MIS Quarterly

SPECIAL ISSUE: NEXT GENERATION IS THEORIES

THEORIES IN FLUX: REIMAGINING THEORY BUILDING IN THE AGE OF MACHINE LEARNING[1]

Monica Chiarini Tremblay
Raymond A. Mason School of Business, William & Mary,
Williamsburg, VA 23187 U.S.A. {Monica.Tremblay@mason.wm.edu}

Rajiv Kohli
Raymond A. Mason School of Business, William & Mary,
Williamsburg, VA 23187 U.S.A. {Rajiv.Kohli@mason.wm.edu}

Nicole Forsgren
GitHub, Inc.
San Francisco, CA 94107 U.S.A. {nicolefv@gmail.com}

Researchers employ methodologies that rely on contemporary technologies to study phenomenon. Recent advances in artificial intelligence (AI), particularly machine learning (ML), have intensified the speed, and our abilities, to create and deploy new knowledge for constructing theories (Abbasi et al. 2016). The availability of big data and ML tools is no longer the sole domain of academics. Business processes generate large amounts of data and now practitioners also deploy ML-enabled methodologies to create new knowledge through working theories.

This presents an opportunity for Information Systems (IS) academics to collaborate with practitioners by addressing their business problems while also creating new theories that have the potential to serve as building blocks toward a generalizable theoretical contribution. As IS is an applied discipline, IS academics must uphold methodological rigor when new technologies, such as ML, offer new methods of knowledge creation.

Both practitioners and academics share the view of theory as abstracted knowledge about the world, and both seek rigor in theory building methods. Academics consider methods to be rigorous when such methods result in parsimonious, generalizable and repeatable[2] theories. Practitioners apply rigorous methods, such as experimentation and A/B testing, to create theories from situated abstractions that can be acted on and can be used to defend their actions. Practitioners value theoretical precision because the consequences of their decisions are costly, for example, in how to price products or when to launch advertising campaigns.[3] They weigh the degree of acceptable uncertainty with the costs of delay. A delay of even one month to act on findings could add up to millions of dollars in lost revenue. Therefore, practitioners are willing to accept greater uncertainty in a theory that is *good enough,* then iterate and improve.

We propose that ML offers an opportunity to reimagine the theory building process. By rapidly generating numerous situated abstractions that can be discarded or refined in pursuit of a generalizable theory, researchers can iterate quickly. As such, we can expect a new form of theory that remains in

[1]This paper was invited and editorially reviewed by the Special Issue Senior Editors: Andrew Burton-Jones, Brian Butler, Susan Scott, and Sean Xin Xu.

[2]United States of America National Institutes of Health (NIH) define scientific rigor as "the strict application of the scientific method to ensure robust and unbiased experimental design, methodology, analysis, interpretation and reporting of results … [so that] others may reproduce and extend the findings" (https://grants.nih.gov/policy/reproducibility/ guidance.htm). It reinforces the notion that rigorous research is parsimonious, generalizable, and repeatable.

[3]We thank the senior editors for this insight.

flux and then evolves. By collaborating with practitioners, academics can apply theories to fill the knowledge gaps when practitioners are unable to rationalize relationships among the variables analyzed. Similarly, practitioners can test and (dis)-confirm academic theories when a technological change results in changes in human behavior, such as how privacy calculus in online shopping explains drivers' willingness to share data in Internet of Things (IoT)-based connected cars. We argue that managerial sense-making, combined with ML, will accelerate our ability to generate new theories that are relevant, rigorous and *good enough* to be useful. We refer to these as theories in flux (TIF).

What Is a Theory in Flux?

We define TIFs as evidence-based inferences that emerge from analyzing large amounts of data or big data, often gathered from business processes and in partnership with practitioners. ML and big data present an opportunity to generate numerous TIFs in several ways. For example, an academic can be an active participant who collaborates with the practitioner organization as an action researcher to develop and refine the theory. Alternatively, the academic may simply obtain data or provide findings and remain as objective and independent as possible. A TIF generally takes shape when a pattern of a phenomenon emerges from the analysis of data. The role of the academic researcher is to further refine TIFs into generalizable theories through engaged scholarship and subsequent scholarly validation. There is a place for, and indeed a need for, targeted and contextual micro-theories that comprise TIF. Under contemporary constraints of academic rigor, TIFs perish along with possible future new theories, resulting in a loss to academics and practitioners.

Practical relevance is the key to a TIF. Therefore, targeted relevance appeals to practitioners and emerges in the form of bounded generalizability, fast feedback, and iterative refinement. These characteristics are inherent in TIFs, and academics should reflect upon their merits to reimagine theory building. In this way, TIFs will augment theory building methods deployed by academics and better explain a phenomenon by providing new paths to discovery that are consistent with academic standards of rigor, relevance, and generalizability.

Machine Learning and TIF

With ML, contemporary organizations can process large amounts of data to create new knowledge about customers' behavior, product quality, and the effectiveness of delivery processes. Practitioners apply ML to big data to model, for example, consumer behavior, to rapidly build TIFs, test them in a redesigned process, then adopt or refine in favor of emergent TIFs that better explain the phenomenon of interest, which in turn improves decision quality. Even when a TIF is discarded, the emergent learning informs future TIFs, just as an unsupported hypothesis informs future hypotheses to advance scientific discovery.

Following a series of interviews with executives of leading digital organizations, we observed the divergence between how contemporary organizations use big data and ML to generate practical insights in the form of TIF, and how IS academics build theory. A vigorous panel discussion at the International Conference on Information Systems and recent journal editorials (for example, Rai 2016) indicate that many academics in the IS discipline hold similar views.

By reimagining theory building, academics can recouple theory building and theory testing that were traditionally viewed as two separate activities. For example, an ML-driven IT artifact enables reciprocal learning between the artifact and the researcher to create new knowledge and build theories. Researchers can interact with ML and quickly explore large data sets and examine a number of relationships to create multiple TIFs. Subsequently, they can discard or refine TIFs, and provide the most promising theory-supported guidance to decision makers, who can then test the efficacy of TIFs in practice. Furthermore, after decision makers take action, the outcomes can be fed back into the model and analyzed for emergent learning.

Some may argue that the iterative methodology is not new and ML is simply a tool that cannot, and should not, replace human imagination in theory building. We agree. However, we propose that the speed with which ML can discover initial patterns, test, and develop theory can pave the way to create relevant and rigorous theories in the future. Furthermore, with reciprocal learning, ML can help uncover obscure features and relationships in a problem setting, expand researchers' imagination and motivate further theory development. Our failure to take advantage of emergent developments, such as ML, risks missing opportunities to make important discoveries (Maass et al. 2018). Therefore, reimagining how to leverage the emerging nexus of big data and ML to build TIF will expand goals to construct actionable and useful theories.

Reimagining Methodological Rigor in TIF

Generalizability: In traditional theory building, generalizability is a core objective. Without it, a theory may apply to only one or a few instances, requiring scarcely available resources to create new knowledge. Generalizability requires theorists to be precise in defining and using constructs and variables. In practice, the application of theory across contexts rarely meets the original definition or assumptions, which then requires adjustments to tailor the theoretical specifications to fit the new context. In the ML and big data environment, generalizability is not a concern, or in the words of a director of a digital organization that utilizes ML capabilities, "generalizability is not a virtue." Organizations that deploy ML find that the costs to customize past theoretical insights to a new context often outweigh the costs of seeking insights tailored to the new context. By reimagining the scope of generalizability, ML can quickly provide a large number of context-specific situated insights for further validation. Multipurpose generalizable theories, practitioners argue, are like a Swiss army knife that offers general guidance in typical contexts, while insights from a TIF are like a scalpel to perform precision surgery, customized to a context. The availability of large datasets and abundant processing power has made it possible to generate quick, precise, custom TIFs for a targeted context.

Replicability: Replicability implies that other researchers should arrive at similar findings when testing theories. Replication builds confidence in the integrity of logical relationships among variables and ensures that findings indeed describe a phenomenon and are not an artifact of the method or the context. For TIFs, replicability of findings across time or contexts isn't a requirement because what matters is fast, useful analysis, *good enough* for making decisions. Indeed, artifacts of the method or context may be a virtue because they can illuminate how different methods provide more precise guidance under certain conditions. However, when the phenomenon is in flux, replicability is neither expected nor desired. For example, property rental company Airbnb must deal with fluctuations in rental markets and with changes in consumer preferences depending upon location, events, and time of the year. In this case, findings of consumer preferences are neither replicable nor desirable because each dataset will produce different and customized theories, or TIFs, for each consumer population. Should a common construct for consumer preferences emerge, TIFs have the potential to coalesce into a generalizable theory while still creating new knowledge relevant for a specific consumer segment.

Parsimony: Academics aim for theories to be parsimonious, that is, demonstrate high explanatory power using the fewest variables or constructs. This criterion overcame resource scarcity because data collection was costly, intrusive, or otherwise difficult, and processing costs were high. For TIFs, the abundance of data, compression of time to capture large amounts of data, and few processing constraints make parsimony no longer necessary.[4] Indeed, including more variables in the analysis leads to deeper insights, and more TIFs, because ML can uncover latent relationships that may *prima facie* appear to be unrelated.

Despite the emerging methods to build TIFs, researchers will continue to play a crucial role in interpreting the analysis (e.g., explainable AI) and to ensure that the findings are not spurious or inadvertently biased.[5] We contrast the traditional and TIF evaluation criteria in Table 1.

In our review of recently published papers, we found that the theory building process is expanding to a rapid, pragmatic theory development that, like TIF, can be quickly and iteratively refined through the use of ML. It is clear that authors, reviewers, and editors of these papers were flexible about generalizability, replicability, and parsimony, so that such papers (four papers discussed here) were published in the leading IS journals. Zhou et al. (2018) and Adamopoulos et al. (2018) used ML (textual analysis) and econometric techniques to propose TIFs. The TIFs are drawn from a specific time frame in an online platform and may not be generalizable to other time frames or platforms, yet the authors uncovered interesting insights for their specific context. Lin et al. (2017) and Dong et al. (2018) used the design science paradigm to extract useful features through ML algorithms. These papers are not framed as theory-building papers per se; however, the authors uncovered new relationships that are consistent with our definition of TIF. Such papers demonstrate that the IS community is likely to benefit from reimagining generalizability, replicability, and parsimony for theory building.

Opportunities and Risks of Theories in Flux

Science embraces discovery. ML advances the discovery process by providing academics with an opportunity to observe phenomena in near real-time and to build TIFs that open new paths to discovery and provide a promising pipeline for the development of future grand theories. We see an example in biology where the electron cryo-microscopy, an emergent tool, provided biologists the opportunity to observe cell activity in real-time and to uncover new paths to discover how cancer progresses.

[4]In ML, parsimony is applied by selecting the simplest model from all possible models that provide similar performance. In this paper, we will use the definition used by IS theorists and outlined in the paper.

[5]For example, Amazon conducted an internal AI study to help speed identification of people to select for interviews. What they found instead— upon human inspection—was that it was biasing against women.

Table 1. Summary of Methodological Evaluation Criteria

Criteria	Traditional Criteria	TIF Criteria
Generalizability	Generalizability is a core objective	Focused, context-specific solution is valued over generalizability.
Replicability	Replicability is required	Replicability is not pursued. The goal is context-specific new knowledge.
Parsimony	Parsimonious solutions are preferred, both for simplicity in analysis and for explanation	ML computational power enables analysis of many features or variables from large datasets and facilitates latent patterns to emerge. Transparency will sometimes be desired.

Timely development of theories encourages their use by practitioners. However, academics must be aware of risks when engaging with practitioners. Practitioners may be unwilling to share proprietary data, their expertise, or access to processes because they fear exposing the organization's intellectual property. Even when practitioners engage with the academic community, they may do so for reasons that serve their business interests, not for scholarly pursuits such as theory building and publications. Academics can mitigate the risks by signing nondisclosure agreements (NDA), agreeing how to assign ownership of intellectual property, discussing how to publish academic findings while protecting propriety information, and by involving professionals from their institutional knowledge transfer office to ensure that the relationship with practitioners is transparent and enduring.[6] We believe that, with such safeguards, the rewards of collaboration are worthy of the potential risks.

A Call to Action

IS academics are uniquely positioned to use ML and TIFs for theory building: we are an applied discipline, with technical proficiency, and the methodological expertise in areas such as grounded theory and design science. The history of the IS discipline is that of solving business problems with IT, and advancing the discovery process as new technologies emerge. Now ML-enabled TIFs have the potential to further that role. And yet, IS academics have been slow to embrace ML in theory building, while other management disciplines such as marketing and finance are already utilizing ML to build theories. Below, we propose three actions that the IS community should take to endorse TIF as a viable method for theory building.

First, to avail opportunities to conceive new theories that impact practice, academics must be able to publish TIF-based research. Journal and conference editors and reviewers must be open to ML-enabled TIFs to promote emergent learning, whether through micro-theories, algorithms, models, or lessons learned. We must be open to TIF's context specificity, reciprocity, and bounded generalizability as acceptable standards of rigor. When a scholarly community determines, say through the review process, that a method demonstrates a logical link between the research question and the answer, the method is accepted as evidence of rigor by the community.

Second, academics and practitioners in the IS community can be pragmatic in adopting evidence-based TIFs, just as medicine and healthcare disciplines adopt proven practices even as controlled trials further refine the findings. This demonstrates the value of sharing emergent knowledge quickly and in ways that are accessible and useful to both academics and practitioners. The IS discipline has thrived when it has adapted methods to leverage emergent tools of data collection and hypothesis testing. A TIF enables IS researchers to combine theory development with contextually informed rapid data collection and analysis to appropriate the benefits of ML.

Third, IS academics can serve as pioneers for other business fields by identifying new and better ways to develop, refine, and manage TIFs. This could include tackling the nontrivial problem of opening the black-box embedded in ML theory development. With proper safeguards to mitigate risks, academics can make ML-based TIFs transparent, explainable, and reversible such that others can rapidly build and test theories that the practitioners can apply with confidence. Generally, IS researchers can take the lead in deploying TIF-related research to accelerate, iterate, and build the cycle of scientific discovery, and to benefit science, in general.

[6]For examples, see https://tlo.mit.edu/ for the Massachusetts Institute of Technology's Technology Licensing Office and https://www.ukri.org/ for UK Research and Innovation described as "works in partnership with universities, research organisations, businesses, charities, and government to create the best possible environment for research and innovation to flourish."

References

Abbasi, A., Sarker, S., and Chiang, R. H. 2016. "Big Data Research in Information Systems: Toward an Inclusive Research Agenda," *Journal of the Association for Information Systems* (17:2), pp. i-ii.

Adamopoulos, P., Ghose, A., and Todri, V. 2018. "The Impact of User Personality Traits on Word of Mouth: Text-Mining Social Media Platforms," *Information Systems Research* (29:3), pp. 612-640.

Dong, W., Liao, S. Y., and Zhang, Z. J. 2018. "Leveraging Financial Social Media Data for Corporate Fraud Detection," *Journal of Management Information Systems* (35:2), pp. 461-487.

Lin, Y. K., Chen, H. C., Brown, R. A., Li, S. H., and Yang, H. J. 2017. "Healthcare Predictive Analytics for Risk Profiling in Chronic Care: A Bayesian Multitask Learning Approach," *MIS Quarterly* (41:2), pp. 473-495.

Maass, W., Parsons, J., Purao, S., Storey, V. C., and Woo, C. 2018. "Data-Driven Meets Theory-Driven Research in the Era of Big Data: Opportunities and Challenges for Information Systems Research," *Journal of the Association for Information Systems* (19:12), pp. 1253-1273.

Rai, A. 2016. "Editor's Comments: Synergies Between Big Data and Theory," *MIS Quarterly* (40:2), pp. iii-ix.

Zhou, S. H., Qiao, Z. L., Du, Q. Z., Wang, G. A., Fan, W. G., and Yan, X. B. 2018. "Measuring Customer Agility from Online Reviews Using Big Data Text Analytics," *Journal of Management Information Systems* (35:2), pp. 510-539.