Unlock Your Content

With NoSQL and Semantics

Matt Turner
CTO Media and Publishing
matt.turner@marklogic.com
Introductions

Me: Matt Turner, CTO Media and Publishing, MarkLogic

MarkLogic: New generation database software powering information products and mission critical applications
PUBLISHING

CHANGE IS THE ONLY CONSTANT
FROM PUBLISHING

- Company Data
- Industry Data
- Filings
- Reports

Product A Dedicated Infrastructure (database + search engine)

Product B

Product C

FORM BASED PRODUCTS

DEDICATED PRODUCT INFRASTRUCTURE
TO INFORMATION PROVIDERS

Deliver the right content, to the right user, in the right format, in real time.
Top 5 Requirements for Information Publishers

1. Getting data *IN* fast isn't the problem – it's getting insights *OUT* Faster!
2. Data is complex – but users want complexity hidden!
3. Not everyone has permission to access all the data…
4. Repurpose, repurpose, repurpose. Repeat
5. Once you attract them – you must be reliable
Top 5 Requirements for Information Publishers

1. Getting data IN fast isn’t the problem – it’s getting insights OUT Faster!
2. Data is complex – but users want complexity hidden!
3. Not everyone has permission to access all the data…
4. Repurpose, repurpose, repurpose. Repeat
5. Once you attract them – you must be reliable
Traditional Technology

Rows and columns for content strip information?

<table>
<thead>
<tr>
<th>Title</th>
<th>Publication Date</th>
<th>Category</th>
<th>Abstract</th>
<th>Section</th>
<th>Section 2?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Science Article 1</td>
<td>3/1/14</td>
<td>Biology</td>
<td>Abstract text...</td>
<td>Section text</td>
<td>Section text</td>
</tr>
<tr>
<td>Research Book</td>
<td>6/4/13</td>
<td>Surgery</td>
<td>Abstract text...</td>
<td>Section text</td>
<td>Section text</td>
</tr>
<tr>
<td>Science Article 2</td>
<td>6/4/05</td>
<td>Chemistry</td>
<td>Abstract text...</td>
<td>Section text</td>
<td>Section text</td>
</tr>
</tbody>
</table>
Traditional Technology

Rows and columns for content strip information?

Hierarchical taxonomies overlap and don’t capture the complexity?
We Are The New Generation Database

HIERARCHICAL ERA
“For your application data!”
- Application & hardware specific

RELATIONAL ERA
“For all your structured data!”
- Normalized, tabular model
- Application-independent query

ANY STRUCTURE ERA
“For all your data!”
- Schema-agnostic
- Massive scale
- Query and search
- Analytics
- Application services
- Faster time-to-results
Model Content with NoSQL
Semantics to Model Relationships

Data model to manage relationships and link together data

‘triples’ describe single facts

Collections of facts describe complex real-world scenarios

"John Smith" livesIn "London" isIn "England"

livesIn
Ontologies Instead of Taxonomies
NoSQL and Semantics Unlock Content

Flexibility + Efficiency = Complete Picture
Use Case: Master Data

Foundational data to drive product development and innovation
Use Case: Enhance Digital Products

Link together and present content with relationships
Use Case: Go Beyond Search

Concepts, ontologies drive rich discovery experiences
Use Case: ‘Everything Else’

Tailored to the users role, activity and location
Top 5 Requirements for Information Publishers

1. Getting data IN fast isn’t the problem – it’s getting insights OUT Faster!
2. Data is complex – but users aren’t!
3. Not everyone knows how to use your data...
4. Repurpose, repurpose, repurpose. Repeat
5. Once you attract them – you must be reliable
Time’s Up!

About your speaker:
Name: Matt Turner
Company: MarkLogic
Email: matt.turner@marklogic.com
Social Media: @matt_turner_nyc

Link:
http://www.marklogic.com/blog/listicle-top-5-requirements-todays-information-providers/

FREE BEER
After the Event
Bloomsbury Tavern
2 minute walk