết bị, thông tin quản lý điều hành mạng...
- Tích hợp các kiểu dữ liệu khác nhau: từ file,...