Data Harvesting and Analysis

Turning data into meaningful action for academic publishers

Chris Leonard
Head of Product / Emerald Publishing
Hi!
Discover new things

Journals, books and case studies - a library of real world research at your fingertips

Accounting, Finance & Economics
Business, Management & Strategy
Education
Engineering
Health & Social Care
HR, Learning & Organization Studies
Information & Knowledge Management
Library Studies
Marketing
Operations, Logistics & Quality
Data is an opportunity for Emerald

- Opportunity to get to know customer behaviour better (by analyzing what they do, rather than what they say they do)
- Opportunity to tailor content display to certain users
- Opportunity to tailor product collections for librarians
- Opportunity to spot and solve problems before they become problematic
But also...

- An opportunity to rethink what a publisher does
- Supporting researchers across the continuum from ideation and funding to impact and real world change
- Data opportunities are almost infinite
Some examples from the Now, Next and Future time horizons

1. Using data to explicitly map a user’s pathway to and through the Emerald Insight platform, and influencing behaviour for maximum usage through data (and UX).
2. Using usage and turnaway data to optimize product makeup for a librarian.
3. Using data about individuals to map their interests to funding opportunities, research outputs and collaborators
Data as a way of optimizing useful interactions on your site
Data as a way of optimizing usefulness of your product offering
Observations on ‘Products’ for libraries

We sell fixed content packages, for fixed time periods, in all international markets.

Library budgets under strain.

How to demonstrate value of purchases?
Need to add flexibility to our content packages

One commonly heard piece of feedback: we need flexibility to define journal collections.

A key part of Emerald’s future will lie in successfully addressing this request, but how?
With data!

- Librarian defines, with a Emerald business manager, a list of titles which they think matches their patrons’ needs - and is affordable for the library. T=100
- Each 6 or 12 months, we analyze usage patterns and investigate bottom X% used titles.
- Also analyze turnaways from unsubscribed titles
- If there is scope for improving usage, Emerald will proactively contact the librarian to suggest that titles A, B, and C are replaced with titles X, Y, and Z.

Usage up: Librarian happy they have maximum bang for their buck, Emerald happy with more usage and happy customers.
Using data to look at researcher support beyond publishing
Publishing is just one part of the research continuum

What happens before and after publishing is an opportunity too.

Researchers need to acquire funding, collaborators, and demonstrate impact outside of the publishing process.
Matching people to services

Person A: Has written 5 papers, classified as covering

1: AI, ecommerce, mobile tech, online shopping
2: ecommerce, subscriptions, millennials, Australia
3: psychology, ethics of persuasion, drought, fraud
4: sales phenomena, discounts, bulk purchasing, ecommerce
5: stockpiling, disaster relief, online, food security

ORCID (if up to date) suggests she is based in Adelaide, has been publishing regularly for last 4 years, has developed a niche in psychology and user behaviour of online purchasers in times of panic.
What can we do with this data?

Match to people with similar profiles (in terms of career length and subject matter) who have not been co-authors in the past.

Highlight their skills to people applying for grants in this area, or proactively to grant-awarding bodies.

Her work has lower bibliometric figures than peers - maybe some training on impact literacy would be ideal for this person.
### Evolution of Products

<table>
<thead>
<tr>
<th>NOW</th>
<th>FUTURE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Products = Content</td>
<td>Products = Content</td>
</tr>
<tr>
<td></td>
<td>= Data</td>
</tr>
<tr>
<td></td>
<td>= Platform</td>
</tr>
<tr>
<td></td>
<td>= Content + Data</td>
</tr>
<tr>
<td></td>
<td>= Platform + Data</td>
</tr>
</tbody>
</table>