

2020

Systemic Analysis of the use of Artificial Intelligence (AI) In Regulating Terrorist Content on Social Media Ecosystem Using Functional Dependency Network Analysis (FDNA)

Alaina Roman
Old Dominion University

C. Ariel Pinto
Old Dominion University

Follow this and additional works at: <https://digitalcommons.odu.edu/ourj>



Part of the [Risk Analysis Commons](#), and the [Systems Engineering Commons](#)

Recommended Citation

Roman, Alaina and Pinto, C. Ariel (2020) "Systemic Analysis of the use of Artificial Intelligence (AI) In Regulating Terrorist Content on Social Media Ecosystem Using Functional Dependency Network Analysis (FDNA)," *OUR Journal: ODU Undergraduate Research Journal*: Vol. 7 , Article 7.

DOI: 10.25778/75E6-2612

Available at: <https://digitalcommons.odu.edu/ourj/vol7/iss1/7>

This Article is brought to you for free and open access by ODU Digital Commons. It has been accepted for inclusion in OUR Journal: ODU Undergraduate Research Journal by an authorized editor of ODU Digital Commons. For more information, please contact digitalcommons@odu.edu.

Systemic Analysis of the use of Artificial Intelligence (AI) In Regulating Terrorist Content on Social Media Ecosystem Using Functional Dependency Network Analysis (FDNA)

Cover Page Footnote

This research is supported in part under NSF grant DGE-1723635

SYSTEMIC ANALYSIS OF THE USE OF ARTIFICIAL INTELLIGENCE (AI) IN REGULATING TERRORIST CONTENT ON SOCIAL MEDIA ECOSYSTEM USING FUNCTIONAL DEPENDENCY NETWORK ANALYSIS (FDNA)

By Alaina Roman* and C. Ariel Pinto

I. FRAMING THE PROBLEM

In 2019 there were over three billion social media users worldwide. The number only continues to grow as the age in which individuals post their first online introduction decreases and the normalization of an online presence continues to increase. The popularity of social media has introduced a variety of problematic scenarios that society has never encountered historically. Its ubiquity has been exploited for malicious purposes like facilitating terrorism, enabling exploitation of vulnerable members of society, unduly influencing political processes, and other malicious purposes. The sheer volume of information and covert behavior of malicious agents makes identifying, removing and blocking malicious content from social media platforms a great challenge. One emerging technical solution is artificial intelligence (AI). This solution comes with issues that reach beyond AI technology itself and permeates other areas of the larger ‘ecosystem’ where the issue exists in society – policies, laws & regulations, international relations and economics. This research examined artificial intelligence being used on social media platforms to regulate terrorist content using the risk analysis concept using Functional Dependency Network Analysis (FDNA).

* This research is supported in part under NSF grant DGE-1723635.

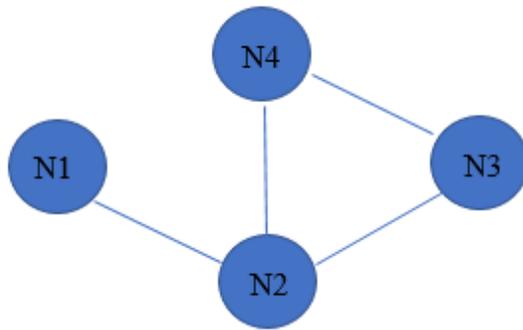


Figure 1: Sample FDNA modeling schema

1.1. Functional Dependency Network Analysis

Functional Dependency Network Analysis or FDNA is a risk management tool that identifies dependencies and critical points within a general network – including the social media ecosystem. FDNA was originally developed to identify and calculate risks in large scale information system networks by Garvey and Pinto (2009), and later extended by Ozdemir, et al. (2019). FDNA allows an infrastructure with several cooperating counterparts to be examined for risk as a whole subject (Servi & Garvey, 2017).

For example, Figure 1 displays four different nodes connected to one another (N1, N2, N3, N4). Using the FDNA algorithm you can determine how dependent each node is on one another. You could also calculate which node or node(s) is the critical point(s) in the infrastructure. A critical point is an element or node within the network/organization that's proper function is required to maintain the overall purpose of the larger scale organization. A critical point can also be identified as a point of failure in a network. Node 2 would be the critical point of the organization because without Node 2 working properly the other nodes would fail their purpose as well. FDNA will additionally conclude through computation that nodes N3 and N4 are less dependent on N2 compared to N1 node's dependency on the critical point N2. The

use of FDNA as a tool allows for a deeper understanding of an organization with various counterparts. The FDNA concept will be applied to this research as a tool used to depict and break down the social media ecosystem, terrorists' interaction online, elements of AI, and cyber ethics.

1.2. AI Ecosystem

Artificial intelligence has been around in different forms and theories since the 1950s. As the evolution of AI theories and logistics continues to develop, the definition and expectations of AI has changed. In its more recent form, artificial intelligence is the simulation of human styled intelligence through the use of data collection and algorithm. This emerging technology is being used for atomization, analysis of big data, and improvement of business analytics. This is why the use of a risk management concept, FDNA, is being applied to examine the relations of all components of AI in a scientific research manner.

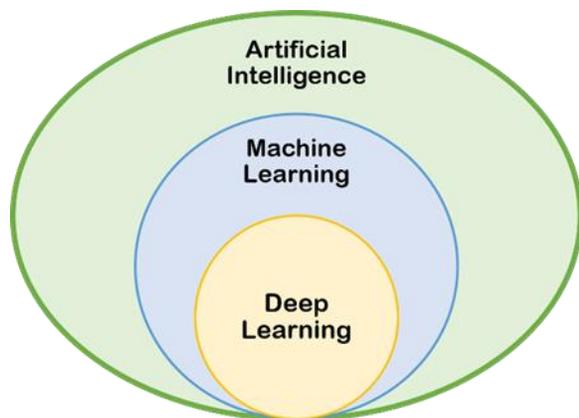


Figure 2: Distinct layers of AI

AI can be broken down into three distinct layers: Machine Learning, Deep Learning, and AI (represented in Figure 2). The term AI has been glamorized by society and used to reference all layers. Understanding each layer, will create an understanding of the needs of the AI ecosystem.

The current issues with AI technology can be summarized as a seven-point criteria as described in Ethics Guidelines for Trustworthy AI (HLEG, 2019). They are enumerated and briefly described in below:

1. Human agency and oversight, e.g., AI should enhance human welfare.
2. Technical robustness and safety, e.g., technological integrity and resilience is assured.
3. Privacy and data governance, i.e., assured data confidentiality, availability, and integrity of data
4. Transparency, i.e., algorithm is documented and traceable to humans.
5. Diversity, non-discrimination and fairness, e.g., do not create nor reinforce unfair bias in the use of input data as well as algorithm design.
6. Societal and environmental well-being, e.g., sustainable and environmentally friendly.
7. Accountability, e.g., documented trade-offs and ability for redress for the aggrieved

II. GENERAL ECOSYSTEM MEMBERS

2.1. Cybersecurity Risk Management

Risk can be generally described as the departure from the *desirables* (Pinto, et al., 2012). Risk management is used to evaluate these *undesirables* in terms of their likelihood and consequences, and the overall measures to address them. Risk management is applied to cybersecurity when an organization or groups in the larger society evaluate specifically their cyber-related risks. For example, a social media platform should apply cybersecurity risk management to their organization and information system infrastructure. Being able to calculate risk will enable a company to administer security measures to reduce risk. The cybersecurity risk management discipline will be used in this research to identify the effectiveness of AI in terms of

risk evaluations. A risk management approach will also be used as a means of calculating problems and risks between different disciplines.

2.2. Ethics of Cybersecurity

Ethics is the discipline that governs human behavior by creating morals. Morals implemented into information systems and frameworks is called cyber ethics. Cyber ethics is applied through policy, user behavior interactions, and within the application programming. An example of cyber ethics would be the application of censorship. This has stood as an ethical problem for many social media platforms. Censorship is an authoritative decision which determines what can be said, shared, or spoken about. Cyber ethics is applied to this scenario in two different applications. One, cyber ethics theory and concept will debate whether it is morally sound for one authoritative figure to determine what can and cannot be said. The other way it could be applied is how to ethically determine what is acceptable to censor. Cyber ethics is a complex and interdisciplinary concept that combine cybersecurity and philosophical theories with the intent on creating moral accountability in the online environment.

2.3. Laws and Regulations

Technology is a fast-paced and ever-changing field. Legislation on the other hand is not. It is relatively slow and static for periods of time compared to the rapid changes in the form and usage of technologies like AI. This can result to conflict when new technology such as AI or social media platforms progressively develop and change at an exponential rate. This contributes to the difficulty of applying legal ramifications to enforce ethics in newer cyber concepts. In addition to the inability for law to keep up with changing cyber law demands, the cyber world is international. This means that there will be a diversity of laws and norms attempting to compromise with other perspectives on law and ethics. One nation can have vastly different

opinions on what is ethically correct for the cyber environment. While nations can create and implement legislation within their borders effectively, domains, websites, and social media platforms are not always confined to specific borders. Therefore, conflict arises when international entities such as social media interact with multiple borders. Legislation must be broad and have the ability to change in order to effectively work--as well as international agreement, which is near impossible. Overall, this is why cyber law, until recent, has been a huge grey space for law and regulation.

2.3. Threat of Cyber Terrorism

Criminals have been exploiting the cyber environment since its early development. If done correctly, the cyber environment can be used to anonymously communicate and interact at an international level. One area the cyber environment and social media platforms have unintentionally aided is terrorism. It has coined the term cyber terrorism. The objective of terrorism is to spread fear to particular groups or populations. Terrorist groups used to have to find secretive methods of in person communication, recruiting, and organizing. Now terrorist groups can do it anomaly and internationally. This means the cyber environment and the use of social media platforms have been abused by terrorist groups to internationally conduct their business. They can now recruit people from across the world and have access to various resources. Terrorist groups being able to operate at this level possesses a greater threat than ever before. Agencies who monitor, detect, and prevent terrorism have evolved dramatically over the past years to accommodate these changes. Ethically, debates have been made about whether large cyber organizations such as social media should assist or are required to help prevent terrorism (Dean, Bellm & Newsman, 2012). Social media is the platform in which terrorists can easily hide and communicate. Therefore, if social media platforms have legal obligations to

detect and restrict terrorist activity, it would greatly affect a terrorist group’s efficiency. Overall, cyber terrorism has impacted the capacities of terrorism forcing additional laws and regulations to be created.

III. DESCRIBING THE IDEALIZED PROBLEM-SOLVED SCENARIO

A significant aspect of systemic risk management is the explicit narrative of the idealized or (i.e. problem-solved) scenarios for various members of the system, which for the case of this research are the members of the AI ‘ecosystem’. The checklist below summarizes the idealized, problem-solved scenario for each of the ecosystem members described in the earlier sections (i.e., which AI characteristics will be necessary for the satisfaction of the interests of ecosystem members).

Idealized AI Characteristics	Risk Management	Ethics	Laws and regulations	Security (cyber)
Human agency and oversight		✓	✓	
Technical robustness and safety				✓
Privacy and data governance		✓		✓
Transparency		✓	✓	
Diversity, non-discrimination and fairness		✓		
Societal and environmental well-being		✓		
Accountability	✓		✓	

IV. REPRESENTATION OF ECOSYSTEM AS FDNA

To create an FDNA representation for a complex multidisciplinary problem, “nodes” must be identified. In this scenario, the problem was further broken down into ecosystems. An ecosystem in this case is a community or organization of interacting elements. Social media platforms, artificial intelligence technology, cyber ethics, and cyberterrorists are the segmented ecosystems that will be examined using FDNA. Figure 2 is the representation of an artificial intelligence ecosystem connected to the various other ecosystems. The complexity of the issue allows for various FDNA depictions of the issue to be presented from different perspectives. This, however, does not affect the determinations of dependencies and critical points calculated. Using this representation in Figure 3, the dependencies between the ecosystems are clear and important elements to each ecosystem’s function are displayed. Since AI is the critical point in this scenario, it can be concluded that without a proper functioning AI, the other displayed ecosystems experience a change in their overall function. For example, if AI was applied to a social media platform but was significantly ineffective, all other ecosystems would experience an impact. A rise in cyberterrorism could occur along with rejection of technological methodologies, the social media platform’s corruption of moral, and an overall impact on the platform’s business structure. AI’s inability to perform its functions, debilitates the other ecosystems to preforms their function because all dependencies on AI as the critical point are strong. To further understand dependencies derived from Figure 2, an examination between the dependence of AI, ethics, and the system as a whole is necessary. If the ethics ecosystem was completely removed from the system, AI’s function would be impacted. Yet, the dependency alone to ethics is not valued enough to damage the other ecosystem’s functions. AI without the application of ethics would shift the function capabilities of AI. AI’s connections to other

ecosystems will allow it to maintain an acceptable amount of function. A deeper examination of the dependencies illustrated below will allow for an understanding of a complex scenario that will allow for proposals of future solutions.

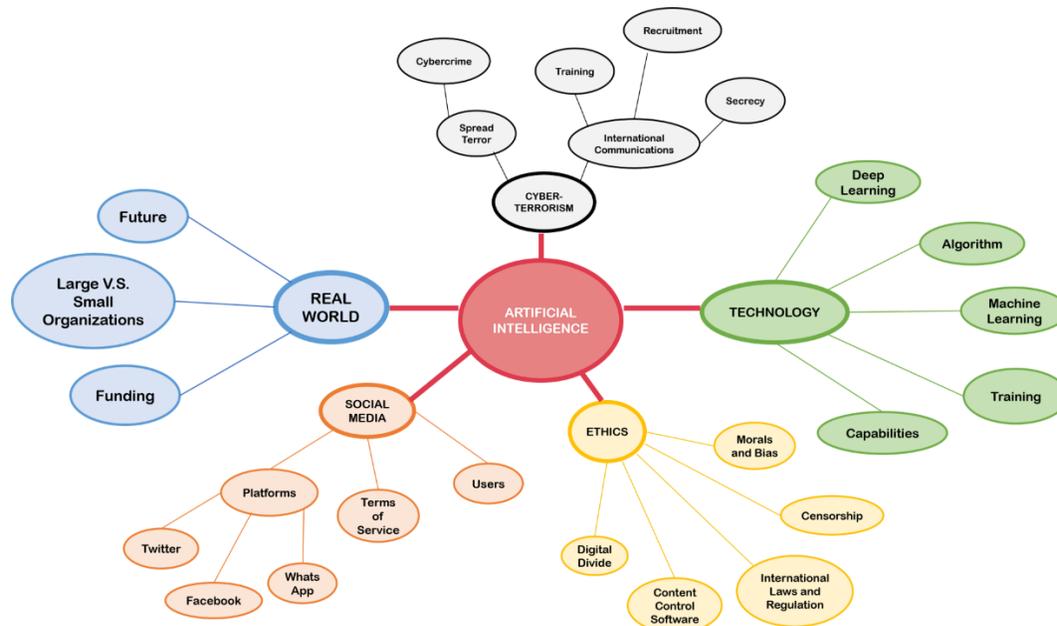


Figure 3

V. CASE STUDIES

The use of applying knowledge to case studies is to examine the issue while applying real world factors. All information derived is speculation and not intended to cause blame or fault for any examined parties. This is not meant to spread a story, but to identify key factors that can be taken as prevention methods in future incidents. The purpose of exploiting the case study is to solidify, from an FDNA perspective, true critical points and dependencies of the separate ecosystems functioning as a whole system.

5.1. Case Study 1: Facebook

Facebook is a social media platform that is used internationally to connect with friends and family through sharing posts, instant messaging, and video chat capabilities. Similar to other social media platforms, Facebook has created its own Terms of Service, user agreements, and legislation. Facebook has had a controversial past with the ethics and legislation behind their Terms of Service (TOS) and their user agreement which they call “Community Standards.” The TOS of Facebook clearly outlines that users are agreeing to their conditions. Any user that wants to be part of the platform must sign the TOS. Facebook’s user agreement or Community Standards specifically states a user’s posted content becomes property of Facebook. Hence, if a user posts an image to Facebook, legally Facebook has rights over the image. Officially stated it in the Community Standards as follows:

...when you share, post, or upload content that is covered by intellectual property rights (like photos or videos) on or in connection with our Products, you grant us a non-exclusive, transferable, sub-licensable, royalty-free, and worldwide license to host, use, distribute, modify, run, copy, publicly perform or display, translate, and create derivative works of your content (consistent with your privacy and application settings). This means, for example, that if you share a photo on Facebook, you give us permission to store, copy, and share it with other content you delete may persist for a limited period of time in backup copies (though it will not be visible to other users). In addition, content you delete may continue to appear if you have shared it with others and they have not deleted it (Newcomb, 2018).

Additionally, Facebook states that they are the determining authority on what is acceptable to post, leaving the statement vague and open for their interpretation. They use phrases such as

“good faith belief” to leave the interpretation up to Facebook. Also, when signing Facebook’s Community Standards, users are agreeing to their information being saved, stored, and compiled in a database to be bought or used in the future. Facebook asks and informs users through TOS and their Community Standards of their personal standards and policies. When either document is updated, users will have to review and accept it again to continue participating on the platform. Lastly, Facebook closes their Community Standards with the following statement: “at times we will allow content that might otherwise violate our standards if we feel that it is newsworthy, significant, or important to the public interest. We do this only after weighing the public interest value of the content against the risk of real-world harm (Newcomb, 2018).” This is Facebook retaining control by officially stating that they reserve the right to ultimately decide what content may be public. Overall, Facebook uses their TOS and Community Standards to control the content on their platform as well as legal protection, including international laws.

As stated prior, Facebook is well-known for being involved in many cybersecurity, cyber ethics, and cyber world scandals. It was also stated that Facebook has video capabilities on its platform. Facebook Live Stream is a Facebook specific feature that allows users to live stream to the Facebook platform. Its purpose was to create real-time connections between users. It has grown to be used by startup businesses, groups, and individuals. The tool was designed to be user-friendly and open for all Facebook users. Soon after the tool was released, incidents of abusing live stream capabilities started to occur. In the year 2017, an influx of incidents were reported on Facebook Live Stream. Individuals were live streaming criminal activities such as rape, murder, mental and physical abuse, and suicide. Well-known incidents include the Uppsala Rape, Robert Godwin Shooting, and the Wongtalay Family murder-suicide (Newcomb, 2018). These events are among the many incidents to occur from 2017 to present day. Facebook has

been pushed by the public to grasp a handle on regulation and prevent these types of incidents on their platform. Facebook's goal is to create an interactive and connective environment for their users, so the idea of creating restrictions on the platform was not ideal. Facebook used flagging through user reports, and deletion of any copied or shared posts to satisfy the public need for proper regulation.

In addition to dealing with the Live Stream incidents, Facebook has been undergoing international legal conflict. Cyber law has begun cracking down on international platforms such as Facebook. In January of 2018, Europe's data protection laws enforced that Facebook must respect their cyber laws and apply it to their European users. In February of the same year, courts from Germany filed against Facebook due to the platform not legally confirming to the users that they could collect their data and clarifying the purpose for which it is collected. Germany has data privacy legislation that prohibits business from taking data from users unknowingly. Within the same month, Belgian courts were also taking Facebook to court. Belgian accused Facebook of tracking users across the Internet after leaving their Facebook page (Newcomb, 2018). In March of 2018, the Federal Trade Commission launched an investigation against Facebook. Then in October of the same year, the UK Information Commissions Office fined Facebook the max punishment. Facebook was then called out publicly the following month by several lawmakers from different countries for not attending an international committee hearing on misinformation (Newcomb, 2018). Between these two events, Facebook released its first version of its Community Standards documentation. As cyber laws continue to catch up with the cyber environment, large platforms will continue to have to adapt and implement new legislation in order to stay afloat.

5.2. Case Study 2: New Zealand Terror Attack

On March 15, 2019 at the Al Noor Mosque in New Zealand, a terror attack took place. During Friday prayer, Brenton Tarrant attacked and killed 51 people, live streaming the whole incident on Facebook Live Stream. The live stream lasted 17 minutes, and was not flagged until 12 minutes after the video had ended (Facebook, 2019). Then the police were alerted of its presence on Facebook, to which the police then had to notify Facebook. Facebook took the video down and began deleting any reposts. There was international upset and heartbreak over not only the incident but the manner in which Facebook had dealt with the issue. Many were troubled by Facebook's AI's inability to detect the mass shooting on Facebook Live stream. Society was additionally troubled by the Facebook's delay in any type of response to the incident, such as how long it took to completely remove all copies from the platform—especially due to Facebook's history with Facebook Live Stream incidents. Facebook has been using AI technology to help detect and monitor activity on their platform. The big question is why Facebook's AI efforts were unable to detect a mass murder being live streamed. AI can determine what images are using machine learning and can break down images with representation layers. This takes processing power and time for the AI to identify images. Videos are multiple frames of images which take even more time for an AI to digest—nonetheless live streams. This is the first problem with monitoring a platform with current AI technology. It is unable to congest the amount of content large platform such as Facebook hold. Facebook's AI was also unable to detect the live stream due to the style of content. The mass shooting was recorded from the terrorist himself; therefore, it was presented as a first-person shooter style video. Other innocent content in first-person shooter format is allowed to be live streamed such as video games. Yet, this is misleading for AI technology; it is unable to determine distinctive

factors between first-person shooters games and first-person shooters in a real-life harmful instance. Lastly, Facebook released a statement after the terrorist attack as to why the incident was not identified sooner (Further, 2019). Facebook indicated that was due to misconfigure the reporting logic of flagged video. After the video was flagged, it was not categorized into priority to where it would be reviewed by Facebook staff. Therefore, Facebook has had to re-evaluate their logic, AI training, and algorithm behind their AI technology.

After this incident, social media platforms have been altering their AI techniques to deplete the presence of terrorist activity online. Laws and regulation are still ethically discussed to understand and morally instill a sense of responsibility to prevent these types of incidents. New Zealand has enacted legislation that aids in stopping the story of the terrorist (Holmes, 2019). They did not release the manifesto, speak about the shooter's story, and focused on victim's recovery. This is to prevent the terrorist from making a lasting impression on anyone else, reducing the likelihood of future incidents. Lastly, through this case study we can determine the critical points and dependencies. It can be concluded that AI technology needs adequate legislation that is ethically backed to be more functional. It can also be concluded that terrorist content has the potential to be regulated if the critical points of technology and ethics are met.

VI. SUMMARY AND CONCLUSION

AI has been discussed as an emerging tool which will greatly benefit society. Yet, AI technology has proven not to be marketable for all consumers currently. An ethnically-based discussion has been presented regarding whether the AI ecosystem is stable based on the seven AI criteria for idealized scenarios (i.e., Human agency and oversight, Technical robustness and safety, Privacy and data governance, Transparency, Diversity, non-discrimination and fairness, Societal and

environmental well-being, and Accountability). Ecosystem members like social media platforms could benefit from automation and regulation brought by AI, but only to specific extents. On the other hand, viability to some ecosystem members may still be ambiguous due to doubts in the robustness and ethics associated cost of AI tools. AI is a powerful tool that can be beneficial in many aspects of society, but it still faces challenges in sustaining a viable ecosystem.

REFERENCES

- Byman, D., Chambers, S. T., Isacson, Z., & Meserole, C. (2018). Regulating Internet Content: Challenges and Opportunities. doi:10.14324/111.9781787351714
- Britton, B. (2019, March 16). How the New Zealand terror attack unfolded. Retrieved August 20, 2019, from <https://www.cnn.com/2019/03/15/asia/new-zealand-christchurch-attack-what-we-know-intl/index.html>
- Dean, G., Bell, P., & Newman, J. (2012). The dark side of social media: review of online terrorism. *Pakistan Journal of Criminology*, 3(3), 103-122.
- Facebook, (2019, March 20). A Further Update on New Zealand Terrorist Attack. Retrieved August 19, 2019, from <https://newsroom.fb.com/news/2019/03/technical-update-on-new-zealand/>
- Garvey, P. R., & Pinto, C. A. (2009). *Introduction to functional dependency network analysis*. Paper presented at the The MITRE Corporation and Old Dominion, Second International Symposium on Engineering Systems, Massachusetts Institute of Technology, Cambridge, Massachusetts.
- Holmes, L. (2019, April 30). Youth Suicide Rates Increased Following '13 Reasons Why' Debut. Retrieved August 20, 2019, from https://www.huffpost.com/entry/13-reasons-why-suicide-increase-report_1_5cc712e0e4b04eb7ff99514f
- HLEG, A. (2019). High-Level Expert Group on Artificial Intelligence. Ethics Guidelines for Trustworthy AI.
- Minsky, M. (1992). Future of AI technology.

- Newcomb, A. (2018, March 24). A timeline of Facebook's privacy issues - and its responses. Retrieved August 19, 2019, from <https://www.nbcnews.com/tech/social-media/timeline-facebook-s-privacy-issues-its-responses-n859651>
- Ozdemir, Halil I; Pinto, C Ariel; Unal, Resit; Keating, Charles B; Britcher, Colin (2019). Supporting technology selection via portfolio readiness level and technology forecasting, Proceedings of the International Annual Conference of the American Society for Engineering Management; Huntsville, Al.
- Perez, S. (2016, March 24). Microsoft silences its new A.I. bot Tay, after Twitter users teach it racism [Updated]. Retrieved August 20, 2019, from <https://techcrunch.com/2016/03/24/microsoft-silences-its-new-a-i-bot-tay-after-twitter-users-teach-it-racism/>
- Pinto, C.A., McShane, M.K. and Bozkurt, I., (2012). "System of systems perspective on risk: towards a unified concept," *Int. J. System of Systems Engineering*, v.3:1, pp.33-46.
- Servi, L. D., & Garvey, P. R. (2017). Deriving Global Criticality Conditions from Local Dependencies Using Functional Dependency Network Analysis (FDNA). *Systems Engineering*,20(4), 297-306. doi:10.1002/sys.21394