

THAM LUẬN HỘI THẢO

TỰ CHỦ DỮ LIỆU TRONG BÁO CHÍ

Nhìn từ thực tiễn triển khai — Bối cảnh, Rào cản và Hướng đi

Trình bày bởi:

AMI Group · Nền tảng Neotam.vn



I. BỐI CẢNH DỮ LIỆU TRONG BÁO CHÍ HIỆN NAY

Năm lớp dữ liệu

1

Content data

Dữ liệu nội dung

Nội dung nào được sản xuất, dưới định dạng nào, theo chủ đề nào

2

Audience data

Dữ liệu khán giả

Ai đang xem, khi nào, ở đâu

3

Engagement data

Dữ liệu tương tác

Hành vi tương tác cụ thể — xem hết hay bỏ giữa chừng, bình luận, chia sẻ

4

Commercial data

Dữ liệu thương mại

Doanh thu, quảng cáo — giá trị tạo ra từ nội dung và khán giả

5

Operational data

Dữ liệu vận hành

Hệ thống đang chạy ra sao — hạ tầng, hiệu suất kỹ thuật

Lớp thứ sáu — dữ liệu niềm tin xã hội và rủi ro thông tin là dữ liệu rất quan trọng nhưng đòi hỏi cách tiếp cận riêng, nằm ngoài phạm vi chia sẻ hôm nay.



I. BỐI CẢNH DỮ LIỆU TRONG BÁO CHÍ HIỆN NAY

Dữ liệu phân tán, cục bộ

Dữ liệu trong báo chí không phải khái niệm mới



Nhiều loại dữ liệu

- Hành vi xem, nội dung, phân phối, tương tác
- Mỗi loại đo một thứ khác nhau



Nhiều công cụ

- Google Analytics,
- Nielsen
- CMS nội bộ, mạng xã hội
- Mỗi nơi báo cáo theo chuẩn riêng



Nhiều nền tảng

- Truyền hình Truyền thống
- OTT
- Báo điện tử, website
- Mạng xã hội
- Mỗi nơi một định nghĩa lượt xem



Nhiều nguồn sở hữu

- Dữ liệu bên thứ nhất
- Bên thứ hai (đối tác)
- Bên thứ ba (nhà dịch vụ nghiên cứu, đo lường)

Dữ liệu hiện diện tại nhiều nơi, đang được sử dụng với nhiều mục đích khác nhau và phân mảnh



II. DỮ LIỆU ĐANG ĐƯỢC DÙNG THẾ NÀO?

Mức độ trưởng thành rất khác nhau giữa các lớp

Lớp dữ liệu	Bộ phận chịu trách nhiệm	Mục đích sử dụng	Dự địa khai thác
1 — Nội dung	<i>Ban Biên tập / Sản xuất</i>	Quản lý kho lưu trữ, gắn metadata cơ bản — thể loại, thời lượng, ngày phát	Gắn nội dung với hiệu quả thực tế theo Content Health score để biết nội dung nào tạo ra giá trị theo dữ liệu Tái khai thác tài nguyên nội dung
2 — Khách giả	<i>Kinh doanh / Quảng cáo</i>	Ratings, reach, GRP, View, share — ngôn ngữ quen thuộc trong kinh doanh quảng cáo	Phân khúc theo hành vi thật, không chỉ theo nhân khẩu học bề mặt Tự động hoá trong quảng cáo hướng đối tượng
3 — Tương tác	<i>Biên tập + Digital/Social</i>	Đếm lượt bình luận, chia sẻ, lượt thích — chủ yếu là số lượng	Tỷ lệ rời bỏ nội dung, giữ chân khán giả, biết người xem rời đi ở đâu, khi nào và vì sao
4 — Thương mại	<i>Kinh doanh</i>	Doanh thu quảng cáo theo GRP truyền thống, hợp đồng theo gói	Định giá theo khán giả thực, đo hiệu quả quảng cáo đến hành vi sau xem
5 — Vận hành	<i>IT / Kỹ thuật</i>	Theo dõi tình trạng hệ thống, dữ liệu kỹ thuật cơ bản	Ứng dụng vận trù học bằng cách liên kết hiệu suất vận hành với chất lượng dữ liệu ở bốn lớp còn lại

Việc ra quyết định tại từng khối có gắn liền với lớp dữ liệu tương ứng?



III. DỮ LIỆU CŨNG CẦN NGÔN NGỮ CHUNG

Việc ra quyết định dựa trên dữ liệu có khả thi?

Một lượt xem trên app/web, một lượt xem trên YouTube, một lượt xem theo trên nền tảng MXH — là ba định nghĩa khác nhau của cùng một khái niệm.
Vấn đề: định nghĩa, cộng dồn, so sánh, tin cậy.



Xuyên suốt

Áp dụng được cho mọi nền tảng, không đổi nghĩa khi đổi kênh phân phối



Nhất quán theo thời gian

Không đổi cách tính giữa các kỳ báo cáo, so sánh được theo chuỗi thời gian



Kiểm chứng được

Đòi hỏi dữ liệu sử dụng có thể truy vấn ngược và xác minh

Ví dụ điển hình:

Một chuyên mục A: cần mời gọi quảng cáo, phục vụ mục đích kinh doanh

→ Báo cáo lượng khán giả theo dõi: các chỉ số về lượt view hoặc rating...

Vẫn chuyên mục A đó, cần đánh giá nội dung thu hút khán giả ra sao, nên tiếp tục hay nên dừng:

→ Báo cáo lượng khán giả theo dõi, ranking chương trình, chuyên mục với cũng các chỉ số về lượt view hoặc rating...

Bộ dữ liệu nào, phục vụ mục đích gì, ai là người ra quyết định, có đủ tin cậy để ra quyết định dựa vào bộ dữ liệu đó không?



Phương pháp luận, giải pháp công nghệ luôn “bị” ưu tiên

01

Dữ liệu có thể làm được nhiều thứ, sẽ có tác dụng ngay

Tập trung thiết lập công cụ, xây dashboard trước khi biết hiện đang có dữ liệu gì, ở đâu, ai dùng, chất lượng đến đâu.

02

Dữ liệu không được sử dụng để ra quyết định

Có công cụ, có dashboard nhưng không đúng người dùng, không dùng được để quyết định.

03

Dữ liệu là tài nguyên, nhạy cảm, không chia sẻ ra ngoài phạm vi của tổ chức

Một kho dữ liệu đóng kín không tạo ra giá trị.

V. BẮT ĐẦU TỪ ĐÂU?



Kiểm toán dữ liệu là một quá trình tìm hiểu và cấu trúc của hoạt động chuyên môn của 3 nhóm: nhóm tạo ra dữ liệu, nhóm dùng dữ liệu, và nhóm ra quyết định dựa trên dữ liệu.

Vòng lặp nhỏ với ba câu hỏi

1 Lớp dữ liệu nào?

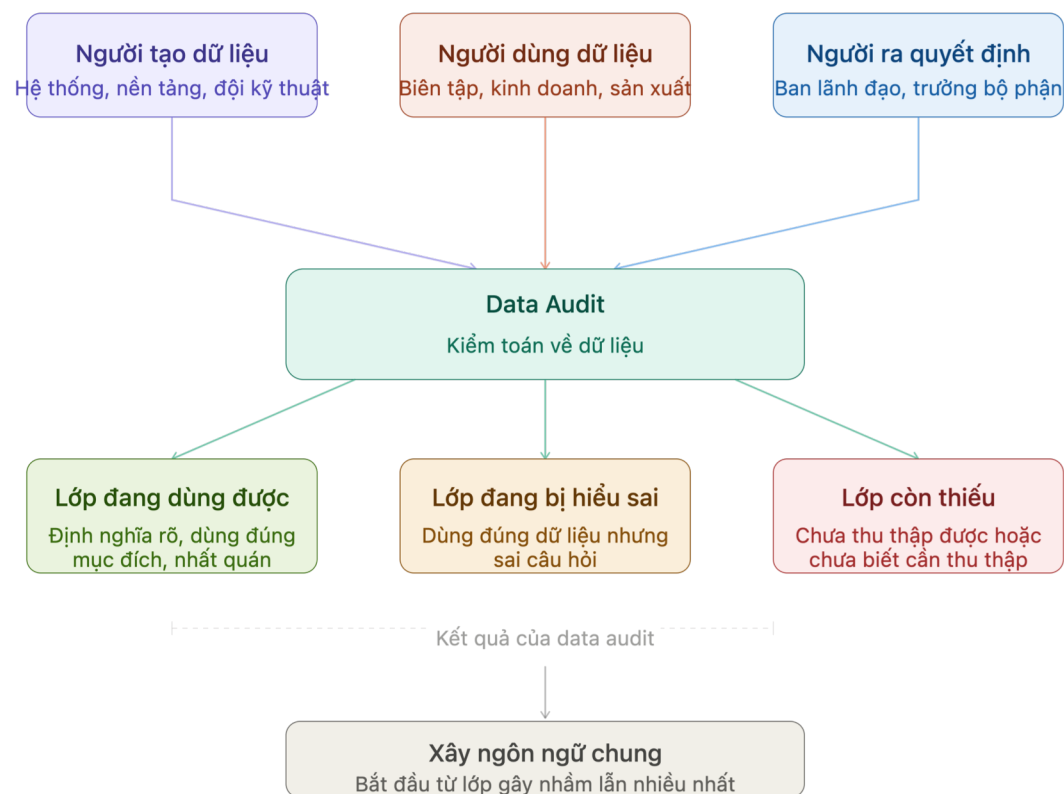
Lớp nào đang có sẵn nhiều nhất, dễ chuẩn hóa nhất, gần với một quyết định cụ thể nhất.

2 Đối tượng nào dùng?

Không phải là "toàn bộ tổ chức"
Ban biên tập phụ trách chuyên mục mục
Người quản lý hoạt động kinh doanh

3 Dùng cho việc gì?

Quy trình hoá việc sử dụng dữ liệu, trong đó quyết định dựa vào dữ liệu thực tế
Giữ hay bỏ một định dạng
Định giá quảng cáo theo cách nào.



KẾT

Tự chủ dữ liệu là năng lực cốt lõi trong quản trị dữ liệu

và các cơ quan báo chí nên tiếp cận theo cách hiệu quả nhất là bắt đầu từ một vòng lặp nhỏ và phát triển dần theo quy mô, phạm vi của tổ chức

Xin trân trọng cảm ơn Ban Tổ chức và toàn thể đại biểu.