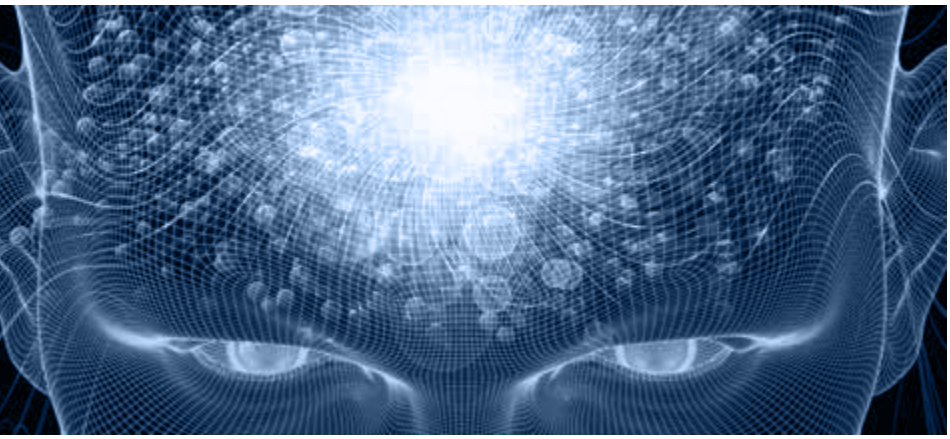
The Automation Conundrum

Do you have something to say about this article?
Visit the *Journal* pages of the ISACA® website *(www.isaca.org/journal)*, find the article and click on the Comments link to share your thoughts.

日本語版も入手可能
www.isaca.org/currentissue

"The computer is a moron. And the stupider the tool, the brighter the master must be," claimed Peter Drucker, in an often-quoted 1967 article.[1] Although this statement lends itself to a bit of hyperbole, the argument was clear and perhaps relevant at the time, because computers were only replacing clerical chores.

Fifty years later, artificial intelligence (AI) systems, riding on the exponential increases in computing power and the availability of big data, are outperforming humans in numerous domains. These intelligent systems continue to penetrate every industry sector and are delivering enormous benefits in the form of new business opportunities, deeper customer insights, improved efficiency, enhanced agility and so forth.



Phillimon Zongo
Phil Zongo is a cybersecurity consultant based in Sydney, Australia. He has more than 10 years of technology risk consulting and governance experience working with leading management consulting firms and large financial institutions. He has practical experience advising senior business and technology stakeholders on how to manage critical risk in complex technology transformation programs. He also authored "Managing Cloud Risk Top Considerations for Business Leaders," published in volume 4, 2016, of the *ISACA® Journal*.

US-based Memorial Sloan Kettering Cancer Center is using IBM Watson to compare patient medical information against a vast array of treatment guidelines, published research, journal articles, physicians' notes and other insights to provide individualized, confidence-scored recommendations to physicians.[2] In Canada, Bank of Montreal deployed robo-advisors to provide automated, algorithm-based portfolio management advice to its customers.[3] Massachusetts Institute of Technology (MIT) (USA) developed an AI system that can detect 85 percent of cyberattacks by reviewing data from more than 3.6 billion lines of log files each day and informing about anything suspicious.[4]

Adoption of AI systems is expected to accelerate over the next few years. A December 2015 report by Bank of America Merrill Lynch Research predicted that the robotics and AI solutions market will grow to US $153 billion by 2020—comprising US $83 billion for robotics and US $70 billion for AI-based analytics. The same report estimates that this exponential growth can boost productivity by up to 30 percent and cut manufacturing labor costs by 18 to 33 percent.[5]

While some organizations are still experimenting with AI using insignificant business tasks, others are taking ambitious strides by delegating mission-critical roles to AI algorithms. One such example is Deep Knowledge Ventures, a Hong Kong–based venture capital firm, which, in May 2014, took a leap of faith and appointed an AI algorithm to its board of directors.[6] The algorithm, named Vital, automates due diligence by scanning financing, clinical trials, intellectual property and previous funding rounds of prospective enterprises then votes on whether to invest in the enterprise or not; a role with significant responsibility and consequence.

The proliferation of AI raises intriguing opportunities; however, associated risk exists—and should it prevail, its impacts can result in significant consequences. A number of strategic concerns have been documented regarding the rise of AI; however, this article highlights three crucial risk

concerns that leaders face when adopting AI within their businesses and provides practical insights to minimize business exposure while maximizing AI potential. These risk concerns are:

- Critical business decisions based on flawed or misused AI algorithms
- Cultural resistance from employees whose roles are vulnerable to automation
- Expanded cyberthreat surfaces as AI systems replace more vital business functions

Flawed or Misused AI Algorithms

A well-designed AI system can significantly improve productivity and quality, but when deployed without due care, the financial and reputational impacts can be of epic magnitude. In banking and finance, flawed algorithms may encourage excessive risk taking and drive an organization toward bankruptcy. In the health care sector, flawed algorithms may prescribe the wrong medications, leading to adverse medical reactions for patients. In the legal sector, flawed algorithms may provide incorrect legal advice, resulting in severe regulatory penalties. In 2012, Knight Capital Group, a US-based market-making firm, provided an unsettling insight into the likely impacts of such risk when it lost more than US $440 million in just 30 minutes as a result of an untested change to its high-frequency trading algorithms. Dubbed “the mother of all software glitches,” the incident cost the firm four times its 2011 net income.[7]

In contrast to traditional rule-based systems where errors can be rolled back with minimum business impact, minor errors in critical AI algorithms can result in severe consequences. Further complicating this risk is the probability that AI systems can behave unpredictably when interacting with humans or the external environment. As intelligent systems increasingly take on vital business roles, the risk that crucial business decisions might be based on flawed algorithms invariably rises. Therefore, the need for the AI system concepts to match those of its human designers increases as the AI system becomes more powerful and autonomous.[8]

The three key critical steps that can help businesses to maximize AI value while managing risk are:

- Align AI adoption with business strategy and risk appetite
- Experiment with low-risk functions
- Test rigorously

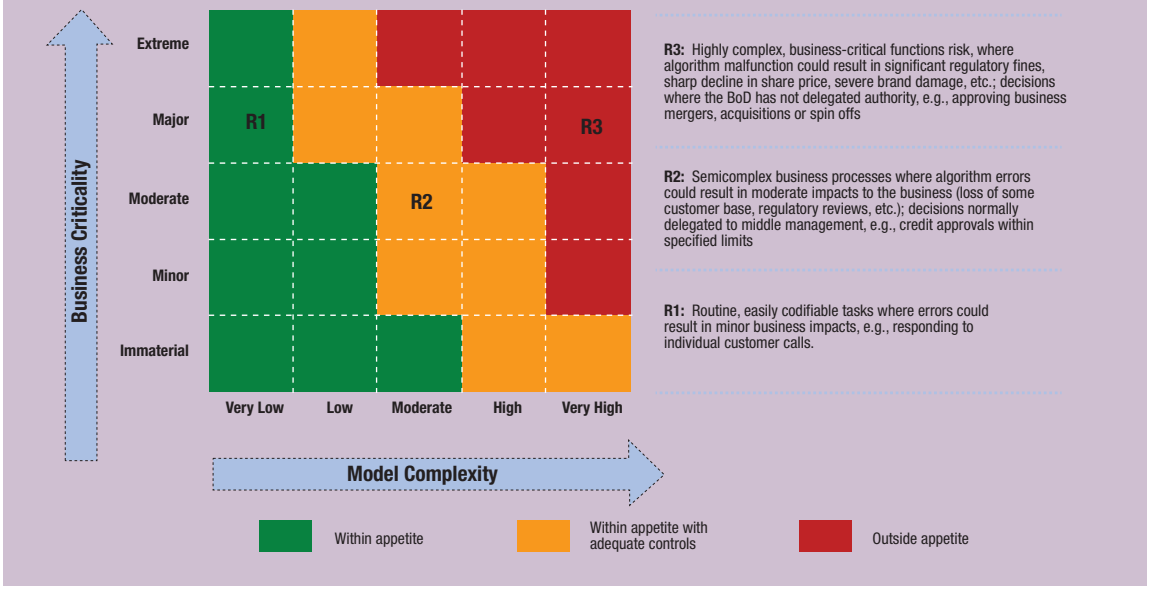
Align AI Adoption With Business Strategy and Risk Appetite

Business leaders should be mindful of key risk that is inherent in AI adoption, conduct appropriate oversight, and develop principles that articulate the business roles that can be partially or fully automated. Equally important, the board should approve the automation of high-risk business functions, ensuring that the business is not exposed to risk beyond its capacity or risk that does not contribute to the business strategy.

“In contrast to traditional rule-based systems where errors can be rolled back with minimum business impact, minor errors in critical AI algorithms can result in severe consequences.”

A simple way to conduct this assessment is illustrated in **figure 1**, which models risk exposure along two factors: criticality of the business function being automated and complexity of the associated model. In the example in **figure 1**, a financial institution may decide to automate some call center functions (R1) and avoid automation of business acquisition or spin-off approvals (R4), based on different risk exposures. Routine or clerical business roles are naturally easier to automate and pose less business risk compared to complex functions such as those requiring intellectual reasoning, creativity, interpersonal skills or emotional intelligence.

Figure 1—Risk Assessment for Business Process Automation



Source: P. Zongo. Reprinted with permission.

A clear understanding of regulations that govern specific business functions is also vital because full automation of some business functions might be prohibited in certain jurisdictions. For example, in April 2016, the Massachusetts (US) Securities Division published a policy statement in which the division questioned the ability of robo-advisors to act as state-registered investment advisers. The securities regulator stated, "It is the position of the Division that fully automated robo-advisers, as currently structured, may be inherently unable to carry out the fiduciary obligations of a state-registered investment adviser."[9] The division's argument was that a fully automated robo-adviser may not act in the best interest of its client, does not conduct sufficient due diligence, provides advice that is minimally personalized and may fail to meet the high standard of care.[10] This policy position underscores the importance of carefully considering the legal implications that are associated with automating a business function, including anticipated reforms, before committing any project capital.

An effective risk assessment requires business leaders to answer the following crucial questions:

- How can intelligent systems advance the enterprise business strategy and what does success look like?
- What are the plausible financial, reputational or regulatory risk if the AI system malfunctions, and does the business have enough capacity to absorb associated impacts if the risk materializes?
- What are competitors doing in this space, and how far have they advanced in pursuit of these goals?
- Is the business willing to take a leadership role or wait until the benefits of AI are fully proven?
- Does the organization have demonstrable expertise in managing the risk? If this is being outsourced, has the identified vendor successfully delivered AI transformation programs of similar or larger scale?

Although AI adoption introduces significant challenges, it can also be a catalyst for risk reduction. The first industrial robot, Unimate, created in 1961 by American inventor George Devol, was designed for that purpose. The 4,000-pound robotic arm transported die castings from an assembly line and welded these parts onto automobile bodies. This was a high-risk task for workers who could be poisoned by exhaust gas or lose a limb if they were not vigilant.[11] A similar, but more current, example is the IBM Watson system, which is being used by companies operating in heavily regulated industries to keep up with ever-changing legislation and compliance standards.[12]

Experiment With Low-risk Functions

Delegating a crucial task before attaining a solid theoretical understanding of the associated outcomes has high risk.[13] Therefore, organizations should experiment, learn and adapt using low-risk, low-cost and easily codifiable tasks. After the underlying assumptions are validated, competences are proven and major uncertainties are resolved, organizations can gradually automate more complicated functions.

Test Rigorously

Due to their high degree of uncertainty, intelligent systems require more extensive testing than traditional applications. When constructing intelligent systems that learn and interact with all complexities of reality, it is not sufficient to verify that the algorithm behaves well in test settings. Additional work is necessary to verify that the system will continue working as intended in live environments.[14] This testing should be performed by employees with appropriate qualifications and motivations. Likewise, detailed testing should be performed after the AI system has been modified, or after it has acquired new intelligence, and the conditions under which these tests are conducted should reflect a real-life environment.

Cultural Resistance

Any significant transformation program can be deeply unsettling for employees. AI programs amplify this risk, because employees whose jobs are vulnerable to automation—especially those performing less-skilled and repetitive tasks—may be worried about the fate of their jobs. Consequently, these employees may dig in to protect their turf and actively resist change, derailing AI program success. Revolts against innovation are not new. One of the most famous examples is the Luddite movement of the early 19th century, during which a group of English textile artisans protested the automation of textile production by seeking to destroy some of the machines.[15] Furthermore, lack of clear and consistent communication from leaders leaves employees open to confusion and distrust of important AI transformation programs.

> **“To successfully lead an AI transformation, business leaders must create an environment of trust and ensure high levels of employee engagement, buy-in and support.”**

A 2011 report emphasized that the “reshaping of employee attitudes and behaviours is just as critical to the success of a transformation as the implementation of process changes.”[16] To successfully lead an AI transformation, business leaders must create an environment of trust and ensure high levels of employee engagement, buy-in and support. To do this, business leaders should:

- Communicate a compelling change story that motivates employees and promotes a shared automation vision for the future
- Identify segments susceptible to automation; assess impact on employees and identify alternative job opportunities

- Establish a dedicated change management team consisting of senior business leaders, human resources, and change professionals to communicate the transformation agenda, anticipate challenges, and minimize attrition rates. Change management communications should also be targeted and allow for employee feedback.
- Identify opportunities for employees to work alongside AI systems and formulate strategies to maximize those synergies. Knowledge jobs generally consist of a range of tasks, so automating one activity may not make an entire position unnecessary.[17] For example, algorithms can perform routine tasks, freeing time for humans to manage customer relationships or derive deeper business insights. Also, highly regulated tasks might not be completely replaced by machines.
- Engage legal teams for due diligence to understand applicable job protection laws and appropriate responses if the program intends to completely automate some jobs
- Establish incentives to promote behavioral changes and keep people engaged

> **"The ability of AI systems to fully transform business hinges on the effectiveness of their security and privacy controls."**

Businesses will continue to automate tasks that were performed by humans to drive down costs, improve efficiency and reduce operational errors. Given the disturbing impact that automation can have on an organization's most valuable assets—its employees—it is essential for business leaders to anticipate potential risk early to minimize possible negative impacts. Employees also have a part to play: up-skilling themselves to remain relevant in the face of disruptive innovation. Researchers have predicted, "As technology races ahead, low-skill workers will reallocate to tasks that are non-susceptible to computerisation—i.e., tasks requiring creative and social intelligence. For workers to win the race, however, they will have to acquire creative and social skills."[18]

Expanded Cyberattack Surface

The ability of AI systems to fully transform business hinges on the effectiveness of their security and privacy controls. Failure to provide these assurances can inhibit their acceptance. The Bank of America Merrill Lynch Research report states that cyber security and privacy concerns, and other critical factors such as regulation, insurance and cost, remain primary hurdles to self-driving-car adoption. The report cites that 54 percent of buyers fear that connected cars will be hackable, and 30 percent do not want to use a connected car because of privacy concerns.[19] In 2015, a group of Virginia (USA)-based researchers successfully hacked into a driverless car system and took control of a vehicle, highlighting the significant threat posed by unsecured AI systems.

Cyberrisk continues to increase in frequency and business impact, and has gained significant attention from boards of directors, regulators and policy makers. Public and private-sector enterprises are already struggling to keep up with relentless, sophisticated and well-resourced cybercriminals. AI further complicates this struggle with the issues that are described in the following sections.

Vulnerabilities

To date, no industry standards exist to guide the secure development and maintenance of AI systems. Further exacerbating this lack of standards is the fact that start-up firms still dominate the AI market. A recent MIT report revealed that, other than a few large players such as IBM and Palantir Technologies, AI remains a market of 2,600 start-ups. The majority of these start-ups are primarily focused on rapid time to market, product functionality and high return on investments. Embedding cyberresilience into their products is not a priority.

Inadvertently, vendors ship solutions with basic security controls and easily exploitable vulnerabilities such as default passwords or weak authentication techniques. These weaknesses not only provide easy targets for cybercriminals to exploit, but also potentially refute layers of existing network security controls. The Verizon *2016 Data Breach Investigations Report* highlighted that 63 percent of confirmed breaches involved weak, default or stolen passwords.[20]

The self-learning capabilities of AI systems also present unique challenges. Cybercriminals might successfully predict the data that are used to train an algorithm and deliberately manipulate its behavior, contrary to its design objectives. The results of a recent Microsoft live experiment with an AI chat-bot, named Tay, offers a cautionary tale about the dangers of exposing vulnerable AI systems to the Internet. In March 2016, Microsoft admitted that it had made a critical oversight when a coordinated attack exploited vulnerability within its experimental AI algorithm. Tay was designed to mimic a teenage girl, interact with people on social media and learn from them. Unfortunately, Microsoft's oversight left Tay open to a specific vulnerability that was exposed by the attack and resulted in Tay sending wildly inappropriate, offensive and hurtful tweets and images, including racial slurs misrepresentative of Microsoft's values and Tay's design.

A Zero-sum Game

Intelligent systems are already playing a crucial role in combating cybercrime, for example, through automated fraud detection and spam detection. However, this role may prove to be a zero-sum game, because the same technology can be used to perpetrate highly sophisticated and evasive cyberattacks against critical systems. This sentiment was echoed by more than 75 percent of respondents who were polled in a 2014 survey that was jointly conducted by McKinsey and the World Economic Forum (WEF), including chief information officers (CIOs), chief risk officers (CROs), chief technology officers (CTOs), regulators and business unit executives, who conceded that the sophistication or pace of cyberattacks would grow faster than their own defensive capabilities.[21]

Therefore, an important question is: Will these malefactors continue to outsmart security vendors and develop superior and elusive AI programs that will unleash advanced persistent threats against critical systems, manipulate stock markets, perpetrate high-value fraud and consistently steal intellectual property, and, in doing so, destroy associated forensic evidence?

If current cybercrime trends continue unabated, residual business cyberrisk exposure may continue to rise.

Building Cyberresilient Intelligent Systems

To support business innovation and maximize its value, comprehensive cyberresilience for intelligent systems is vital. Unified efforts by policy makers, business leaders, regulators and vendors are a prerequisite for long-term success. However, before these concerted standards come to realization, business leaders should:

- Use existing, industry-accepted industry standards where possible. Although these are not specifically designed for intelligent systems, they can help businesses to identify common security risk and establish a solid baseline for securing new technologies. Notable frameworks include:
  - **Open Web Application Security Project (OWASP) Top 10**[22]—A list of the 10 most current critical web application security flaws, along with recommendations to ensure that web applications are secured by design.
  - **US National Institute of Standards and Technology (NIST) Cyber Security Framework**[23]—Consists of standards, guidelines and practices to promote the protection of critical cyberinfrastructure.
  - ***COBIT® 5 for Information Security***[24]—Provides detailed and practical guidelines for security professionals to manage and govern important information security, and make more informed decisions while maintaining awareness about emerging technologies and the accompanying threats.

- Engage experienced security consultants to review critical controls for AI products (including detailed penetration testing) and fix any exploitable security vulnerabilities before going live
- Conduct due diligence to determine vendor security capabilities, product security road map and frequency of security updates—with a long-term commitment to product security being a critical success factor

> **"In today's dynamic business environment, organizations need to experiment with new digital capabilities and accept risk in pursuit of new product offerings and to remain relevant to their customers."**

- Deploy robust encryption to protect sessions between AI systems and critical records from compromise (commonly referred as man-in-the-middle attacks)
- Grant minimum system privileges and deploy strong controls to protect service accounts that are used by AI systems to execute critical tasks from abuse—especially those with administrator—equivalent privileges
- Adopt defense in depth to ensure that a failure in one control layer will not result in a system breach

Conclusion

Looking ahead, numerous challenges remain for the full adoption of intelligent systems, like any emerging technology. These challenges may pale in comparison to the consequences of missing opportunities presented by AI.

In today's dynamic business environment, organizations need to experiment with new digital capabilities and accept risk in pursuit of new product offerings and to remain relevant to their customers. To do so, organizations need to align their innovation strategies with their risk appetite, anticipate major pitfalls and embed the right governance structures into transformation programs. For this to succeed, executive buy-in and oversight is paramount to AI success.

Author's Note

The author thanks Gina Francis, Innocent Ndoda, Shingi Muvenge, Kathleen Lo and Andrew Strong for their valuable feedback, which helped to improve this article.

Endnotes

1 Drucker. F. P.; "The Manager and the Moron," *McKinsey Quarterly*, 1967, *www.mckinsey.com/business-functions/organization/our-insights/the-manager-and-the-moron*
2 IBM Corporation, "Memorial Sloan-Kettering Cancer Center: IBM Watson Helps Fight Cancer With Evidence-Based Diagnosis and Treatment Suggestions," January 2013, *www-935.ibm.com/services/multimedia/MSK_Case_Study_IMC14794.pdf*
3 Alexander, D.; "Bank of Montreal Jumps Into Robo-Advising Ahead of Other Lenders," Bloomberg Technology, 18 January 2016, *www.bloomberg.com/news/articles/2016-01-18/bank-of-montreal-jumps-into-robo-advising-ahead-of-other-lenders*
4 Conner-Simons, A.; "System Predicts 85 Percent of Cyber Attacks Using Input From Human Experts," Massachusetts Institute of Technology,

USA, 18 April 2016, *www.csail.mit.edu/System_predicts_85_percent_of_cyber_attacks_using_input_from_human_experts%20*

5 Bank of America Merrill Lynch, "Thematic Investing: Robot Revolution—Global Robot and AI Primer," press release, November 2015, *www.bofaml.com/content/dam/boamlimages/documents/PDFs/robotics_and_ai_condensed_primer.pdf*

6 Wile, R; "Venture Capital Firm Just Named an Algorithm to Its Board of Directors—Here's What It Actually Does," *Business Insider*, 14 May 2014, *www.businessinsider.com.au/vital-named-to-board-2014-5*

7 Philips, M.; "Knight Shows How to Lose $440 Million in 30 Minutes," Bloomberg, 2 August 2012, *http://www.bloomberg.com/news/articles/2012-08-02/knight-shows-how-to-lose-440-million-in-30-minutes*

8 Sotala, K.; "Concept Learning for Safe Autonomous AI," Machine Intelligence Research Institute, 2015, *www.aaai.org/ocs/index.php/WS/AAAIW15/paper/download/10131/10137*

9 Massachusetts Securities Division, "Robo-Advisers and State Investment Advisor Registration," Policy Statement, 1 April 2016, *www.sec.state.ma.us/sct/sctpdf/Policy-Statement--Robo-Advisers-and-State-Investment-Adviser-Registration.pdf*

10 *Ibid*.

11 Mickle, P.; "1961: A Peep Into the Automated Future", *www.capitalcentury.com/1961.html*

12 Kelly, E.; "Computing, Cognition and the Future of Knowing How Humans and Machines are Forging a New Age of Understanding," IBM, 2015, *www.research.ibm.com/software/IBMResearch/multimedia/Computing_Cognition_WhitePaper.pdf*

13 Soares, N.; B. Fallenstein; "Aligning Superintelligence With Human Interests: A Technical Research Agenda," Machine Intelligence Research Institute, 2015, *https://intelligence.org/files/TechnicalAgenda.pdf*

14 *Ibid*.

15 Autor, H. A.; "Why Are There Still So Many Jobs? The History and Future of Workplace Automation," *Journal of Economic Perspectives*, 2015, *http://pubs.aeaweb.org/doi/pdfplus/10.1257/jep.29.3.3*

16 Aiken, C.; D. Galper; S. Keller; "Winning Hearts and Minds: The Secrets of Sustaining Change," McKinsey & Company, 2011, *www.ru.is/media/opni/frettir/Winning-hearts-and-minds-McKinsey.pdf*

17 Manyika, J.; M. Chui; J. Bughin; R. Dobbs; P. Bisson; A. Marrs; "Disruptive Technologies: Advances That Will Transform Life, Business, and the Global Economy," McKinsey Global Institute, 2013

18 Frey, C. B.; M. A. Osborne; *The Future of Employment: How Susceptible Are Jobs to Computerisation?,* 17 September 2013, *http://www.oxfordmartin.ox.ac.uk/downloads/academic/The_Future_of_Employment.pdf*

19 *Op cit*, Bank of America Merrill Lynch

20 Verizon, *2016 Data Breach Investigations Report*, 2016, *www.verizonenterprise.com/verizon-insights-lab/dbir/2016/?utm_source=pr&utm_medium=pr&utm_campaign=dbir2016*

21 Bailey, T.; J. Kaplan; A. Marcus; D. O'Halloran; C. Rezek; *Beyond Cybersecurity: Protecting Your Digital Business*, Wiley, USA, 2015, *www.wiley.com/WileyCDA/WileyTitle/productCd-1119026849.html*

22 The Open Web Application Security Project, "Category: OWASP Top Ten Project," 2016, *https://www.owasp.org/index.php/Category:OWASP_Top_Ten_Project*

23 National Institute of Standards and Technology, "Framework for Improving Critical Infrastructure Cybersecurity," USA, 2014, *https://www.nist.gov/sites/default/files/documents/cyberframework/cybersecurity-framework-021214.pdf*

24 ISACA®, *COBIT® 5 for Information Security*, USA, 2012, *www.isaca.org/COBIT/Pages/Information-Security-Product-Page.aspx*