

Creating a Culture of Data in Your Media Organization

Presented by Joel Hughes



Howdy, I'm Joel.

- I've been working on the tech and ops side of B2B media/publishing for nearly 20 years.
- Current Omeda customer for Omail, Audience Database, Print Circ fulfillment
- COO at EnsembleIQ. We serve the retail and CPG industries with content that helps them do their jobs
- Migration to Omeda last fall is the 3rd time I've built a centralized database
- What we're going to talk about today is in various phases of implementation, some done, some half-baked, some future thoughts
- Thanks to my team for a lot of this stuff

You might be in the wrong presentation if...

- You are killing it in print revenue
- Your editors and mail room are overwhelmed with fan mail like the North Pole
- You are turning away excess event attendees and sponsors regularly
- If client marketing budgets dry up, you're all good
- You're running a call answering service, vs. outbound call center for subscriptions and registrations. Like a bustling daily telethon
- Unlimited resources and cash

Defining Data Culture

A culture in which brand strategy, content strategy, sales, and audience management are all informed by purposefully meaningful data.

We still say this stuff...

Above the fold!

If Google is blocking pop-ups how else do we offer high impact ads?

Hmm, how can we serve more ads?

Write once and publish everywhere!

We'll just make a pop-up modal to ask unknown visitors who they are. I'm sure they will tell us in exchange for our amazing newsletter!

Hire a data scientist!

Homepage Redesign!

We should get into Podcasts

We should increase newsletter frequency with more news for more ads

Giant screens and stats displayed in the office will whip editorial into shape!

Programmatic and remarketing! That's the future!

This CMP/DMP solves all our problems! Drop the tags in ASAP!

We need more video. Can we do more video?

Somebody reading about this topic must clearly be a hot lead! Surveillance marketing is the next big thing!

Instead of this stuff.

How do we help our audience do their jobs every day?

Do we have the right folks in our audience?

Can we serve less ads?

How do we futureproof content?

How will this content work in a post-mobile world? Will we have to redo everything?

Where's "below the fold" on a voice device like an Alexa?

How do visitors actually use our content?

What assumptions are we making?

How can we think beyond "the page"?

What information do people need and in what format(s)?

What information might help editorial create really useful content?

What content formats would help our audience the most?

What new titles and roles are showing up in our industry?

Are we betting the farm on surveillance marketing?

Our industry has enough tools And SaaS Products.

 BlueConic

 Bombora

 BrightEdge

 OptiMonk

 CrazyEgg

 MediaMind

 33 Across

 LiveRamp

 Tynt Tracer

 Google Analytics

 Pardot

 MediaMath

 Rapleaf

 Dynatrace

 Everest Technologies

Analytics and Tracking

 Omniture SiteCatalyst
Marketing Automation

 Eloqua
Marketing Automation

 Rapleaf
Marketing Automation

 Adobe Dynamic Tag Management
Tag Management

 Chartbeat
Visitor Count Tracking

 Adobe Marketing Cloud
Audience Measurement · Marketing Automation

 Google Analytics
Application Performance · Audience Measurement · Visitor Count Tracking

 MOAT
Advertiser Tracking

 Google Universal Analytics

 Lotame Crowd Control

 MediaMath
Advertiser Tracking · Demand-side Platform

 Bombora
Advertiser Tracking

 LiveRamp
Lead Generation · Marketing Automation

 Facebook Signal

 Yahoo Web Analytics
Audience Measurement

 Everest Technologies

 OwnerIQ
Marketing Automation

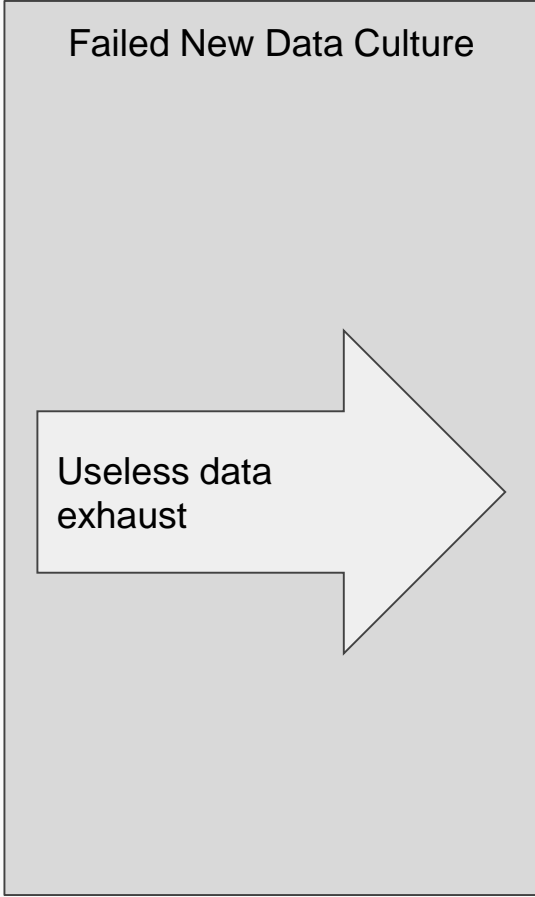
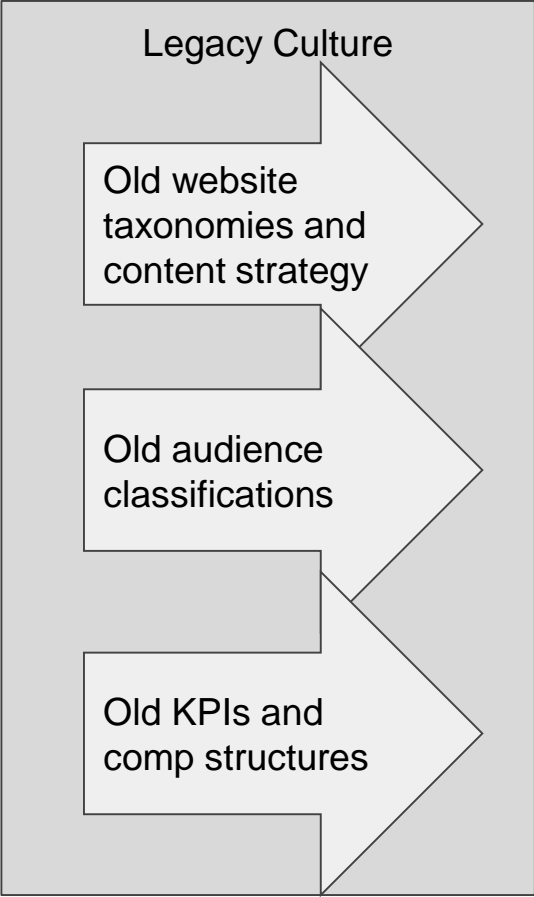
 comScore
Advertiser Tracking · Audience Measurement · Conversion Optimization · Site

We tend to either...

Feed these tools less-
than-meaningful data

And/or... Pile on more
tools

And/or... Ignore what
comes out of the tools.



OK, so how do we build a data culture?

1. Learn what we don't know about our **audience** and our **content**
2. Agree upon an organizational **lexicon**
3. Create systems that **classify** audience and content based on steps 1 & 2
4. **Educate** and **inform** your organization continuously
5. Create new KPIs and variable compensation structures that create **accountability** particularly with content creation

Learning what we don't know: Correcting Content Metadata/Tagging

Audience interactions with content may not be creating **accurate** or **actionable** metadata.

- Inconsistent topic tagging/classification by different editorial groups or websites, compounded over time as different content editors wash in and out of your organization
- Outdated or vague topic vocabularies
- Content tagged for non-topical reasons such as “to appear in a certain area of the website” or to trigger some ad targeting, etc.
- Vast swaths of totally unknown content still hanging around on a website

Learning what we don't know: Running all content through an AI-driven classification engine

- We have created an AI-driven **content classification engine**. Easier said than done.
 - The bulk of the effort here is **training the AI** for the correct predictions within each content vertical. Otherwise the machine will rapidly go haywire/bonkers with classification.
 - Unsupervised ML has unearthed classifications that are **deeper** than simple parent/child taxonomies. E.g. Vaping legislation vs. vaping dangers vs. vaping profitability vs. vaping sentiment
 - **Entity recognition** also has to be carefully trained and managed to make correct and relevant entity predictions

Beyond topics: Looking at the Why

Why is someone consuming our content?

- Becoming generally aware of a trend or issue they will need to solve for
- Actively trying to solve a problem they are already aware of
- They have knowledge of the problem, and the solution, and are now whittling down a short list of vendors and solutions

Learning what we don't know: Re-thinking Audience Metadata

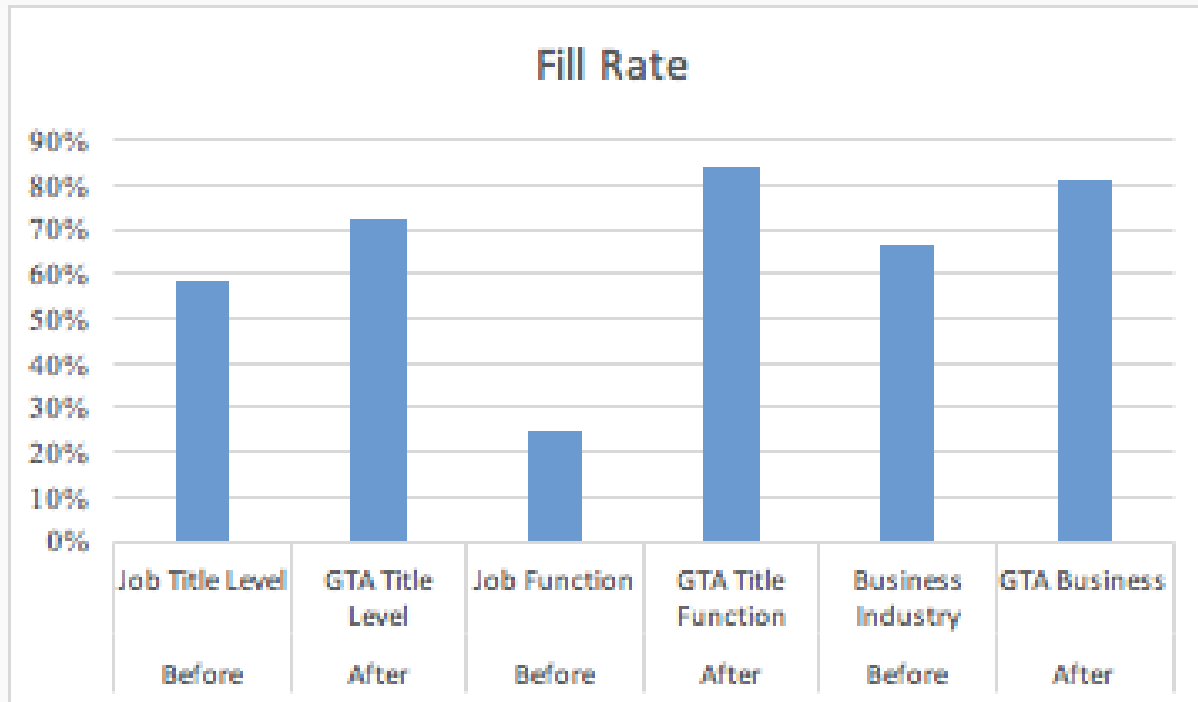
Ground Truth Data (GTD) Initiative: Unsupervised ML on Write-In titles -> Meetings with internal stakeholders showing them what they might not know about their audience -> Development of new corporate audience lexicon

Audience Data Carwash: Create ruleset and systemized classification of new and updated audience members. Entity recognition and specifically being smart about company identification and decoupling of company and individual data

Learning what we don't know: GTD Detail

- First we called together all internal **brand stakeholders** to come to consensus on a global lexicon for lead job function, job level, and business type, ignoring all historical classifications and assumptions.
- These meetings were assisted with AI visualizations using unsupervised ML to analyze hundreds of thousands of write-in titles and company names and doing a **cluster analysis** to show us blind spots in our audience.
- These meetings combined with the AI discoveries resulted in a **rules engine** to classify and standardize incoming leads and kick exceptions out for manual review.

GTD Audience Incremental Tagging: Initial Results



Beyond titles/levels: Looking at the Who

Who ultimately are we trying to convert to a lead and how do we help them with content? How do we classify these personas in addition to function/level/company? How will they find us?

- New to the industry. Coming out of school, in school, or a career change.
- Industry veteran but new to our brands.
- Industry veteran but in a new role or on the other side of the table.
- Consultant to the industry looking for temporary help for a project/client.

Keeping Humans in the Loop when using ML

Decide what content classification jobs are really best for humans to do.

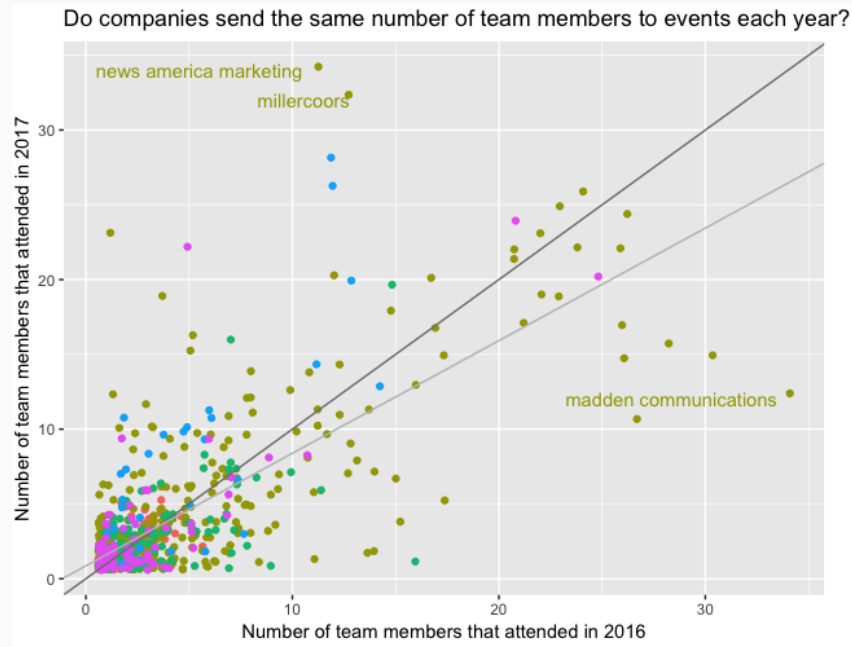
- Exception processing
- Intent path classification where applicable
- Constant re-training/maintenance of any ML system

“Life of a Lead” Roadshow - Internal Corporate Training

- Spread the information to the organization
- Two tracks: Sales, and everyone else
- Internal webinars and live presentations at key meetings
 - Sales Track: Teach sales team who we really have and what they really do
 - Everyone Else Track: Teach editors, brand leadership, accounting, and the reception desk where leads come from both current and future state
- Create internal certifications for demonstrated knowledge on the “Life of a Lead”

Continual Training

Create an Internal “Visualization of the week” newsletter



On Data Scientists

- **Is Not:** Somebody's smart nephew that makes charts and graphs
- **Is:**
 - First and foremost a software engineer, familiar with application architecture, dev environments/workflows, databases
 - Able to prep/groom data for use elsewhere, this being the bulk of the work and an art in itself
 - Is an expert at ML, text mining, and training AI models
 - Statistical modeling expert
 - Full of curiosity and passion for data and problem solving

What will we do with optimized audience data + content data + tools?

- Informed **content** creation, first and foremost.
- Informed **sales** conversations. Teach clients and prospects something they might not know (emerging titles, decision makers, buying teams, and trends).
- Useable engagement data to drive anything from **insights** to potential **purchase intent** path identification*

**Which is not a panacea BTW*

Questions?

Joel Hughes

joel@joel-hughes.com

<https://www.linkedin.com/in/joeldhughes/>