

**Uncertainty
Research,
LLC**

The History of Business Strategy: An NLP Perspective

A natural-language-processing
exploration of the thematic landscape
of the field of Business Strategy,
1980-2020

General Assembly DSI-10 Capstone Briefing

March 13, 2020

Neil A. Hamlett, D.Sc., MBA

Consulting Strategic Data Scientist

neil.hamlett@uncertainty-research.science +1(703)349-7560

 <https://www.linkedin.com/in/neil-hamlett-strategic-quant/>

2651 Park Tower Drive, Vienna, VA 22180 USA

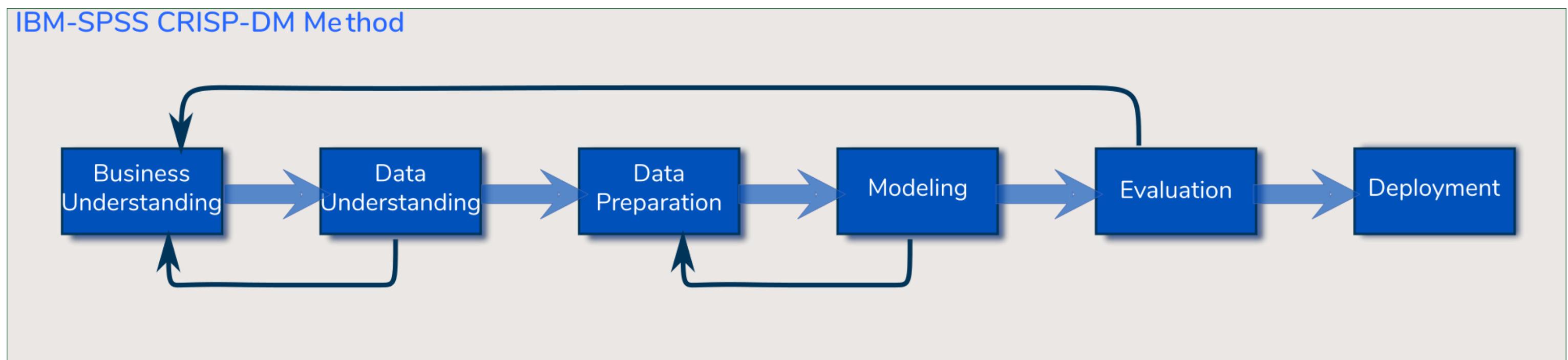
Purpose and Outline.

- Purpose. Describe a study of the application of Topic Analysis to published literature on Business Strategy.
- Outline.
 - Statement of business problem.
 - Approach to study.
 - ◊ Business understanding.
 - ◊ Data understanding.
 - ◊ Data collection.
 - ◊ Modeling.
 - ◊ Model evaluation.
 - Future directions.

Statement of business problem.

- Essential problem. Detect in published business-strategy literature the major themes from the topic.
- Target Audience.
 - Business consultants.
 - Academic researchers.
- Potential benefits.
 - Ability to provide more-insightful guidance to clients.
 - Better understanding of the body of research.
 - Recognition of opportunities to apply methods to other problems.

This study's approach was guided by IBM's CRISP-DM method.

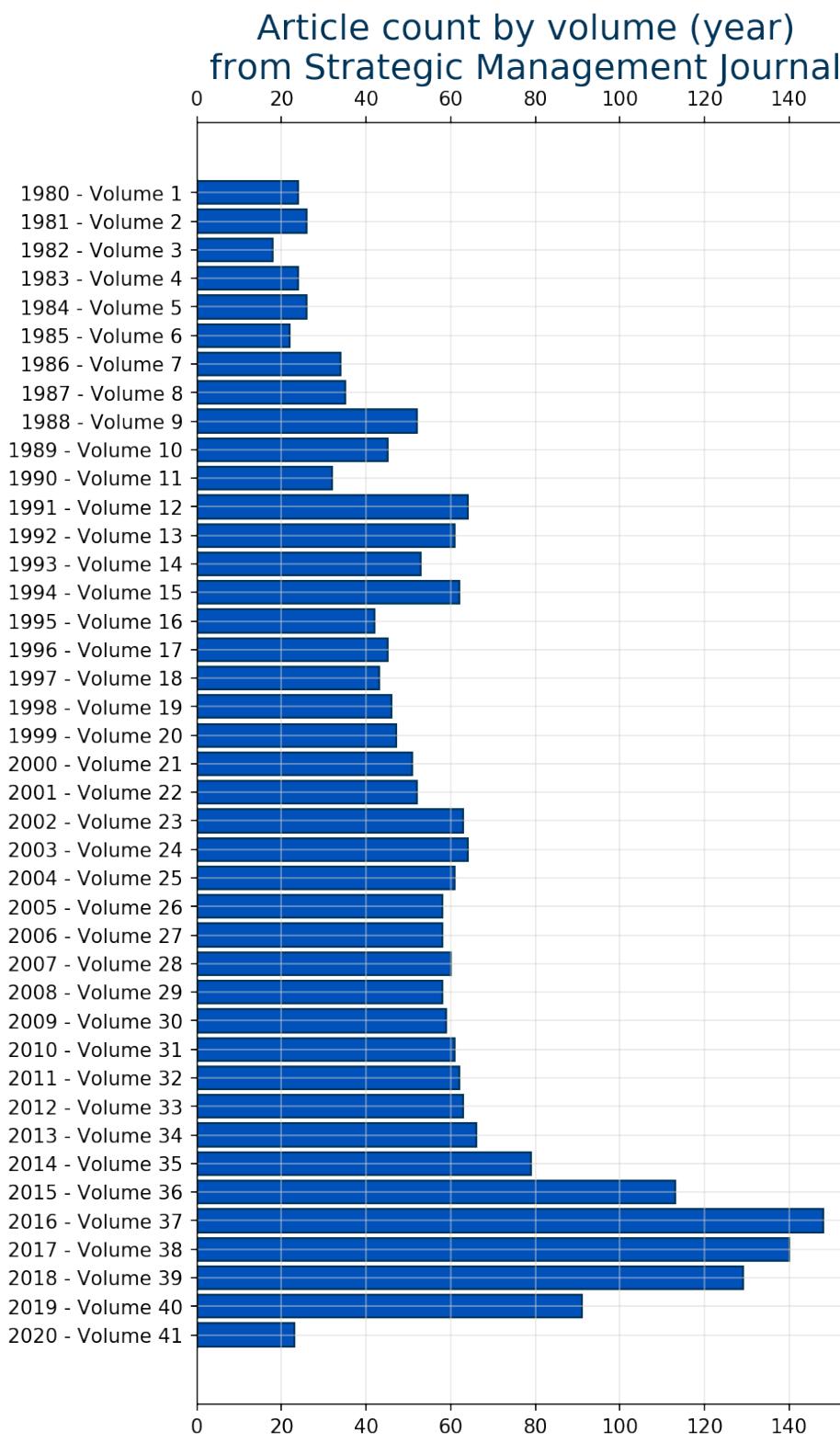


Business understanding: Five key themes dominated the business-strategy discussions over the last century.

	Key Ideas	Early Originators	Period of Prevalance
Operations Research and Management Science	<ul style="list-style-type: none"> ▷ Operational efficiency is the key. ▷ Leads to narrow focus on cost, price. 	<ul style="list-style-type: none"> ▷ Fredrick Taylor. ▷ Patrick Blackett 	Late 1800s to 1980s
Competitive Advantage	<ul style="list-style-type: none"> ▷ Factors other than price can provide bases for differentiation. ▷ Emphasizes industry structure, resources, positioning. ▷ Explains industries in "steady state". 	<ul style="list-style-type: none"> ▷ Michael Porter. 	1980s to early 2000s
Disruptive Innovation	<ul style="list-style-type: none"> ▷ Recognizes that industry structures are fragile. ▷ Describes specific mode in which low-cost firms "eat" incumbents markets from the bottom. ▷ Originator complained that the term "disruption" is overused, misused. 	<ul style="list-style-type: none"> ▷ Clayton Christensen 	Late 1990s to present
Dynamic Capabilities	<ul style="list-style-type: none"> ▷ Adaptability, resilience, agility are keys in industries subject to periodic change. ▷ Describes three key classes of capabilities: Sensing, seizing, managing competitive landscape. 	<ul style="list-style-type: none"> ▷ David Teece 	Early 2000s to present
Information as Key Differentiating Asset	<ul style="list-style-type: none"> ▷ Information, Knowledge began to provide greater return than traditional asset classes as early as 1960s. ▷ Explains power of platform operating models. 	<ul style="list-style-type: none"> ▷ Peter Drucker ▷ Martin Reeve ▷ N. Hamlett 	Emerging

"Uncertainty is a key source of competitive advantage."

Data Understanding: The business-strategy literature is concentrated both in publications and in time.



Candidate Journal	Years Published	Digital Accessibility	Status
<i>Strategic Management</i>	since 1980	Html by leading publishing firm	complete ✓
<i>Strategic Entrepreneurship</i>	since 2007	Html by leading publishing firm	complete ✓
<i>Journal of Strategic Organization (Sage)</i>	since 2003	Html by leading publishing firm	complete ✓
<i>Business Management and Strategy</i>	since 2010	Less than professional-grade html	complete ✓
<i>International Journal of Business Strategy</i>	since 2007	Html since 2012; pdf prior	complete ✓
<i>Journal of Business Strategies</i>	since 2000	Mostly-complete html	complete ✓
<i>California Management Review</i>	Since 1970	Html since 2012; pdf prior	complete ✓
<i>Journal of Business Strategy</i>	since 1980	Professionally-maintained HTML	complete ✓
<i>Strategy Science</i>	since 2016	Html by leading publishing firm	Not attempted: insufficient content ✖
<i>Business and Strategy Development</i>	since 2018	Html by leading publishing firm	Not attempted: insufficient content ✖
<i>Academy of Strategic Management Journal</i>	since 2002	Html since 2017; pdf prior	Not attempted: insufficient html content ⚠
<i>Journal of Strategic Information Systems</i>	since 2001	Html by leading publishing firm	Not attempted: peripheral to subject ⚠
<i>strategy+business</i>	since 1995	Html by leading publishing firm	Pending *
<i>Harvard Business Review</i>	Since 1956	Javascript by leading publishing firm	Not attempted: insufficiently specialized ⚠

“Uncertainty is a key source of competitive advantage.”

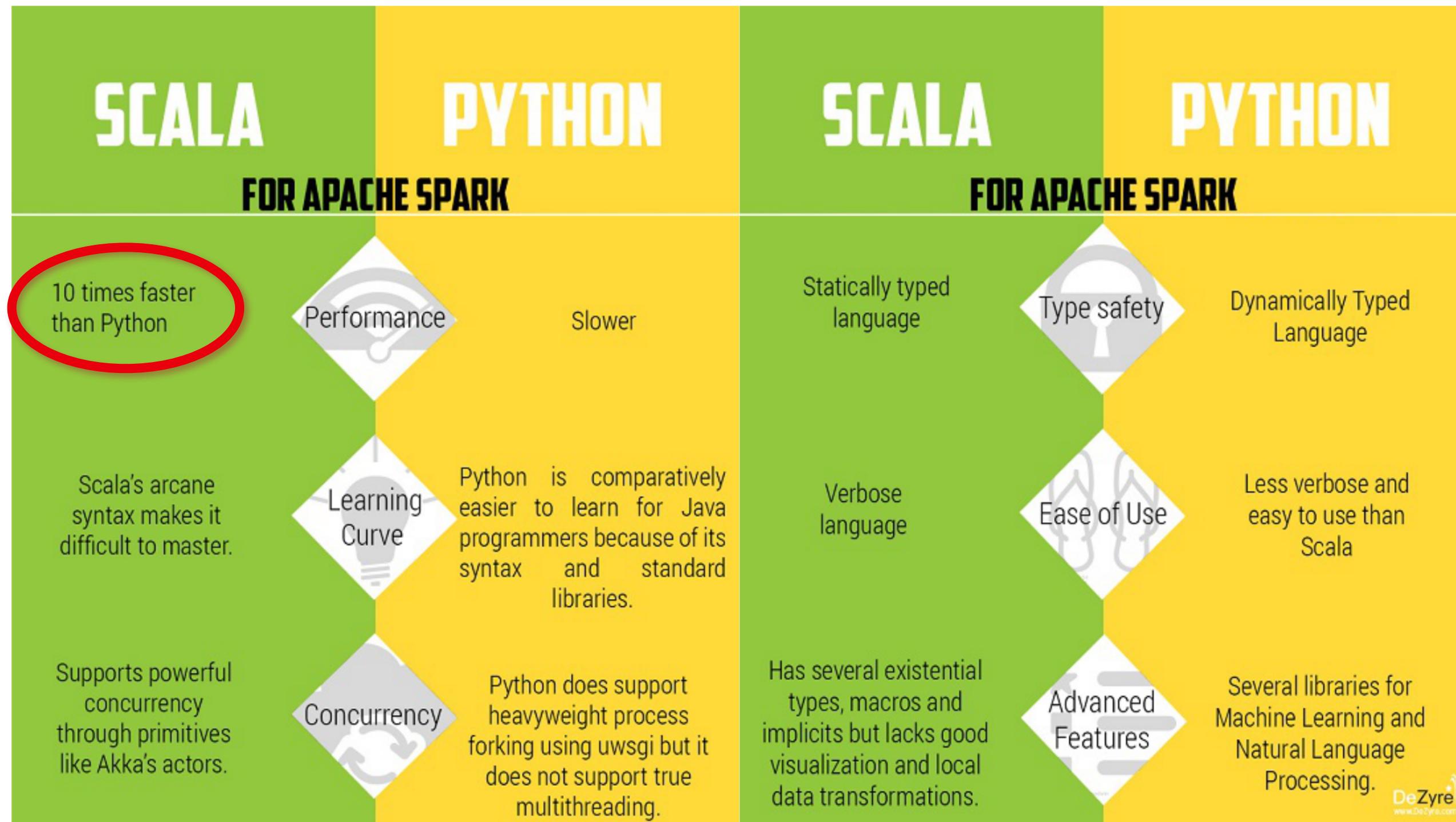
Data Preparation: A mixture of technologies and methods was employed to assemble the corpus.



{JSON}



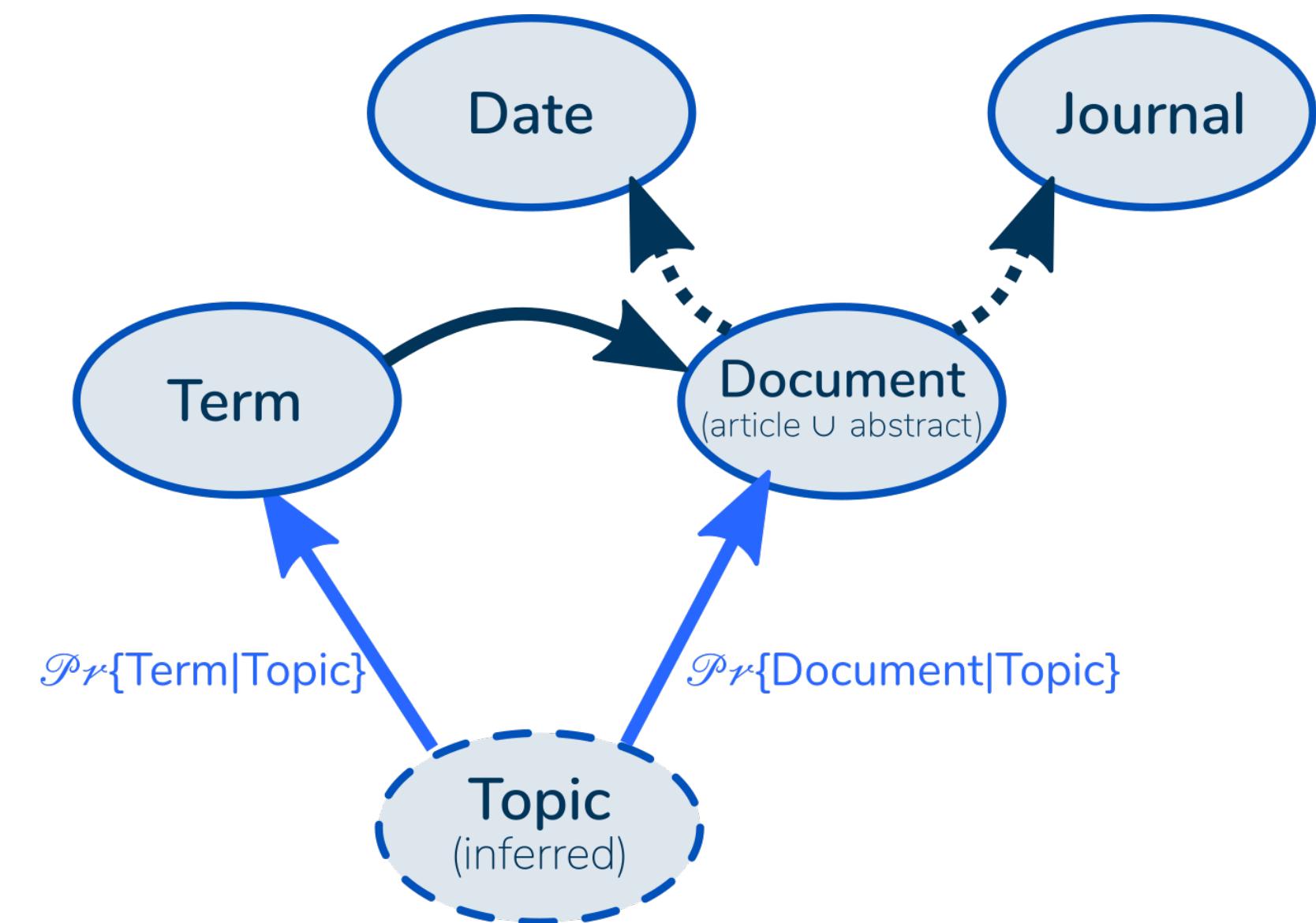
Modeling: The modeling itself was performed in Spark Scala, due to its computational-performance advantages.



After: Scala vs Python for Apache Spark, ProjectPro, February 12, 2016,
<https://www.dezyre.com/article/scala-vs-python-for-apache-spark/213>

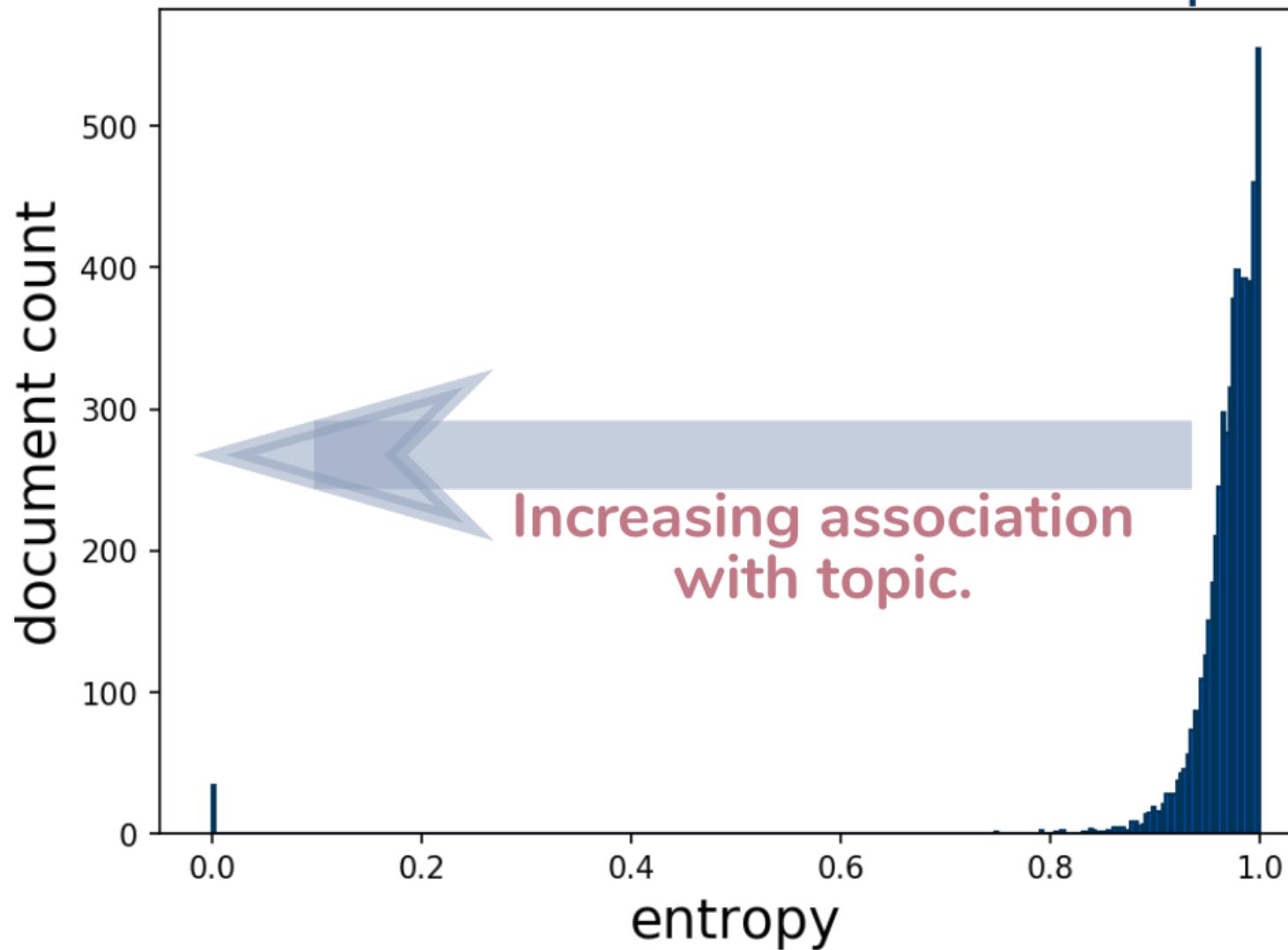
Modeling: Latent Dirichlet Allocation (LDA) — a Bayesian unsupervised-learning method — is commonly used for topic analysis.

- » **Latent:** The Topic attribute is “latent”: Not directly observable.
- » **Dirichlet:** The Topic probabilities are described by a Dirichlet probability distribution.
- » **Allocation:** We allocate our probabilities across all topics, not definitively assigning to one or another.

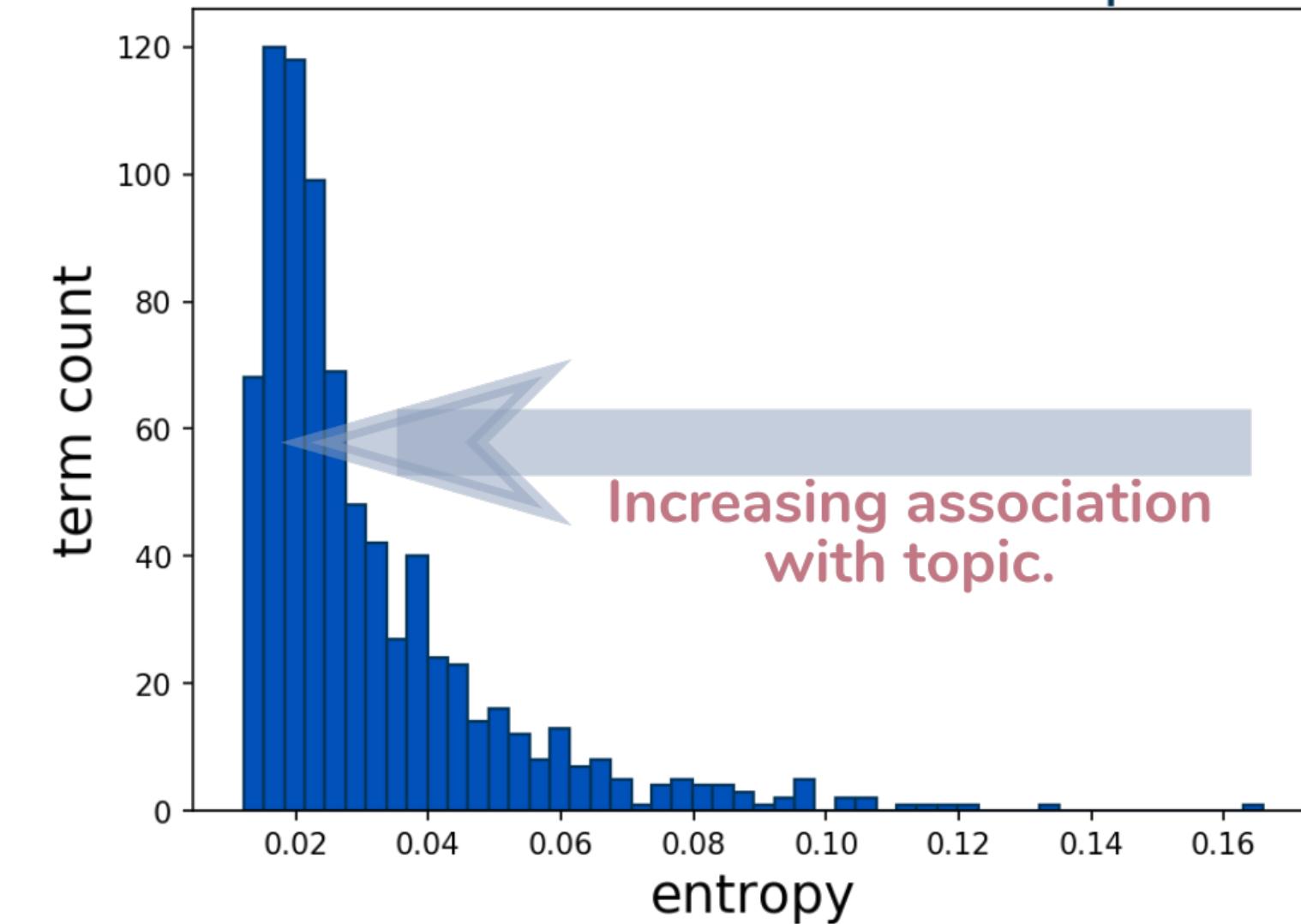


Model Evaluation: Terms (words) are more-strongly associated with topics than the documents they come from.

Distribution of document entropies



Distribution of term entropies



Directions of further work.

March 2020

- Sharpen the model: Incorporate Ngrams, lemmatization.
- Harvest the strategy+business journal.
- Perform closer comparison between topics and major themes.
- Calculate the marginal $\Pr\{\text{Topic}\}$ and calculate cluster “quality” via entropy.

INFORMS Annual Meeting

November 2020

- Harvest full-text journal articles.
- Calculate “perplexity”, statistic described in [\[Brie, et al, 2003\]](#), and compare with entropy.
- Visualize topic clusters’ temporal variation (e.g., violin plot).

Questions?



“About your cat, Mr. Schrödinger—I have good news and bad news.”