Approaches to Realize the Potential of Autonomous Underwater Systems in Concept Development and Experimentation

Thomas Mansfield(✉), Pilar Caamaño Sobrino, Arnau Carrera Viñas, Giovanni Luca Maglione, Robert Been, and Alberto Tremori

NATO STO Centre for Maritime Research and Experimentation, Viale San Bartolomeo, 400, La Spezia, Italy
{thomas.mansfield, pilar.caamano, arnau.carrera, giovanni.maglione, robert.been, alberto.tremori}@cmre.nato.int

Abstract. Recent NATO reports highlight the rapid progress being made in the development of autonomous underwater systems. In contrast, national reports indicate that their benefits are not being fully realized in a timely manner in operational scenarios. One approach to improve NATO's adoption of these systems is to provide guidance in the NATO concept development and experimentation process specially aimed at articulating autonomous system behaviors and allowing efficient experimentation with their capabilities. This position paper reviews the latest techniques and approaches for articulating and testing autonomous system capabilities in industry, academia and within NATOs national militaries. Discussed techniques focus on encouraging and developing understanding and trust in the commander and operator stakeholder communities as well improving the efficiency of autonomous system testing. Potential future guidance and the structure of these activities within the existing NATO CD&E framework are presented for further discussion.

Keywords: Autonomous underwater vehicles · Concept development and experimentation · Design of experiments · Virtualization · Mine countermeasure

1 Introduction

Advances in sensors, robotics and computing are allowing the development of advanced autonomous systems [1]. These systems offer a wide range of military benefits including the ability to conduct missions in remote and hostile environments without placing personnel in harm's way as well as new human-machine teaming concepts [2]. The benefits of autonomous systems are particularly apparent in the underwater domain, where hazardous activities such as mine counter measure missions must be conducted in uncertain environments with limited or no in-mission Command and Control (C2) infrastructure [3].

To develop new operational capabilities, NATO typically uses its Concept Development and Experimentation (CD&E) process [4]. While this process has been

J. Mazal (Ed.): MESAS 2018, LNCS 11472, pp. 614–626, 2019.
https://doi.org/10.1007/978-3-030-14984-0_46

designed to incorporate a wide range of technologies, the increased technical and conceptual complexity of operations involving autonomous systems has not allowed the rapid advances in autonomous system capabilities to be considered in a timely manner [5].

This position paper reviews recent advances from industry, academia and national defense to provide a summary of the latest techniques and approaches that may compliment and bolster the existing NATO CD&E process. The paper continues to propose a possible update to the NATO CD&E toolset to enable the efficient consideration of autonomous systems in future operations.

Section 2 of this paper summarizes the key components of the NATO CD&E process. Section 3 discusses the key concepts and challenges presented by autonomous systems in the underwater domain. A review of the latest developments in autonomous system design and test are presented in Sect. 4. Section 5 provides conclusions and links between the review findings and the CD&E process steps.

2 The NATO CD&E Process

NATO defines its existing CD&E process as a technology agnostic approach that allows the proposal and test of a range of potential future concepts of operation in all military fields [6].

Key to the concept of CD&E is the use of a spiraling approach, with iterating and separate concept development and experimentation stages. The iterative, spiraling approach is managed in a series of increasing capability maturity levels (CMLs). An overview of the approach is provided in Fig. 1.

The CML groupings represent stages of maturity from the low maturity CML 1, where up to five novel concepts of operation are selected to the more mature concepts discussed at CML 6 where the selected and refined operational scenario is demonstrated and validated implementation requirements are obtained.

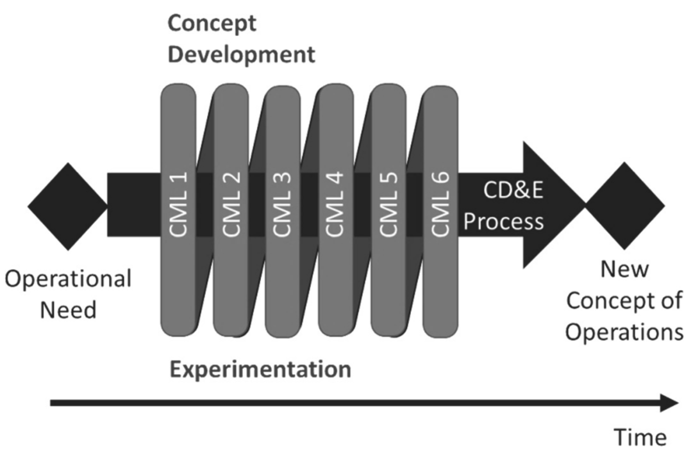

Fig. 1. Summary of NATO's CD&E process.

While the CD&E process does not mandate the use of any specific tools or techniques for the concept development stage, techniques commonly used by NATO are summarized in Table 1 [7].

Table 1. Summary of existing and commonly used NATO concept development techniques

Existing concept development techniques	
Category	Description
Analysis	The definition of the problem statement by a team of end-users and experts in the problem field The activity delivers clear problem statement
Brainstorming	Whiteboard based group sessions in which new knowledge and ideas are generated, discussed and linked. This methodology may apply approaches from the NATO Alternatives Analysis (AltA) handbook [8] The activity delivers a list of potential solutions
Evaluation	Workshops where the brainstorming output is translated to the concept and evaluated by a team of experts in the problem field The activity delivers a shortlist of potential solutions and rationale for their selection

Further, the NATO tools commonly used in the experimentation phase of the NATO CD&E process are summarized in Table 2 [7].

Table 2. Summary of existing and commonly used NATO experimentation techniques

Existing experimentation techniques	
Category	Description
Table-top gaming	Paper or computer supported games in which the users play a central role The activity delivers further detail about the operation of a process or system in a number of scenarios
Experimentation with virtual/constructive simulation	Computer-based experiments where the focus lies on investigating the detailed behaviour of a modelled system The activity delivers test evidence that indicates the performance of a system in a range of environments or scenarios
Live simulation	Experiments with a real system in the field, using real software and hardware, including operators, in a suitable live test environment The activity delivers test evidence from a representative system in a representative environment

The CD&E approach has been used to excellent effect within NATO to reduce both costs and timescales while increase the quality and end customer value of the final solutions [9]. Specific issues that prevent the timely adoption of autonomous systems are encountered when CD&E techniques are used to discuss and analyze the additional technical and behavioral complexity of autonomous systems [10–12]. Further description of the challenges presented by autonomous systems and the specific challenges in underwater operation are discussed in the following section of this paper.

3 Autonomy in the Underwater Domain

The term 'autonomous system' generally refers to a system that is required to operate with some degree of human independence [13]. Differing from 'unmanned systems', the more complex autonomous systems are able to interact and respond to their environment without the involvement of a human in the loop. This leads to a new level of technical challenge, where not only the systems function but also its behavior needs to be understood and proven [14]. Challenges specific to the underwater domain stem from both the lack of in-mission communications, requiring the operation of the system for long periods without human control and intervention, and the difficulty in sensing and understanding the complex underwater environment.

The key performance parameters for autonomous systems include elements such as how precisely it can observe the environment through its sensors, how effectively it can combine sensor data sources and whether the resulting behavior is the most appropriate given the environment and the required mission objectives.

The evaluation of these key performance parameters must be made from the viewpoint of two key CD&E process stakeholders; operators and commanders [15].

For **commanders**, a key challenge presented by the use of autonomy is in understanding and developing trust in the mission specific concept of operations and the associated trade-offs across a wide range of conflicting parameters of interest. At the level of the commander, concepts are currently both difficult to articulate during concept development and difficult to test, analyze and present following the experimentation stage of the process.

Operators must be able to understand and operate the human-machine interfaces needed to conduct in their mission. Again, the behavior of the system must be understandable and predicable to the operator to build trust and confidence in the system.

3.1 Barriers to the Inclusion of Autonomous Systems in the CD&E Process

The complexity of autonomous systems, combined with the lack of in-mission human supervision to detect and act upon unexpected failures has led to a lack of trust by both commander and operator stakeholder groups and is limiting the adoption of autonomous underwater systems in operations [16].

Further, the additional complexity of autonomous systems requires efficiency improvements in the experimentation and test steps of the CD&E process. The need for further efficiency is driven by the non-deterministic nature of autonomous systems. The number of required test cases, the amount of data generated, and the complexity of the analysis process provide additional barriers to the adoption of autonomous systems in the CD&E process [17, 18].

A range of research activities are currently underway to communicate the potential of complex concepts and technologies to a range of stakeholders. A review of emerging approaches aimed at communicating capabilities, building trust and efficiently testing autonomous systems is presented in the following section of this paper.

4 Techniques for Assessing Autonomous System Capabilities

This section of the paper reviews recent advances from academia, industry and the NATO nations aimed at assessing autonomous system capabilities that may be used as best practice case studies for future NATO CD&E techniques.

4.1 Understanding the States and Interfaces of Autonomous Systems

Due to the non-deterministic nature of autonomous systems and the large number of potential operational unknowns the systems may encounter, guidance is required in the CD&E process to effectively test the system in a representative range of environments. The lack of guidance in this area leads to both complex test phases that does not result in a clear comprehension of the tests coverage in relation to problem space. Solutions to reducing the time and complexity of the test stages for autonomous systems have been pioneered by a range of national programs [19]. One example which will be used to demonstrate work typical to this area is the work carried out by the US Army Robotic Intelligence Evaluation Program [20]. Their work has investigated the use of a design of experiments (DoE) based methodology to both comprehensively test the intelligence of autonomous systems while limiting the number of required test cases.

The first step of the approach uses ‘parameter effect propagation’ to limit the number of tests that are required by running only those tests that provide a unique situation to the autonomous system.

Parameter effect propagation is the process of recognizing each of the individual sets of parameter values (i.e., all the possible scenarios) and estimating the effect on the autonomous system. Central to the identification of parameter values is the recognition that the autonomous underwater vehicle (AUV) decision system can only be affected by its sensor inputs and that the senor inputs are often only limited in their scope.

This initial stage of the analysis results in the population of a critical test matrix where the function of the system (e.g. communication packet loss, battery level, sonar received signal strength) are identified. A pictorial summary of the methodology is shown in Fig. 2.

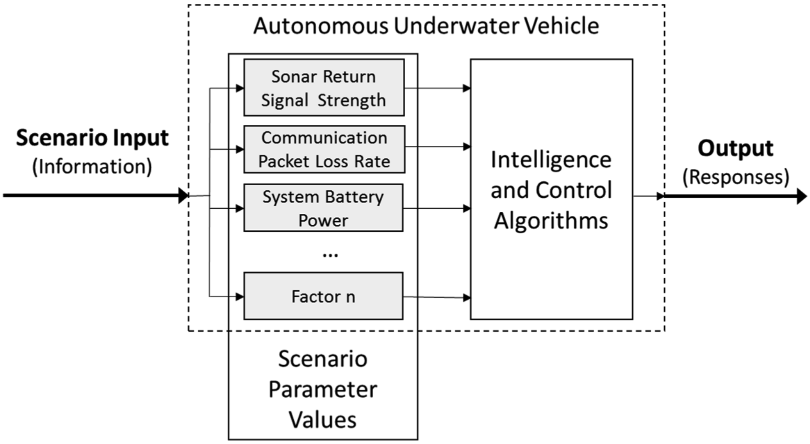

Fig. 2. Factors affected by environmental parameters

Following the identification of the AUV inputs factors, an assessment of the impact of the scenarios on that factor should be made for each scenario.

Before testers can determine how the parameters of a scenario will uniquely affect an AUV, they need to establish the parameters. Using the list of customer defined, testers can produce a list of scenario parameters for each. The list of scenario parameters is determined by examining the variables in a scenario's mission along with all the ways the environment can interact with the system. This approach limits the number of test cases as the environment can only interact with the system in a few ways. Like humans, AUVs can only base decisions on what they detect. The published method may been applied to aid the understanding of AUVs. An example critical test matrix for an example parameter, sonar reflected signal strength, has been populated and is shown in Fig. 3.

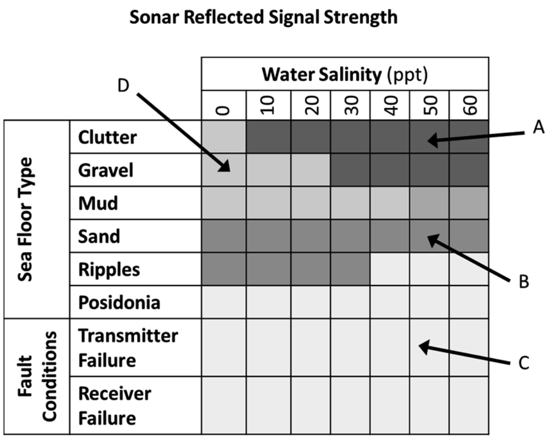

Fig. 3. Critical test matrix test zones

The different shaded regions identified in Fig. 3 each represent a different value for the reflected sonar signal strength and provide a number of unique testing regions. To the system being tested, each similarly shaded case would be identical because the system is unable to realize a difference between them. If testers were to run every case, there would be 56 tests just for environmental effects on reflected sonar strength alone. Instead, choosing one case from each of the unique testing regions leaves testers with four (Labelled from 'A' to 'D') tests and with a significant portion of the information they would have had running all 56 tests. With more complex sensors, the resulting test matrix might be many orders of magnitude less than the original full factorial test design. In addition, each resulting test should propose a unique problem to the AUV resulting in the greatest probability of inducing, identifying and attributing emergent and possibly unwanted behavior. An AUV with a small number of interacting subsystems will typically end up with a limited test set for each scenario, which can be addressed with standard CD&E test techniques.

Further, error and fault conditions may also be considered at this stage. As an example, a high frequency sonar may return much less information when surveying from a posidonia covered sea floor. This scenario is likely to return the same information as broken sonar scenarios, and the AUV will treat them as the same. This is identified as an extension to existing matrix test zones in Fig. 3.

For many of the identified parameter, such as communication packet loss rate, the test plan must specify which values to use. If the parameter is continuous or there is a multitude of discreet values, a traditional approach is for testers to choose values at the 95 ($\approx 2\sigma$) and 99 ($\approx 3\sigma$) percent extremes on the probability distribution function of a parameter. However, with a system containing so many parameters, testing only the extremes would leave possible many common parameter interactions untested. A more comprehensive approach is to select parameter values along a probability distribution function at a given percentage step size. This allows the tester to not only select the extremes but also test more of the most common values. The percentage step size is determined by running a sensitivity analysis on the AUV factors by this particular parameter.

This sensitivity analysis forms the next step of the DoE process. Effects of subsystem factors can be predicted through analytical calculations and then verified through field testing or through empirical experimentation. Testers should not be concerned with the multitude of parameter value combinations, but rather concerned with the sets of factor effects derived from the parameter permutations. A key concept in a number of autonomous system test approaches [17, 20] is that running 100 different tests that induce the same factor effect set in the system will not tell the tester nearly as much as running 100 different tests where each of which induces a unique factor effect set.

A test team should step through each parameter value for each scenario and calculate a corresponding factor effect set. It is important to take the set of parameter values that define a scenario together so that the factor effect set can accurately represent interaction between multiple parameters. For example, range and sea floor type can each affect sonar sensors, but the combination of different values of these two parameters can produce drastic effects on a system. Further, if the sea floor is out of range, it makes no difference to the sonar sensor on an AUV what the sea floor covering is because it may not be able to "see" at all. If all the effects of all the values of one parameter dominate all the effects of all the values of another parameter, there is no reason to have the second parameter and it can be removed from the test scenario.

4.2 Analyzing and Understanding the Experimentation Phase

Once the DoE test environments have been identified and tests conducted, the next stage of the CD&E process mandates the analysis of the results. While the DoE approach limits the number of tests that need to be performed, it is likely that there are still too many test results to review in full with the system commanders. Further work is required to identify approaches to effectively communicate the findings of the experimentation phase, allowing the progression to the next CML level.

A current area of work that may allow this capability is in the development of 3D visualization environments that allow system behavior to be demonstrated to end users [21]. Simulation capabilities have been provided by several available tools [22, 23] for autonomous ground vehicles. These tools typically show, on one screen, the motion and actions of the autonomous system along with a user interface that allows the audience to 'play' with the environment around the autonomous system. An example of this approach can be seen applied to autonomous ground vehicles in Fig. 4.

The addition of an input, also shown in Fig. 4, allows the users trust to be reported. Areas where the user does not trust the system, for example if the user notices that the autonomous system is moving toward home with an explosive still loaded, a flag can be raised and the behavior investigated further in the next CML concept development stage of the process.

Fig. 4. A virtualized environment to demonstrate autonomous system behaviour

Without the users' inputs from the 3D visualization, this potentially dangerous behavior would be much more difficult to detect.

Further to watching and witnessing the behavior of the system, work has also been carried out that allows the system commander and operators to interact in the creation of scenarios. Their involvement in creating the possible scenarios is vital in the concept development phase of the CD&E process. Using Event Sequence Charts [21], Inputs and scenario development can be managed even in complex scenarios.

An example of an Event Sequence Chart for an underwater autonomous system mine detection system can be seen in Fig. 5.

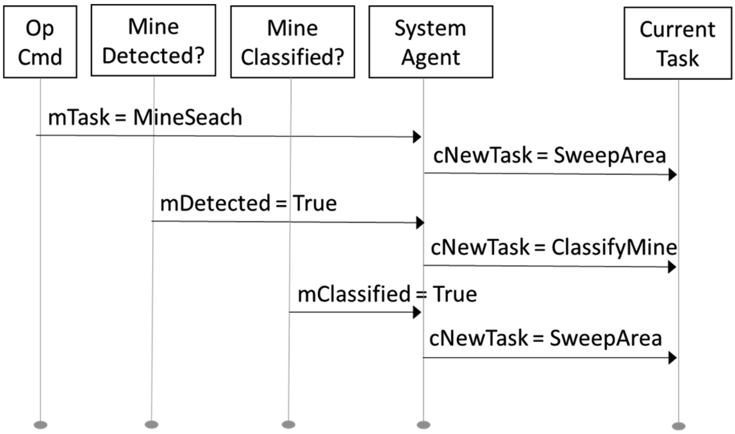

Fig. 5. Event sequence chard for underwater mine identification and classification

Based on UML sequence diagrams, these charts allow the development of sequenced events by the user that can form the basis of either the concept development of experimentation phased of the CD&E process.

4.3 Aiding Communication and Understanding in Concept Development

Following the efficient integration of autonomous system testing into the experimentation analysis of the results by all relevant stakeholders the complexity of autonomous system behavior may benefit from additional tools to clarify new concepts and ways of working. Pioneered by the education sector, a large body of work has been carried out, into allowing the sharing of ideas and concept development aided by AR [24] and VR [25].

Utilizing and building on the models and approaches already discussed in this paper, this approach allows the articulation of complex ideas in an initiative manner to a range of stakeholders. Further, commonly used systems also allow distributed collaborations of specialized personal, encouraging the involvement of the most appropriate personnel efficiently within the CD&E process and driving further improvements to the CD&E process.

5 Conclusions

Recent developments in autonomous system technology have provided an opportunity for new concepts of operation to be developed in a range of NATOs undersea activities. NATOs existing CD&E provides an excellent and adaptable framework for allowing the efficient development integration of emerging technologies. Despite this, the additional complexity of autonomous systems, along with the removal of in-mission humans in the loop, presents a series of challenges that limit the rate of adoption of this improved technology in the existing CD&E framework.

This paper has reviewed the creation of a range of processes and techniques from industry, academia and NATO nations to identify those that may enable autonomous systems to integrate better into the NATO CD&E process.

To enable the experimentation phase, methods for simplifying and clarifying the relevant test cases while increasing the rigor and robustness of testing by identifying common system boundaries has been presented. Building up this with a DoE approach, supported by clear reporting metric matrices, allows both the efficient and fast testing of the system and, more importantly, the results of the testing to be understood in terms of their coverage of the problem space.

Once tested, methods for demonstrating the capabilities and behaviors of the system have been described. These approaches aim to build trust with both system operators and commanders by allowing them to intuitively witness and interact with the system. Areas of the scenario where trust is lost are recorded, highlighted and used to update subsequent CD&E CML stages.

Further, advances spearheaded by the education sector can be used to enhance and streamline the concept development stages. The VR and AR in concept development allow the articulation of complex autonomous system behaviors and can be used to allow subject matter experts to interact with system users as required.

The application of these approaches can be used to enhance the tools available in the existing CD&E process. A summary of the possible improvements identified by this paper and their links to the existing CD&E process is presented Tables 3 and 4.

Table 3. Enhanced approaches and toolsets to aid the CD&E process

Concept development techniques		
Category	Existing techniques	Potential additional techniques
Analysis	The definition of the problem statement by a team of end-users and experts in the problem field. The activity delivers clear problem statement	The generation of a virtualised environment that shows and describes the current operational experiment. The virtualised environment may be combined with VR or AR as required to allow all stakeholders to develop a detailed understanding of the key issues and challenges to be solved

(*continued*)

Table 3. (*continued*)

Concept development techniques		
Category	Existing techniques	Potential additional techniques
Brainstorming	Whiteboard based group sessions in which new knowledge and ideas are generated, discussed and linked. This methodology may apply approaches from the NATO Alternatives Analysis (AltA) handbook [8]	Distributed group sessions with VR and AR tools that involve both subject matter experts and system operators and commanders to collaboratively develop potential solutions
Evaluation	Workshops where the brainstorming output is translated to the concept and evaluated by a team of experts in the problem field. The activity delivers a shortlist of potential solutions and rationale for their selection	Interactive use of a virtualised 3D environment to develop an improved understanding of autonomous system behaviours. This has recently been demonstrated in the development of new anti-submarine warfare concepts that involve multiple autonomous systems in the maritime domain at CMRE

Table 4. Enhanced approaches and toolsets to aid the CD&E process

Experimentation techniques		
Category	Existing techniques	Potential additional techniques
Table-top gaming	Paper or computer supported games in which the users play a central role. The activity delivers further detail about the operation of a process or system in a number of scenarios	Interactive use of virtualised models with end user inputs to highlight areas that require further development to provide a suitable solution
Experimentation with virtual/constructive simulation	Computer-based experiments where the focus lies on investigating the detailed behaviour of a modelled system. The activity delivers test evidence that indicates the performance of a system in a range of environments or scenarios	Targeted testing based on DoE principles to highlight unique areas of operations. The outputs of the experimentation should result in the population of a metric matrix to identify the strengths and weaknesses of the system as well as an idea of the problem space covered
Live simulation	Experiments with a real system in the field, using real software and hardware, including operators, in a suitable live test environment. The activity delivers test evidence from a representative system in a representative environment	Live events carried out when identified by the critical test matrix, maximising the value and efficiency each test

References

1. Dyndal, G.L., Berntsen, T.A., Redse-Johansen, S.: Autonomous military drones: no longer science fiction. NATO Review Magazine, Oslo, Norway (2017)
2. Williams, A.P., Scharre, P.D.: Autonomous Systems - Issues for Defence Policy Makers. NATO Headquarters SACT, Norfolk (2015)
3. Yuh, J.: Design and control of autonomous underwater robots: a survey. Auton. Robot. **8**(1), 7–24 (2000)
4. de Nijs, H.: Concept development and experimentation policy and process. HQ SACT, Norfolk, USA (2010)
5. Boulanin, V., Verbruggen, M.: Mapping the Development of Autonomy in Weapon Systems. Stockholm International Peace Research Institute, Stockholm (2017)
6. NATO: NATO concept development and experimentation (CD&E) process. North Atlantic Military Committee (2010)
7. van der Wiel, W., et. al.: Concept maturity levels bringing structure to the CD&E process. In: Interservice/Industry Training, Simulation and Education Conference, Orlando, USA (2010)
8. NATO: The NATO Alternative Analysis Handbook, 2nd edn. NATO, Brussels (2017)
9. Software V&V Working Group: IEEE STD 1012-2012 - IEEE Standard for System and Software Verification and Validation. IEEE (2012)
10. Pecheur, C.: Verification and Validation of Autonomy Software at NASA. NASA Ames Research Center, USA (2000)
11. Schumann, J., Visser, W.: Autonomy software: V&V challenges and charecteristics. In: IEEE Aerospace Conference, Big Sky, USA (2006)
12. Hodicky, J., Prochazka, D.: Challenges in the implementation of autonomous systems into the battlefield. In: 6th International Conference on Military Technologies, Brno, CZE (2017)
13. Callow, G., Watson, G., Kalawsky, R.: System modelling for run-time verification and validation of autonomous systems. In: Conference in Systems of Systems Engineering, Loughborough, UK (2010)
14. Defence Science Board: The Role of Autonomy in DoD Systems. Office of the Under Secretary of Defence for Acquisition, Technology and Logistics, Washington, USA (2012)
15. Tremori, A., et al.: A verification, validation and accreditation process for autonomous interoperable systems. In: Mazal, J. (ed.) MESAS 2017. LNCS, vol. 10756, pp. 314–323. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-76072-8_22
16. Palmer, G., Selwyn, A., Zwillinger, D.: The "Trust V": building and measuring trust in autonomous systems. In: Mittu, R., Sofge, D., Wagner, A., Lawless, W.F. (eds.) Robust Intelligence and Trust in Autonomous Systems, pp. 55–77. Springer, Boston, MA (2016). https://doi.org/10.1007/978-1-4899-7668-0_4
17. Helle, P., Schamai, W., Strobel, C.: Testing of autonomous systems - challenges and current state-of-the-art. In: INCOSE International Symposium, Edinburg, UK (2016)
18. Hodicky, J.: Autonomous systems operationalization gaps overcome by modelling and simulation. In: Hodicky, J. (ed.) MESAS 2016. LNCS, vol. 9991, pp. 40–47. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-47605-6_4
19. Thompson, M.: Testing the intelligence of unmanned autonomous systems. Int. Test Eval. Assoc. **29**, 380–387 (2008)
20. Ahner, D.K., Parson, C.R.: Workshop report: test and evaluation of autonomous systems. USA Department of Defense, Washington D.C., USA (2016)

21. Heitmeyer, C.K., Leonard, E.I.: Obtaining trust on autonomous systems: tool for formal model synthesis and validation. In: Workshop on Formal Methods in Software Engineering, Florence, Italy (2015)
22. Heitmeyer, C.L., Archer, M., Bharadwaj, R., Jeffords, R.D.: Tools for constructing requirements specifications: the SCR toolset at the age of 10. Comput. Syst. Sci. Eng. **20**(1), 19–35 (2005)
23. Knexus: eBotworks. http://www.knexusresearch.com
24. Phon, D.N.E., Ali, M.B., Halim, N.D.A.: Collaborative augmented reality in education: a review. In: International Conference in Teaching and Learning in Computing and Engineering, Kuching, Malaysia (2014)
25. Carruth, D.W.: Virtual reality for education and workforce training. In: International Conference on Emerging eLearning Technologies and Applications, Stary Smokovec, Slovakia (2017)