



# Guest Editorial: Recent Trends in Reuse and Integration

Thouraya Bouabana-Tebibel<sup>1</sup> · Stuart H. Rubin<sup>2</sup> · Lydia Bouzar-Benlabiod<sup>1</sup>

Published online: 12 February 2019

© Springer Science+Business Media, LLC, part of Springer Nature 2019

## 1 Introduction

The term, “reuse” means far more than code reuse, which it subsumes. It includes all manner of abstractions – from models to dynamic plans, simulations, knowledge bases, software, firmware, and hardware. In particular, the term includes semantic repositories, which enable the reuse of knowledge stored in episodic memory, rules, neural weight sets, connections in hardware, and the like. A problem with deep learning and neural networks in general is that they don’t facilitate transfer learning, which is a very human way to learn. (Hosseini et al. 2017) recently demonstrated that unlike humans, neural networks are unable to learn to recognize images that are equivalent under simple transformation (e.g., integers and their Polaroid negatives). Without transfer learning, neural networks cannot take advantage of randomization for accelerating their learning and increasing the space of what they know. Moreover, (Lin and Vitter 1991) proved that if a neural network has at least one hidden layer, then it is *NP-hard* to train. Geoffrey Hinton, the inventor of the backpropagation algorithm, has stated that we need to reinvent the neural net if we ever hope to close in on how the brain works and the capability to emulate its diverse abilities.

(Feigenbaum 1984) has written on the knowledge acquisition bottleneck, which is to say that we need to acquire more knowledge with greater alacrity if we are to ever hope to build scalable knowledge-based systems with acceptable ROIs. Semantic repositories are a step in the right direction. Such repositories hold generalized knowledge, which is created through the randomization of instances. The space of recoverable instances then exceeds the original space from which the generalizations spring. In this manner, semantic

repositories support machine intelligence from the perspective of reuse and integration (Fraga et al. 2019). Machine intelligence underpins the AI revolution that started with deep learning and is not limited to any one technology or set of technologies. Machine intelligence depends on software and parallel hardware – the construction of which is dependent upon information reuse and integration. Again, abstraction is key to successful reuse and integration. Abstraction, in turn, depends upon knowledge, which is bootstrapped using various abstractions. We see that information reuse and integration underpins every significant intelligent system ever built or which can ever be conceived (Chang et al. 2019). Indeed, Rubin’s theory of randomization (Rubin 2007), formally predicts that with scale, abstraction allows for an unbounded density of knowledge to be achieved in computer memory. The open question is what density is associated with what scale. Information reuse and integration may have got its start with code reuse, but it is rapidly heading towards knowledge reuse and beyond. That is, designs inherently involve imagination, which can be traced to reuse and domain transference. Indeed the *knowledge revolution* has just begun and will come to vastly overshadow its predecessor – the *industrial revolution*.

## 2 In this Special Issue

Data science has become valuable for prediction – a key capability of AI. Big data mining is based on statistics, but formal, or small data mining, is based on fitting an algorithm or equation to causality. Both approaches are needed to cover the gamut of complexity in the real world. Technically speaking, we say that randomizations are recursively enumerable, but not recursive. This means that search is an inherent part of the non-trivial discovery process. One cannot prove that a better randomization does not exist – only search to see if one can be found.

Another aspect of reuse are the modeling languages, such as UML and XML (Meghzili et al. 2019). The problem with these abstractions is that they reportedly don’t scale well. Thus, researchers are attempting to incorporate Apache Storm and stochastic (colored) Petri Nets to enhance

✉ Lydia Bouzar-Benlabiod  
l\_bouzar@esi.dz

<sup>1</sup> Laboratoire de la Communication dans les Systèmes Informatiques, Ecole nationale Supérieure d’Informatique, BP 68M, 16309, Oued-Smar, Alger, Algérie

<sup>2</sup> Space and Naval Warfare Systems Center Pacific, San Diego, CA 92152-6423, USA

scalability (Requeno et al. 2019). An open question is if it would not be more cost effective to pursue Fourth or Fifth generation (functional) languages as modeling tools in preference to UML. Modular compilation, extensibility, explanation facilities, capabilities for parallel and cyber-secure computation, GUIs, debugging facilities, and readability all factor into the definition of successful modeling languages.

Another logistics problem, faced by the military and certain industries, pertains to the coordination of robotic swarms (Leofante et al. 2019). This is a planning problem, which involves reuse and plan transformation or modification. At this time, there are no effective theories for efficient distributed swarm control. Akin to natural language processing, machine learning seems to have the most to offer here in the way of a solution. Deep learning has been the most successful approach to date; but, deep learning is bottlenecked because it is incapable of transfer learning, which is required for reuse to yield a high ROI here. Similarly, one must be careful not to create bloat and reuse is consistent with optimization when properly done. Deep learning depends on hidden layers – the use of which precludes tractable reuse. For this reason, Professor Geoffrey Hinton is actively seeking a substitute for the backpropagation algorithm, which he invented.

Another popular area for reuse of late is in text processing (Prusa et al. 2019). In the past, we relied on data being stored in relational databases along with SQL, relational algebras, or other access formalisms for its manipulation. Now, we are moving from keyword search to textual understanding and the use of those semantics to not only search for constrained information, but to recognize and predict patterns as well (e.g., predicting the price of West Texas oil futures). As computer speeds increase and the use of parallel processing continues to grow, we can expect natural language driven human interfaces to get better. One application has members of a drone swarm communicating with each other – not in code, but using synthesized natural language. One study predicts that natural language processing will grow to be a \$1.25 billion dollar industry (e.g., for marketing) by 2022. Information reuse and integration play a pivotal role here because text and linguistic semantics have a many to one relation, or a many to many relation in context. Bidirectional recurrent neural networks have been developed as effective real-time language translators and research is exploring their possible use as chat-bots. At present, they cannot scale beyond several lines of chat. Scalability here implies solving the transfer learning problem – perhaps by reinventing the neural network. That task again hinges on successful abstraction through reuse and integration.

One way to enhance scalability is through dimensionality reduction – eliminating irrelevant details. For example, we recognize automobiles despite variances in their manufacturer, model, and year. This differs from data cleansing, which is tasked to minimize noise. Moreover, features can be broken down into reusable parts or fractal equations (Golinko and

Zhu 2019). The problem of transformational learning persists if performed using hidden-layer neural networks. However, supervised symbolic transference is possible today and is appropriate for many domains that can be so labeled.

One of the problems associated with any reuse-based ecosystem pertains to metricizing the salient features (Lima Fontão et al. 2019). For example, if one is trying to optimize a composition of functions, one must be concerned with space, speed, and reliability. How does one determine the best tradeoffs? How will these decisions factor into subsequent compositions? How can these and other decision-making processes be accurately captured for reuse for unrelated problems? How does one go about recognizing that a particular optimization may not be the best for a particular problem? Can one develop an expert system to make the configurations? Can knowledge base segments be reused? How are the rules acquired? Should cases be used in preference to rules? These and a plethora of other questions need to be answer in order that information reuse and integration can advance our intelligent systems of tomorrow.

Finally, there is work going on pertaining to the use of schemas for knowledge extraction (e.g., for web query interfaces) (Jou 2019). Schemas are randomizations, which may be instantiated to create constrained knowledge. The knowledge must be constrained because otherwise it is likely to be in error. The key to scalable schema-based systems is to have knowledge self-apply. This means that the generalization of instances into tractable schemas (under the inverse triangle inequality) and the instantiation of those schemas to yield knowledge must be knowledge-based to be scalable. This defines a network and falls under Rubin's theory of randomization. Not only is the data reusable, but the pattern of knowledge segment calls is too. Symbolic transference is currently realizable. The last paper presents an example using heuristic rules to extract schemas for deep web query interfaces.

### 3 Conclusion

In summary, we have just begun to see how intelligent systems will be developed to change the world for the better. We see that information reuse, integration, and their abstractions, as detailed in this book, will play a fundamental role in the realization of these intelligent systems.

**Acknowledgment** The goal for this special issue is to survey recent trends in reuse and integration by way of presenting nine select papers, which together cover the recent developments. Drs. Rubin and Bouzar-Benlabiod would first like to avail themselves of the opportunity to thank the late Professor Bouabana-Tebibel – without whom there would be no special issue. Her foresight into the need to compile a special issue such as this one was not only insightful, but inspirational.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## References

- Chang, V., Abdel-Basset, M., & Ramachandran, M. (2019). Towards a reuse strategic decision pattern framework – From theories to practices. *Recent Trends in Reuse and Integration. Information Systems Frontiers*, 21(1). <https://doi.org/10.1007/s10796-018-9853-8>.
- Feigenbaum, E. A. (1984). Knowledge engineering. *Annals of the New York Academy of Sciences*, 426(1), 91–107.
- Fraga, A., Llorens, J., & Genova, G. (2019). Towards a Methodology for Knowledge Reuse Based on Semantic Repositories. *Recent Trends in Reuse and Integration. Information Systems Frontiers*, 21(1). <https://doi.org/10.1007/s10796-018-9862-7>.
- Golinko, E., & Zhu, X. (2019). Generalized Feature Embedding for Supervised, Unsupervised, and Online Learning Tasks. *Recent Trends in Reuse and Integration. Information Systems Frontiers*, 21(1). <https://doi.org/10.1007/s10796-018-9850-y>.
- Hosseini, H., Xiao, B., Jaiswal, M., & Poovendran, R. (2017, December). On the limitation of convolutional neural networks in recognizing negative images. In *Machine Learning and Applications (ICMLA), 2017 16th IEEE International Conference on* (pp. 352–358). IEEE.
- Jou, C. (2019). Schema extraction for deep web query interfaces using heuristics rules. *Recent Trends in Reuse and Integration. Information Systems Frontiers*, 21(1). <https://doi.org/10.1007/s10796-018-9863-6>.
- Leofante, F., Ábrahám, E., Niemueller, T., Lakemeyer, G., & Tacchella, A. (2019). Integrated Synthesis and Execution of Optimal Plans for Multi-Robot Systems in Logistics. *Recent Trends in Reuse and Integration. Information Systems Frontiers*, 21(1). <https://doi.org/10.1007/s10796-018-9858-3>.
- Lima Fontão, A., Santos, R., & Dias-Neto, A. (2019). Exploiting repositories in Mobile software ecosystems from a governance perspective. *Recent Trends in Reuse and Integration. Information Systems Frontiers*, 21(1). <https://doi.org/10.1007/s10796-018-9861-8>.
- Lin, J. H., & Vitter, J. S. (1991). Complexity results on learning by neural nets. *Machine Learning*, 6(3), 211–230.
- Meghzili, S., Chaoui, A., Strecker, M., & Kerkouche, E. (2019). Verification of model transformations using Isabelle/HOL and Scala. *Recent Trends in Reuse and Integration. Information Systems Frontiers*, 21(1). <https://doi.org/10.1007/s10796-018-9860-9>.
- Prusa, J. D., Sagul, R. T., & Khoshgoftaar, T. M. (2019, 2018). Evaluating Text Feature Extraction of Expert Reports for the Valuation of West Texas Intermediate Crude Oil Futures. *Recent Trends in Reuse and Integration. Information Systems Frontiers*, 21(1). <https://doi.org/10.1007/s10796-018-9859-2>.
- Requeno, J., Merseguer, J., Bernardi, S., Perez-Palacin, D., Giotis, G., & Papanikolaou, V. (2019). Quantitative analysis of apache storm applications: The NewsAsset case study. *Recent Trends in Reuse and Integration. Information Systems Frontiers*, 21(1). <https://doi.org/10.1007/s10796-018-9851-x>.
- Rubin, S. H. (2007). On randomization and discovery. *Information Sciences*, 177(1), 170–191.

**Prof. Thouraya Bouabana Tebibel** received a Ph.D. in Computer Science from USTHB University (Algeria) in collaboration with Pierre & Marie Curie University (France) in 2007. She has been an engineer/researcher for eight years and then was a full professor of computer science at Ecole nationale Supérieure d'Informatique in Algeria. She was the Director of the Doctoral School in Sciences and Technologies of Information and Communication for four years and director of research at Laboratoire de Communication des Systèmes Informatiques. She has over 90 refereed publications and a book edited by Editions Universitaires Européennes. She was co-editor (with S. H. Rubin) of three Springer books in the *Advances in Intelligent System and Computing* series.

**Dr. Stuart H. Rubin** is a senior scientist at the Space and Naval Warfare Systems Center (SSC) in San Diego, code 71740 (Advanced Concepts & Applied Research). He was previously a tenured associate professor of computer science at Central Michigan University (CMU). He received a Ph.D. in Computer and Information Science from Lehigh University in 1988. He was previously an ONT Post-Doctoral Fellow, at NOSC, for three years. He has over 35 assigned Navy patents, over 295 Refereed Publications, and received SSC-PAC's Publication of the Year Awards in 2007, 2009, 2010, and 2011. He was co-editor (with T. Bouabana Tebibel) of three Springer books in the *Advances in Intelligent Systems and Computing* series. Dr. Rubin is a SIRI Fellow and serves in leadership roles in numerous IEEE technical societies.

**Dr. Lydia Bouzar-Benlabiod** is a researcher at Laboratoire de Communication des Systèmes Informatiques (LCSI), Ecole nationale Supérieure d'Informatique (ESI, Algeria). She received a Ph.D. degree in Computer Science from ESI and from Université d'Artois (France) in 2015. She received a magisterial degree from ESI in 2010. She is a member of The International *Society* of Applied Intelligence.