Editor: **Sarah Gregory**
Intel Corporation
sarah.c.gregory@ieee.org

Requirements Engineering in the Days of Artificial Intelligence

Fabiano Dalpiaz and Nan Niu

ARTIFICIAL INTELLIGENCE (AI) has a long tradition in software and requirements engineering (RE). Over the years, many AI techniques have been employed to represent and analyze requirements, ranging from knowledge representation and reasoning in the 1980s to the use of natural language (NL) processing, machine learning, and deep learning since the 2000s. AI techniques have been successfully applied in practice, for example, to manage large-volume requirements.[1]

The Artificial Intelligence and Requirements Engineering Workshop (AIRE) is the go-to venue for researchers and practitioners who are interested in the intersection of AI and RE. The workshop is conceived as a discussion-oriented event that could foster the cross-fertilization of ideas regarding how AI techniques can be applied to the RE discipline, and how RE techniques can be built to assist the creation of AI-driven systems.

The first instance of AIRE dates back to 2014: Nelly Bencomo, Jane Cleland-Huang, Jin Guo, and Rachel Harrison organized the first edition colocated with the 2014 IEEE International Requirements Engineering Conference in Karlskrona, Sweden. After the first successful event, the second edition took place at RE 2015 in Ottawa, Canada; the participants identified a few grand challenges, including how automated techniques can increase quality in RE. By welcoming two additional organizers, the 2016 edition (Beijing, China) had that challenge as a special theme: do AI-based/automated techniques lead to quality improvement in practice?

AIRE 2017 brought deep learning to the attention of the community: Vincenzo Gervasi's keynote reviewed how different neural architectures could fit different RE problems. Furthermore, AIRE 2017 also revealed an overlap with the crowd-based RE community,[2] leading to the two workshops (AIRE and CrowdRE) joining forces: AIRE 2018 was co-organized by members from both communities and had CrowdRE as its theme. The two keynote speakers represented each of these perspectives: while Lionel Briand reported on successful attempts to apply AI techniques to solve industrial needs, Brian Fitzgerald explained when crowdsourcing can (and should not!) be used as an inexpensive, human-driven alternative to AI-based approaches.

For the past few editions, AIRE has been consistently the highest attended workshop at the IEEE RE conference. The workshop continuity is ensured through a steering committee. Along with presenting cutting-edge research results, AIRE serves as an instrument to raise awareness and to increase the knowledge of AI within the RE realm.

AIRE 2019

The most recent edition of the AIRE workshop took place in Jeju Island, South Korea. This one-day workshop was held on Tuesday, 24 September 2019, as part of the RE 2019 conference week (http://re19.ajou.ac.kr/). Like previous years, the AIRE workshop attracted multiple participants (over 20 people including both program cochairs of RE 2019 in the opening session). We summarize the workshop and highlight the unique challenges that AI and RE pose on each other.

Environment-Centric Intelligence

In his seminal paper, Michael Jackson wrote "Requirements are located in the 'environment,' which is distinguished from the 'machine' to be

Digital Object Identifier 10.1109/MS.2020.2986047
Date of current version: 18 June 2020

built."[3] Conventionally, AI is sometimes called *machine intelligence* to show that machines can algorithmically learn to group objects, to perform classifications, to recognize patterns, to make inferences, and so forth.[4] As machines become increasingly capable (e.g., in terms of processing increasingly large amounts of data), the views evolve in contrasting AI to the natural intelligence displayed by humans. Now that more intelligent machines are built, what about requirements regarding machine intelligence?

Zhi Jin, a prominent researcher in RE, delivered the keynote "Environment-Centric Self-Adaptivity for Autonomous Systems in Smart Spaces," at AIRE 2019. In this talk, Jin elaborated the environment with the smart space that has or shows quick-witted intelligence when responding to the inhabitants in it. Taking the autonomous unmanned systems (AUSs) as an example, the space will be smart by combining perception, cognition, communication, and actuation to sense and operate the reality. When reality is uncertain, it leads to RE challenges in architecture (control loops, monitoring-planning-analysis-execution and a shared knowledge base, and so forth), environment modeling, and situational awareness (perception, comprehension, projection). What are the steps toward RE for intelligent spaces? Read Jin's RE@Next! work in which environment-centric safety requirements ("hazard-elimination" and "conflict avoidance" in particular) are explored and the implications to intelligent AUSs are discussed.[5]

AI for RE

Many AIRE 2019 papers investigated how to develop machine learning and deep learning techniques for RE tasks.[6] Mishra and Mishra used word embeddings for identifying domain-specific ambiguities in NL requirements. Hayes and colleagues presented the use of metadata, such as readability indexes, as a resource for requirements traceability. Fischbach and colleagues discussed the generation of test models from semistructured requirements. Saito introduced a novel process mining technique that can be used to identify stakeholders from business processes execution logs.

Stanik and his colleagues from the University of Hamburg presented one of the most active applications of AI solutions in RE: classifying app reviews to better make sense and use of the user opinions shared in social media. They compared traditional machine learning with deep learning when classifying app reviews and users' tweets into problem reports (potential bugs), inquiries (feature requests), and being irrelevant (noisy feedback). What is the new challenge? Multilingual text due to a project in which the need for analyzing user feedback written in English and Italian arose. Training the models based on 5 million app reviews in English, 5 million tweets in English, and 1.3 million tweets in Italian, they reported that traditional machine learning (e.g., random forest and decision tree) had slightly better classification accuracy than deep learning (e.g., convolutional neural network). Their research also demonstrated that *fastText*, compared to *word2vec*, was more suitable for learning high-dimensional vector representations not only in multilingual settings but also from user-feedback data in which spelling mistakes and transitory terms (Twitter hashtags or emoticons) are commonly encountered. The thread of work on AI for RE, therefore, has made strides in rigorously evaluating the performance of machine learning and deep learning solutions as well as in investigating how general-purpose AI tools can be tailored best for RE tasks.

RE for AI

Although RE presents unique challenges to AI, the converse also holds. As the AIRE workshop evolves, more and more work has begun raising very interesting questions about how RE will contribute to AI. In the instance of AIRE that was held on Jeju Island, Belani and colleagues reviewed the practical challenges that exist in building AI-based systems, which include the lack of suitable development processes. Rahimi and colleagues presented an interesting approach, which was incorporated by Marsha Chechik (coauthor of the AIRE paper) in her conference keynote "Uncertain Requirements, Assurance, and Machine Learning."[7]

Needless to say, vehicle intelligence (e.g., advanced driving assistance systems) must encompass a pedestrian-detection component. Irrespective of what AI solutions are employed for the detection, it turns out that "pedestrian" belongs to the class of hard-to-specify domain concepts, e.g., there is a lack of requirements specifications defining what pedestrian means and how a pedestrian detector will behave. To that end, Rahimi and her colleagues adapted the web-mining approach to augmenting hard-to-retrieve traces and built a visual understanding of pedestrian and related domain terms.[8] What are the surprising results? A benchmark data set on pedestrian detection underspecified "pedestrian in a wheelchair," and it might consequently hurt the machine's proper learning and detection. The work illustrates the important role that requirements specifications play in building dependable AI components in the fast-evolving yet uncertain smart spaces.

Reflections and Visions

To reflect the past six editions of AIRE, we used an online tool (https://www.wordclouds.com/) to generate a word cloud for each year. Specifically, we fed the tool all the abstracts and keywords of the AIRE papers of a given year while manually tagging the domain terms (e.g., "data mining" and "deep learning"). Figure 1 shows these word clouds.

Using the visuals of Figure 1 as a way of learning from AIRE's own history, we make the following observations, along with some visions and challenges for the future.

- Requirements elicitation is probably the RE task that AIRE papers have studied more extensively. One prominently explored task within elicitation is ambiguity detection, which is typically supported via a combination of natural language processing and machine learning. Another topic that received a great deal of attention is the establishment and maintenance of traceability links. A third task that stands out is that of modeling, whether in specific domains, such as the automotive or utilizing notations that borrow primitives from cognitive sciences such as goal modeling.
- Automated classification remains one of the most mature AI techniques applied to tackle RE problems (e.g., classifying rule sentences from business documents, distinguishing functional and nonfunctional requirements, separating bugs from feature requests). Not only was traditional machine learning applied, but deep learning also found its applications in AIRE, e.g., convolutional neural networks (2016), long short-term memory (LSTM) and recurrent neural networks (RNNs) (2017), RNNs (2018), and bidirectional LSTM (2019). However, deep learning is data hungry, which calls for community effort in data sharing, curation, and provenance. More work on understanding the data is necessary, as emphasized by Hayes and colleagues in AIRE'19.
- The automated analysis of user feedback, part of the CrowdRE paradigm in which end users are active contributors to engineering requirements,[2] is gaining traction. In particular, the papers presented at AIRE have investigated how to automatically analyze and organize requirements from user reviews and user forums.

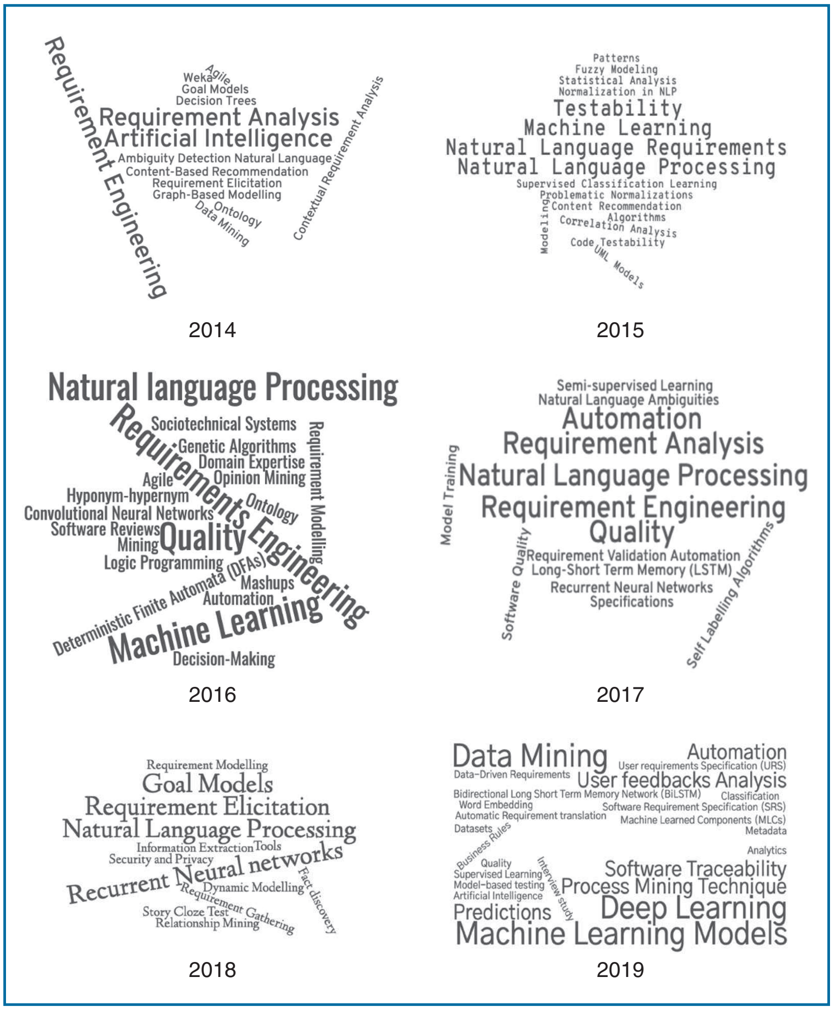

FIGURE 1. The word clouds of the past AIRE editions.

ABOUT THE AUTHORS

FABIANO DALPIAZ is an assistant professor of software systems at Utrecht University's Department of Information and Computing Sciences. Contact him at f.dalpiaz@uu.nl.

NAN NIU is an associate professor in the University of Cincinnati's Department of Electrical Engineering and Computer Science. Contact him at nan.niu@uc.edu.

- Quality is a two-faceted concept in the field. On the one hand, *requirements quality* is one of the objectives of automation, such as in the case of ambiguity detection that was mentioned previously, or a precondition for the generation of test cases. On the other hand, *quality requirements* (roughly, the nonfunctional requirements) can be examined through AI techniques that can support trade-off analysis and negotiation.

AI and RE need a continuous dialog, and AIRE provides such a focused forum to facilitate the exchanges among researchers and practitioners. Although AI helps advance the state of the art of RE in many fronts, requirements engineers have much to offer to the current AI wave. In their paper presented at AIRE'19, Vogelsang and Borg interviewed machine learning experts and showed that performance measures would be good functional requirements to begin with for AI systems, yet we shall be aware of new quality requirements, such as explainability, freedom from discrimination, and specific legal requirements (particularly those regulating the deployment of AI in safety-critical contexts). Can't wait for AIRE updates? Please check out the 2020 edition of the workshop at https://aire-ws.github.io/aire20/.

Acknowledgments

We thank Navjot Singh from the University of Cincinnati for assisting in the word cloud analysis. We also thank Jaspreet Bhatia and Pradeep Murukannaiah for co-organizing the AIRE 2019 workshop as well as the support that we received from the AIRE Steering Committee.

References

1. J. N. och Dag, B. Regnell, V. Gervasi, and S. Brinkkemper, "A linguistic-engineering approach to large-scale requirements management," *IEEE Softw.*, vol. 22, no. 1, pp. 32–39, 2005. doi: 10.1109/MS.2005.1.
2. E. C. Groen et al., "The crowd in requirements engineering: The landscape and challenges," *IEEE Softw.*, vol. 34, no. 2, pp. 44–52, 2017. doi: 10.1109/MS.2017.33.
3. M. Jackson, "The meaning of requirements," *Ann. Software Eng.*, vol. 3, pp. 5–21, 1997. doi: 10.1023/A:1018990005598.
4. Wikipedia, "Artificial intelligence," 2020. [Online]. Available: https://en.wikipedia.org/wiki/Artificial_intelligence
5. Y. Luo, Y. Yu, Z. Jin, and H. Zhao, "Environment-centric safety requirements for autonomous unmanned systems," in *Proc. 2019 IEEE 27th Int. Requirements Engineering Conf. (RE)*, Jeju Island, South Korea, 2019, pp. 410–415.
6. J. Bhatia, P. K. Murukannaiah, N. Niu, and F. Dalpiaz, "The 6th International Workshop on Artificial Intelligence for Requirements Engineering," in *Proc. 2019 IEEE 27th Int. Requirements Engineering Conf. (RE)*, Jeju Island, South Korea, 2019, p. 213.
7. M. Chechik, "Uncertain requirements, assurance and machine learning," in *Proc. 2019 IEEE 27th Int. Requirements Engineering Conf. (RE)*, Jeju Island, South Korea, 2019, pp. 2–3. doi: 10.1109/RE.2019.00010.
8. M. Gibiec, A. Czauderna, and J. Cleland-Huang, "Towards mining replacement queries for hard-to-retrieve traces," in *ASE 2010*, Antwerp, Belgium, 2010, pp. 245–254. doi: 10.1145/1858996.1859046.