

Top 10 Factors for a Successful TDM Project

Mary Ellen Bates

BatesInfo.com

August 13, 2021





Text and data mining (in today's context) is

An automated process

that selects and analyzes large volumes of text or data

to increase discoverability of content and/or discern patterns and relationships



Factor #1: Know the different *types* of KM uses

DISCOVERABILITY

Goal: increase recall with high precision

Outcome: highly relevant full-text articles

Example: Instead of reviewing all articles that had the keywords "Glutamate" and "Obsessive Compulsive Disorder", retrieve only those where glutamate has an effect on OCD



Factor #1: Know the different *types* of KM uses

PATTERN DISCERNMENT

Goal: find patterns and trends across a dataset

Outcome: hypotheses and predictions of likely prospects for therapy, material design, or strategy—*not articles*

Example: Match biochemical properties of molecules to a viral protein's properties to identify a molecule likely to bind to the virus



Factor #2: ID the tangible value points

***Discoverability* projects enhance access to full text; can associate value with output**

***Pattern discernment* doesn't need full text; can't tie value to output**

How will you connect the licensed content to value and outcome?



Factor #3: ID the value of info pros

We understand copyright issues

Greater discoverability = greater demand for content

Ensure license compliance, avoid lockouts

We leverage information for greater ROI

We know how to promote resources internally

Higher ROI for online content subscriptions



Factor #3: ID the value of info pros

We collaborate with other groups

TDM projects involve skills from collections development, cataloging/metadata, IT, outreach

We identify and curate knowledge structures

Internally-developed taxonomies, ontologies sourced from third parties, open access knowledge models



Factor #4: License negotiation

TDM licensing is a challenge for all parties; different perspective on where value lies

Conversations can be lengthy, technical; we're all figuring it out

Info pros bring an enterprise-wide perspective of info use and value



Factor #5: Resist 100% tech solutions

Resist quick fixes and easy wins

Technology is only part of the solution

It takes humans to

understand what question is being solved

inventory existing content

ID needed content and knowledge structure

build/manage a taxonomy



Factor #6: Raise awareness of KM magic

"The magic happens once the content is brought in-house and we figure out how to make it useful. TDM brings intelligence to the data."

Expand users' perspective of information sources

Identify data sets that have consistent metadata

How are users likely to query the content?



Factor #7: Be the Chief Ontology Curator

Identify internal data silos, taxonomies

Sell users on the value of sharing (when appropriate)

Facilitate resource coordination

**Show the value of cross-platform searchability of
resources**



Factor #8: Model good behavior

Be known for authoritative sources and skills

Build a good library portal

Tangible proof of enhanced discoverability

Own your expertise in evaluating new resources & tools

"We can't solve all your KM problems, but we can serve as a template"



Factor #9: Be in the (virtual) room with your users

Find the internal communications channels

What do users need?

What do they (think they) already have/know?

Be proactive! What do users not know they need?

Offer your expertise, resources



Factor #10: Ask more questions

What's the purpose? What's the ideal outcome?

Is this to better uncover existing content or to discover new connections?

What kind of data do you need? What do you already have?

Will you need APIs developed for this project?



Factor #10: Ask more questions

Do you already have an ontology or knowledge graph?

Can or should this project's content or metadata be shared within our enterprise for later uses?

Should we get an institutional license for this content?



Factor #10: Ask more questions

Should other stakeholders be brought in for cost-sharing?

What are your plans for archiving the content and metadata after the project is ended?

How much technical/user support will you need?



White papers on TDM

Managing Institutional Knowledge and Insight (white paper and webinar, 2020) tiny.cc/Springer-TDM

Bringing Insight to Data: Info Pros' Role in Text and Data Mining (white paper, 2018) is.gd/springer_tdm



Slides at BatesInfo.com (or email me)

Mary Ellen Bates

mbates@BatesInfo.com

BatesInfo.com

+1 303 772 7095

Facebook / LinkedIn: [maryellenbates](#)