



Real World Data Curation 101: Understanding and Utilizing Unstructured Data

August 22, 2024



a medical knowledge group company

Magnolia Market Access: Meet Your Presenters



Beni Turner

*Director,
Real World Clinical Insights*



Martine Etschmaier

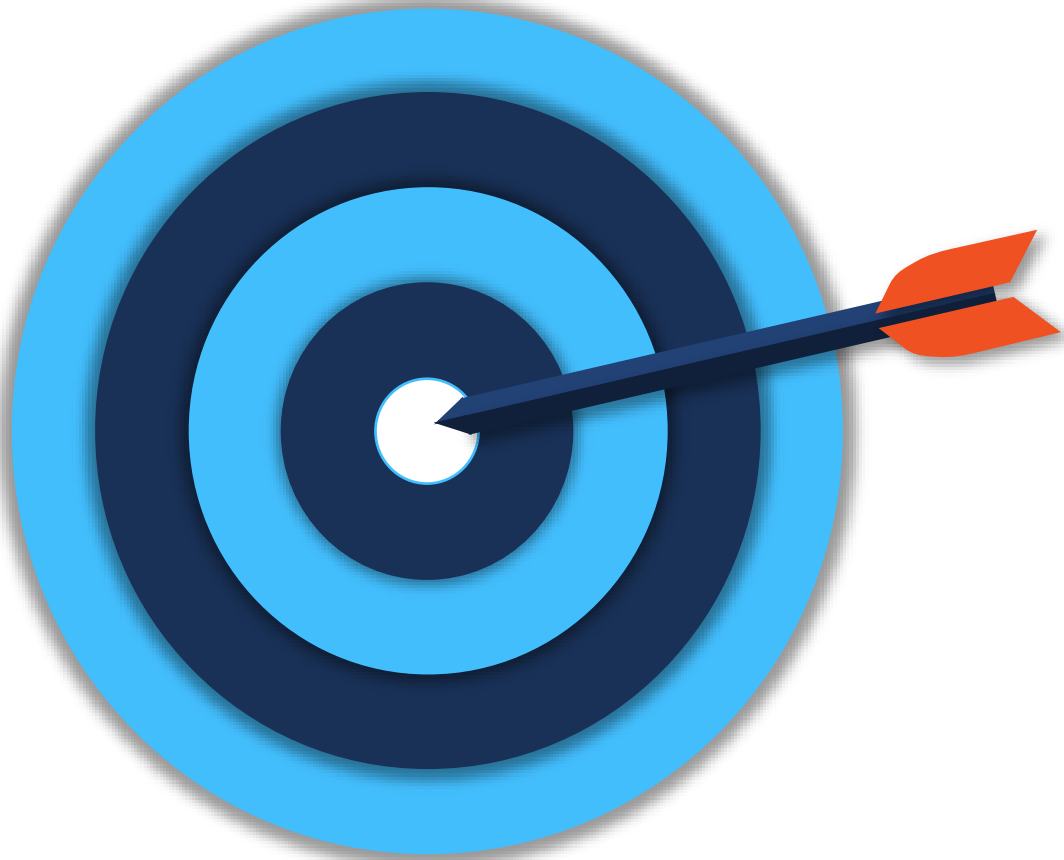
*Senior Director,
Global Market Access & Value Insights*



Pamela Landsman-Blumberg

*Senior Vice President,
Real World Evidence & HEOR Strategy*

Objectives



- 01** Defining Unstructured Data
- 02** Unstructured Data Sources
- 03** Curation Considerations
- 04** Practical Applications
- 05** Future Trends & Innovation Opportunities

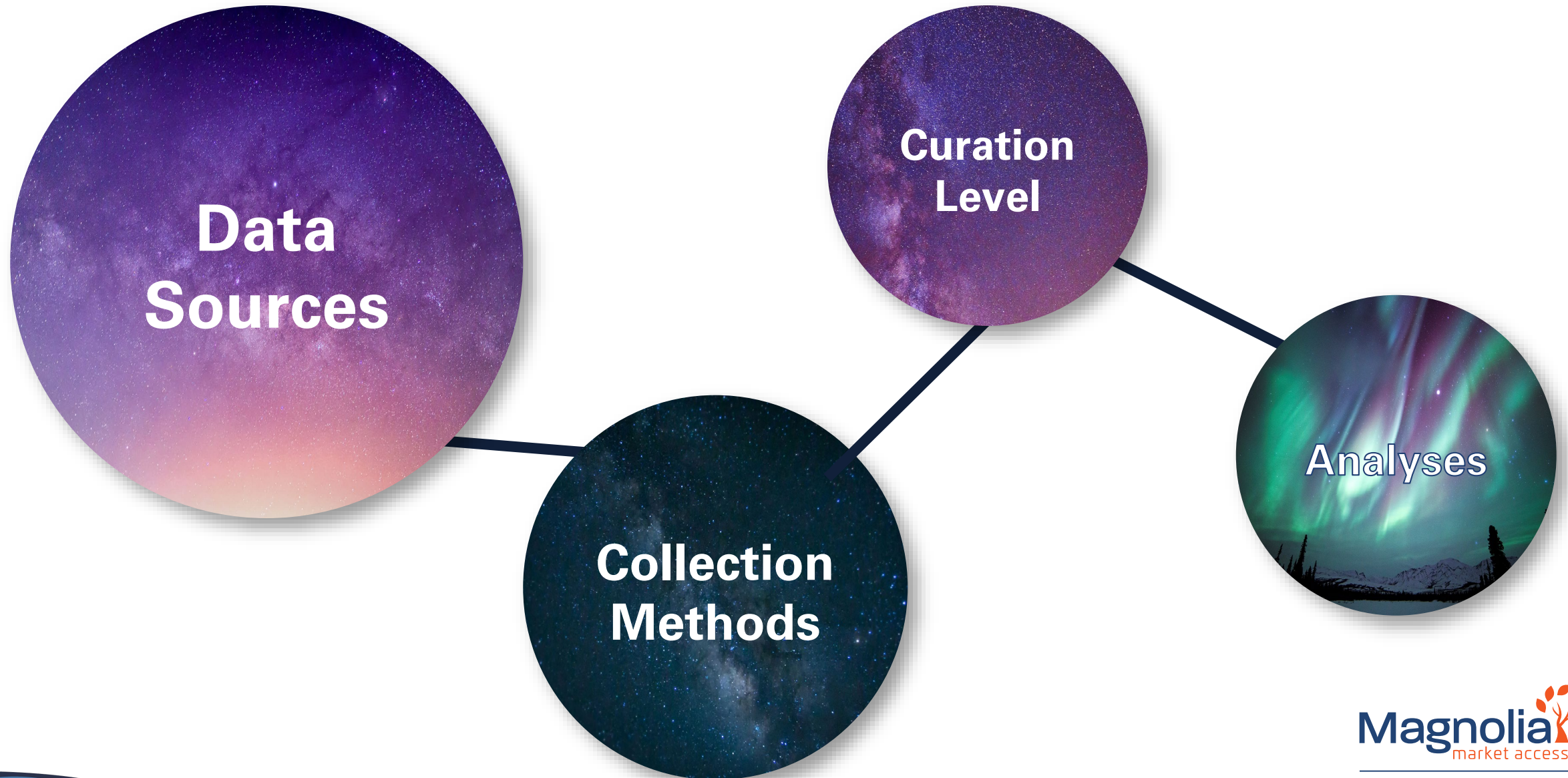
Introduction

Unstructured Data, Curation, & Considerations

The Big Data Era: Unraveling the Vastness of Our Digital Universe

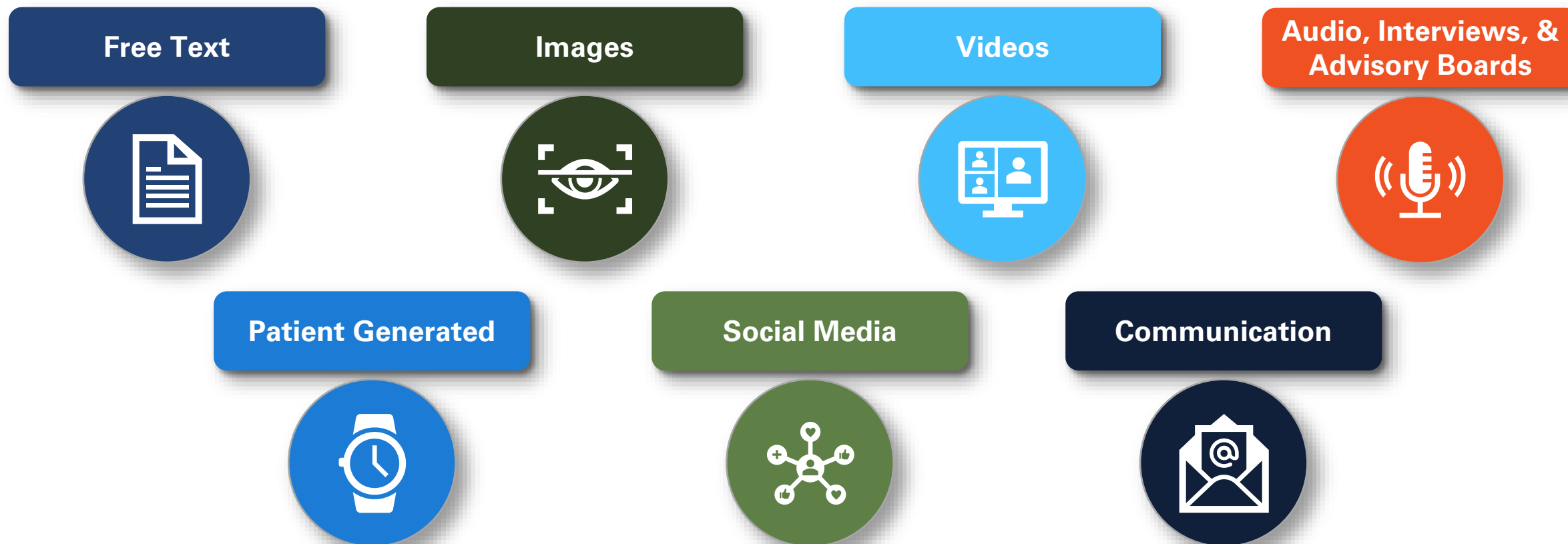


Mapping the Data Cosmos: From Sources to Collection to Insight



Unstructured Data

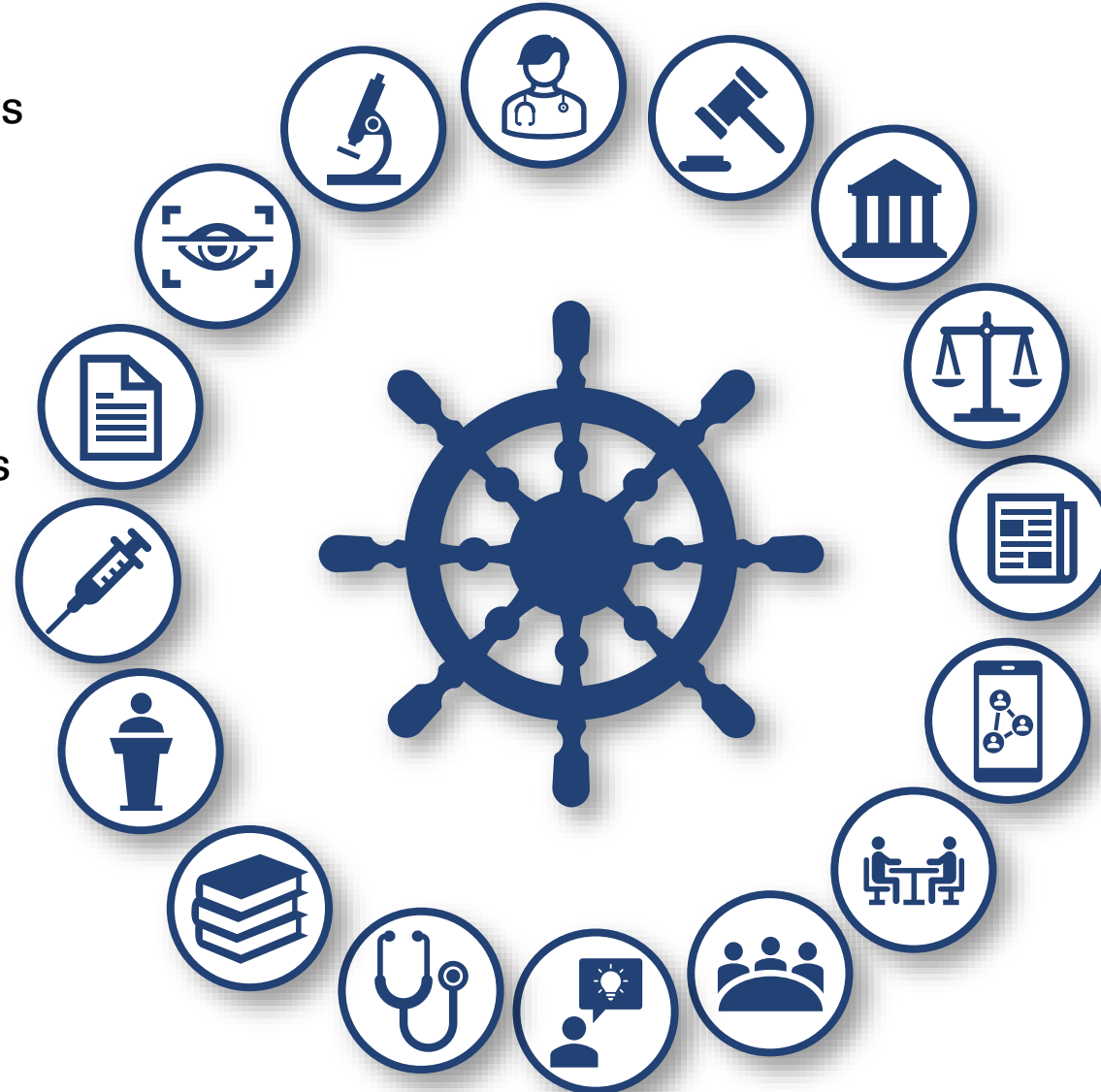
- Information that doesn't fit neatly into a predefined data model or schema
- Has no inherent organization and requires some level of curation for analysis and insight generation



Navigating Unstructured Data Sources

Charting the Course for Market Access Strategy

Healthcare Provider Notes
Pathology Reports
Imaging Reports
Claims Data
Clinical Trial Results
Conference Presentations
Scientific Literature
Clinical Guidelines
Expert Opinions



Patents
Regulatory
Updates
Policy Papers
Industry News
Social Media
Interviews
Online Forums

Unstructured Data to Inform Market Access Strategy – Selected Example Sources

Choosing the Right Methodology for Primary Market Research to Gather Insights From Key Stakeholders



Primary market research provides manufacturers with proof of assumptions, key unknowns in the market, actionable next steps, etc., and can service as a support function for other market access tools

Understanding Data Curation

- The process of providing shape to shapeless data



Important Curation Considerations

- Based on data use and audience, may need to consider:
 - Annotation
 - Linking
 - Compliance
 - Monitoring & Updating

Shaping Chaos: Varying Levels of Structure in Unstructured Data Curation

CHAOS

ORDER



The level of data structure provided through curation efforts exists along a continuum

Shaping Chaos: Varying Levels of Structure in Unstructured Data Curation

CHAOS

ORDER



Loosely structured curation can include:

- Thematic Mapping
- Concept Mapping
- Sentiment Analysis

Shaping Chaos: Varying Levels of Structure in Unstructured Data Curation

CHAOS

ORDER



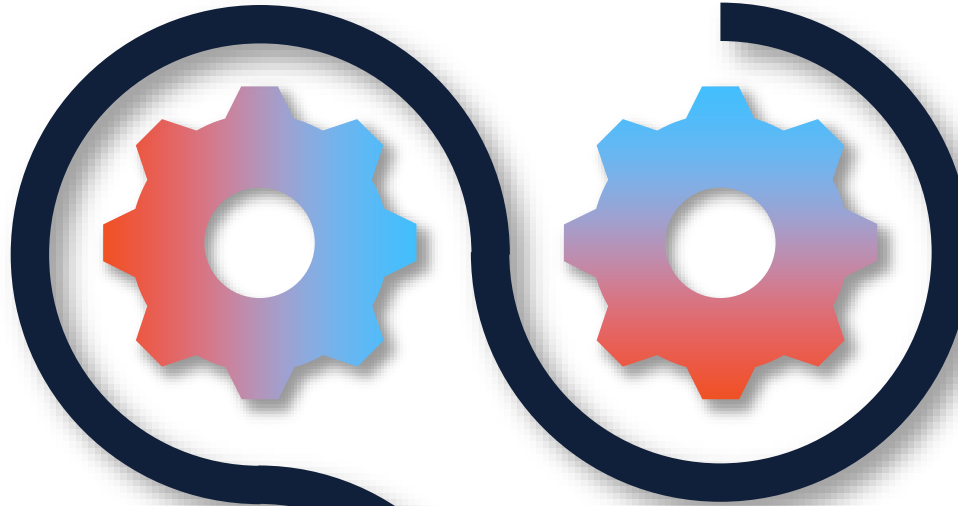
Highly structured curation can include:

- Coding Qualitative Data
- Systematic Literature Review
- Clinical Chart Review or Abstraction

Curation Concerns Begin at Project Ideation

DATA RELIABILITY

Curated data feeds should be continuously tested for reliability and accuracy against source data

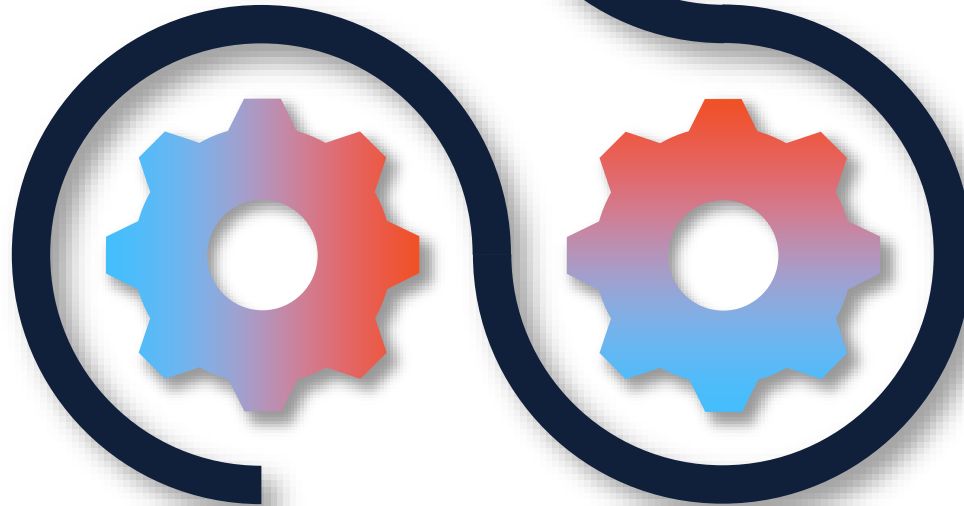


INTERPRETING MISSINGNESS

Lack of or missing data does not necessarily indicate a no

ADDRESSING BIAS

Design of curation plans should be done with an eye for potential sources of bias



CROSS FUNCTIONAL COLLABORATION

Development of data capture forms and collection system logic should be performed with representative input across the project team

Curating Unstructured Data Utilizing AI, NLP, & LLMs

Ensuring Accuracy & Quality in AI Curated Data

1



SOURCING DATA

Data sources should be *diverse*

Reduce risk of bias and false positives, data sources should be relevant and specific to the project, heterogenous, and representative of diverse populations and/or opinions

2



TRAINING

Algorithms require training to ensure *true positive* results

Critical to ensure that algorithms or models used are consistently and specifically trained and validated on data of interest for the analysis

3



REVIEW

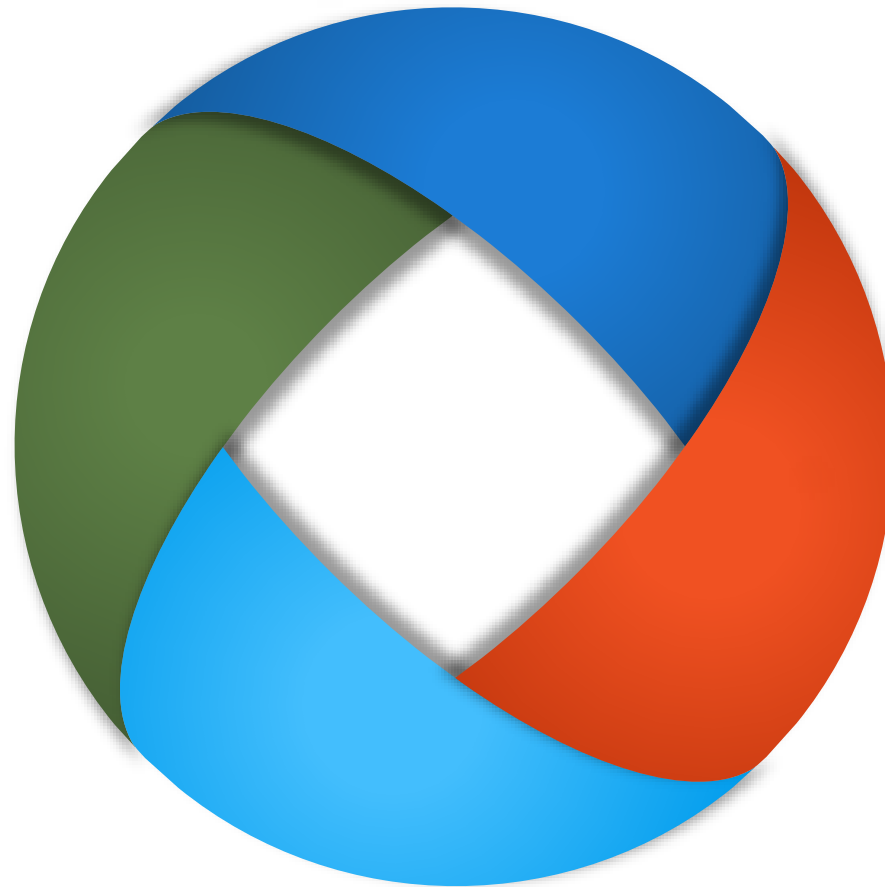
Human monitoring of algorithms is necessary

Comply with regulatory standards through the use of regular, streamlined human review of algorithm curated data

Data Integrity: The Essential Role of Retaining Source Links & Audit Trails

 Accuracy

 Transparency



Reproducibility 

Accountability 

Applications for Unstructured Data

How Can Unstructured Data Solve for Present Day Challenges
In Market Access

Real-World Treatment Performance

Enhancing Traditional Structured Real-World Evidence Insights



Challenge

- Market access decisions often require evidence of a treatment's effectiveness in the real world, beyond highly controlled clinical trial data
- Real-world treatment practices may differ from guidelines



Unstructured Data

- Healthcare Provider Notes
- Pathology Reports
- Imaging Reports
- Claims Data



Curation Benefits

- Demonstrate value
- Support & enhance submission value
- Competitor evaluation
- Enhance pricing & positioning strategies

The Payer Perspective

Strengthen Value Propositions, Evidence Plans & Coding Strategies



Challenge

- Smaller sample sizes limit ability to extrapolate to each payer segment
- Value of product to stakeholders is unclear
- Payers have data needs that differ from data garnered for regulatory review
- Inadequate coding can make patient and treatment identification difficult



Unstructured Data

- Clinical Guidelines
- Expert Opinions
- Interviews
- Policy Papers
- Industry News



Curation Benefits

- Maximize access
- Gain insights for data generation activities
- Uncover barriers
- Validate economic model assumptions
- Confirming price and rebate strategies
- Support creation or affirmation of codes

Competitive Intelligence

Leveraging Unstructured Data to Enhance Competitive Market Intelligence



Challenge

- Staying updated on competitors' market strategies and product performance is critical
- Understanding market landscape requires constant review of unstructured data



Unstructured Data

- Conference Presentations
- Patents
- Clinical Trial Results
- News Articles



Curation Benefits

- Gain intelligence to assist with product positioning
- Identify emerging trends
- Proactive pipeline customized to unmet needs

Providing Context: Size & Content of the EHR

In the US, hospitals alone generate an average of **838,440 terabytes** of patient medical data per day



Providing Context: Size & Content of the EHR

838,440 TB

12,000 times
the size of the
entire Netflix
content library

If the EHR is an Iceberg...



If the EHR is an Iceberg...

Structured Data is what's visible above the water **<20%**

Unstructured Data is below the water's surface **>80%**



If the EHR is an Iceberg: What We Can See Above Water



- **Demographics:** Sex, Age, Ethnicity, etc.
- **Diagnosis and Billing Codes:** ICD-CP ICD-PM, CPT, HPCS, LOINC
- **Numerical Values:** Height, Weight, BP, etc.
- **Categorical Values:** Blood Type, Disease stages of diagnosis, etc.
- **Medications**
- **DE&I**

If the EHR is an Iceberg: What is Below the Surface?



- **Signs & Symptoms**
- **Provider Notes, Results, & Reports (Clinical, Pathology, Operative, Imaging, Discharge Summaries, etc.)**
- **Procedure Details**
- **Problem Lists**
- **Lab Reports**
- **Genomics Reports**
- **OTC Medications**
- **Tumor Size & Characteristics**
- **Minimum Residual Disease**
- **Lines of Therapy**
- **Reasons for Medication Discontinuation**
- **Disease Severity Measures**
- **Patient Reported Outcomes**
- **Social History & Nuanced, Nonclinical Social Needs**
- **Patient Generated Data (wearable, DTx, etc.)**

If the EHR is an Iceberg: What is Below the Surface?



Unstructured Data may be at different depths

Curation Difficulty Increases with Data Depth

If the EHR is an Iceberg: Changes in Water Levels



- **Signs & Symptoms**
- **Date of Death**
- **Cause of Death**
- **Lab Results**
- **Behavioral Risk Factors**
- **Disease Assessment Scales**
- **Progression-Free Survival**
- **Performance Scores**
- **Histology & Genomics**

Real-World Patient Experience

Demystifying the Patient Experience Outside Clinical Trials



Challenge

- Siloed data streams
- Lack of information surrounding patient sentiment, experience, and concerns in structured data



Unstructured Data Sources

- Healthcare Provider Notes
- Social Media
- Online Forums
- Pathology Reports
- Imaging Reports



Curation Benefits

- Reveal important patient details
- Identify emerging trends
- Increased strategy specificity
- Improve access and support

The Regulatory Landscape

Tracking Regulatory Updates, Optimizing HTA Submissions, & Enhancing Clinical Trials with RWD



Challenge

- Constantly evolving guidelines and requirements for regulatory considerations
- Timely information critical to compliance
- Inherent clinical trial enrollment challenges, including lack of diversity, pharmaceutical companies require integration of novel data generation methods



Unstructured Data

- Scientific Literature
- Clinical Guidelines
- Expert Opinions
- Regulatory Updates
- Policy Papers
- Industry News



Curation Benefits

- Proactive & timely strategy adjustments
- Enhance cost evidence
- Ensure continuous compliance
- Maintain competitive advantage

Benefits & Value

Enhancing Market Access Strategy Through Unstructured Data

The Value: Incorporating Unstructured Data in Market Access Strategy

Optimize Pricing & Reimbursement Strategy



Tailor Stakeholder Engagement



Influence Regulatory & Policy Adaptation



Improve Resource Allocation



Patient-Centered Strategies



Incorporate Predictive Analytics & Forecasting



Crisis & Reputation Management



Future Trends

Opportunities for Innovation in Unstructured Data Curation

Future Trends & Innovation Opportunities

Increased Need
for Interoperability
Focus

01



Expanding AI-
Driven Curation

02



Novel Approaches
to Ethical
Considerations

03



07

Enhancements in
Data Quality &
Management

06

Crowdsourcing &
Collaborative
Curation

05

Shifting to Real-
Time Curation

04

Supercharging the
Audit Trail

Join us for our next *Navigating Market Access with Magnolia*

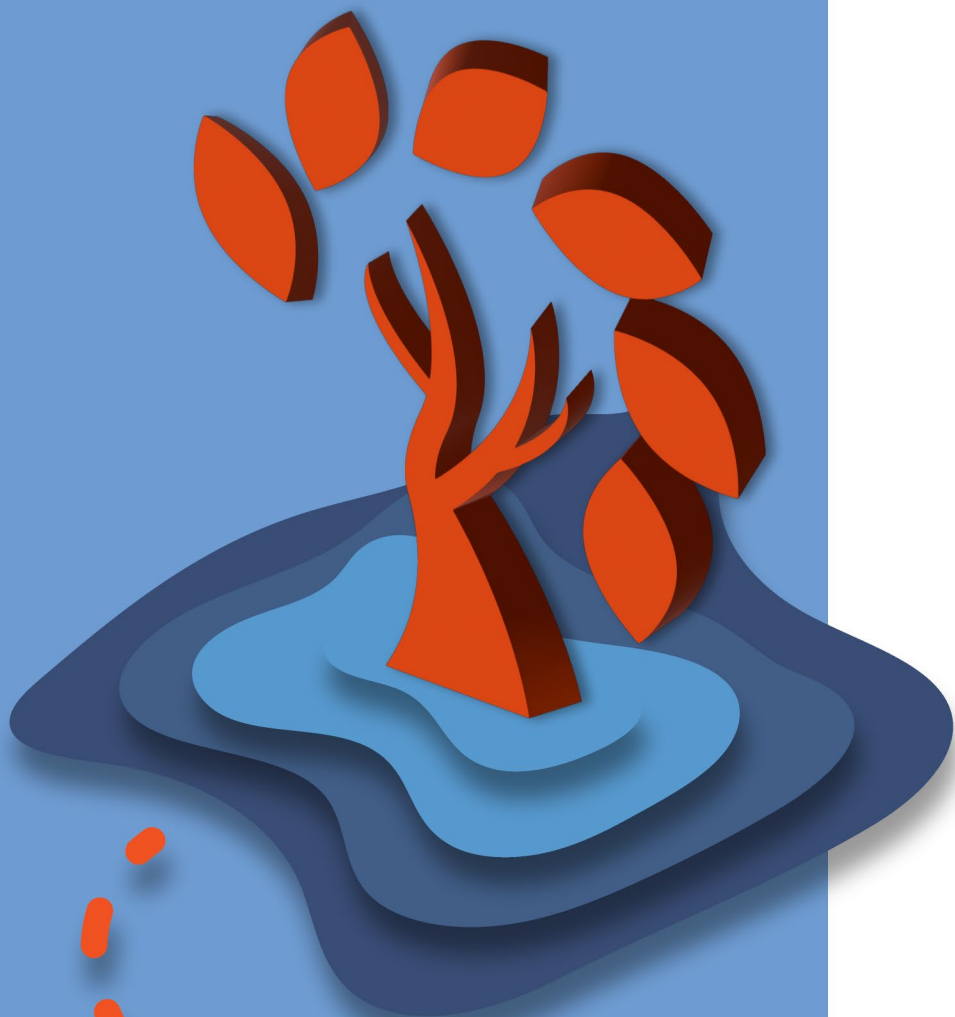
Reimagining Payer Strategies

New Methodologies to Identify Key Payer Accounts & Drive Engagement

Thursday, September 26, 2024
12:00 PM to 1:00 PM Eastern

Magnolia
market access

a medical knowledge group company



Thank you!

Pamela Landsman-Blumberg

SVP, Real-World Evidence & HEOR Strategy

pblumberg@magnoliamarketaccess.com

Martine Etschmaier

Senior Director, Global Market Access &
Value Insights

metschmaier@magnoliamarketaccess.com

Beni Turner

Director, Real-World Clinical Insights

bturner@magnoliamarketaccess.com