AI/ML for Operational Data Management

A conversation with

Masood Khatri
VP, Product Owner
Xoriant CDi

Apnav Agrawal
Director, Data Management
Xoriant CDi
Moderated by Mike Meriton
Co-Founder & COO, EDM Council

- Joined EDM Council full-time 2015 to lead Industry Engagement
- EDM Council Co-Founder & First Chairman (2005-2007)
- Former CEO GoldenSource (2002-2015)
- Former Executive for D&B Software and Oracle
- FinTech Innovation Lab – Executive Mentor (2011 – Present)
Speaker Profiles

Masood Khatri
VP, Head of Digital Practice, Product Owner

Technology Executive & Recognized Thought Leader with passion for building and leading transformational digital products and bringing businesses to the forefront of technology and innovation. Masood has extensive experience in strategically designing, building products, roadmaps & bringing ideas/concepts to life to drive implementation. Masood has developed and deployed cutting-edge cross-functional digital capabilities to needs of fortune 100 enterprises, which exceeded business and customer expectations.

Apnav Agarwal
Director, Data Management and Governance

A DCAM certified professional having over 15 years of experience in Data Management & Governance with proven ability to offer strategic direction/guidance to enterprises for enterprise data strategy development and management. Apnav has experience in guiding the team related to the data warehouse, business rules, data modelling, architecture creation, executive collaborations, and technical requirements. Apnav is leading a team of data management SMEs while developing and implementing SaaS products according to client specifications.
Traditional Methods- Operational Data Management

1. Subscribe to Feeds
2. Transformation & Business Rules
3. Source Information Manually
4. Manual Reviews, 4 Eye Checks
POLL – Challenges
Challenges - Operational Data Management

- Labor Intensive
- Human Errors Impact Data Quality
- Lack of Consistency - Variety of Data Sources and Formats
- Delays - Volume of Data to Manage
Operational Efficiencies in Data Management

Data Sources

- PLATFORM
- PROCESS
- PEOPLE

Clean and Up-to-date data to fulfill business objectives and regulatory compliance

AI/ML to improve Operational Efficiency
AI for Operational Data Management

- DQ Framework to Enhance Data Quality
- Enrichment and Standardization of Addresses
- Automating Hierarchy Data Extraction

AI Techniques Used
- Natural Language Processing
- Machine Learning
- Deep Learning
POLL – AI/ML
DQ FRAMEWORK TO ENHANCE DATA QUALITY
DQ Framework to Enhance Data Quality

Blackrock Group Ltd
Sourced Data
Attribute : Party Type
New Value : Other Financial
CDi Gold Database
Old Value : Investment Advisor

Apply Data Science
• Natural Language Processing
• Machine Learning Models :
  Feature Extraction, Feature Engineering, Classification

CDi DQ Framework Recommendation
Accept
Results

Details of Xoriant Solution

Sample Results

<table>
<thead>
<tr>
<th>#</th>
<th>Attribute</th>
<th>Old Value</th>
<th>New Value</th>
<th>Prediction Probability</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>IMMEDIATE_PARENT</td>
<td>REINSURANCE GROUP OF AMERICA INC</td>
<td>RGA REINSURANCE COMPANY</td>
<td>0.9529</td>
<td>Accept</td>
</tr>
<tr>
<td>2</td>
<td>IMMEDIATE_PARENT</td>
<td>BANK OF HAWAII</td>
<td>First Hawaiian Bank</td>
<td>0.5129</td>
<td>Review</td>
</tr>
<tr>
<td>3</td>
<td>IMMEDIATE_PARENT</td>
<td>BNY MELLON FUND MANAGEMENT LUXEMBOURG SA</td>
<td>BNY MELLON GLOBAL FUNDS PLC</td>
<td>0.0202</td>
<td>Reject</td>
</tr>
</tbody>
</table>

Business Benefit

Our DQ Framework automated exception resolution saving about 60% of the manual effort.
ADDRESS ENRICHMENT AND STANDARDIZATION
Address Enrichment and Standardization

Apply Data Science
- Information Extraction
- Natural Language Processing
- Machine Learning Models:
  - Feature Extraction, Feature Engineering, Classification

Standardize Address Fields
- Address 1: Suite 720
- Address 2: 343 Thronall Street
- City: Edison
- State/Province: NJ
- Country: USA
- Postal Code: 08837

Cleansed Data

Xoriant Corporation
- Registration Authority
- Corporate Site
- Stock Exchange
- SEC
- Regulator Site

Models
- Python
- Word Embeddings
- Pre-trained Models
- Labeled Data

Deep Learning
Edit Address

Enter One Line Address and Select the Country – Click Validate

<table>
<thead>
<tr>
<th>Entity Maintenance</th>
<th>Workflow user history</th>
</tr>
</thead>
<tbody>
<tr>
<td>9006509837</td>
<td>Invensys Middle East FZE</td>
</tr>
<tr>
<td>Basic Classification</td>
<td>Address</td>
</tr>
<tr>
<td>Legal</td>
<td>Business</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>English</th>
<th>Non-English</th>
</tr>
</thead>
<tbody>
<tr>
<td>C/O or Attn</td>
<td>Street Address 1</td>
</tr>
<tr>
<td>Street Address 2</td>
<td>Street Address 2</td>
</tr>
<tr>
<td>Street Address 3</td>
<td>Street Address 3</td>
</tr>
<tr>
<td>Country</td>
<td>State/Province</td>
</tr>
<tr>
<td>CANADA</td>
<td>Select</td>
</tr>
</tbody>
</table>

Comments

Select a comment
Address Standardization

Address Standardization using ML/AI provides enriched address in separate fields for user to review.
Address Standardization

Address automatically applied to respective attributes.
# Address Standardization Accuracy

<table>
<thead>
<tr>
<th>Field Level Accuracy</th>
<th>North America</th>
<th>Europe</th>
<th>Asia</th>
<th>Australia</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Address Level Accuracy</strong></td>
<td><strong>North America</strong></td>
<td><strong>Europe</strong></td>
<td><strong>Asia</strong></td>
<td><strong>Australia</strong></td>
</tr>
<tr>
<td>Address 1 (Building, Unit/Suite, Floor)</td>
<td>Canada 90</td>
<td>US 93</td>
<td>Germany 84</td>
<td>France 92</td>
</tr>
<tr>
<td>Address 2 (Street)</td>
<td>76</td>
<td>97</td>
<td>78</td>
<td>84</td>
</tr>
<tr>
<td>Address 3 (Landmark)</td>
<td>76</td>
<td>99</td>
<td>94</td>
<td>84</td>
</tr>
<tr>
<td>City</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>98</td>
</tr>
<tr>
<td>State or Province</td>
<td>100</td>
<td>99</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Zip or Postal Code</td>
<td>100</td>
<td>99</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

Handling volume of 1 million entities having 3 addresses per records would have taken 1,000 person-days. Supported by 90%+ accuracy through AI/ML this effort is reduced by 900 person-days.
AUTOMATING HIERARCHY DATA EXTRACTION
Automating Hierarchy Data Extraction

Brewin Dolphin Limited
- Registration Authority
- Corporate Site
- Stock Exchange
- SEC
- Regulator Site

Apply Data Science
- Information Extraction
- Natural Language Processing
- OCR
- Machine Learning Models:
  - Feature Extraction, Feature Engineering, Classification

Immediate Parent Identification
- Kabouter Management LLC
- Royal London Asset Management
- Henderson Group PLC
- Aberforth Partners
- Blackrock, Inc.
- FIL Limited
- Kames Capital

Repeat N steps

Ultimate Parent Identification
- Kabouter Management LLC: 6.0% (Holding)
- Royal London Asset Mgmt: 5.95% (Holding)
- Henderson Group: 5.09% (Holding)

With more than 50% holding, Brewin Dolphin Limited is Ultimate Parent.

Cleansed Data
Benefits

- Improved Data Quality
- Increased Productivity
- Speed
- Reduction in Cost
Questions?
FOR MORE INFORMATION:

MASOOD KHATRI
VP, Head of Digital Practice, Product Owner
masood.khatri@Xoriant.com
Cell: 732.718.9547

APNAV AGRAWAL
Director, Data Management & Governance
apnav.agrawal@Xoriant.com
Cell: 646.895.3538

Visit us at https://cdi.xoriant.com to learn more