DISSERTATION

QUALITY ATTRIBUTES OF DIGITAL TWINS

Submitted by

Jim Scheibmeir

Department of Systems Engineering

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Fall 2021

Doctoral Committee:

Advisor: Yashwant Malaiya

Erika Miller
Hilla Skiba
Thomas Bradley

ABSTRACT

QUALITY ATTRIBUTES OF DIGITAL TWINS

Digital twins are virtual representations of their physical counterparts and offer modeling, monitoring, and prediction as common conveniences. Digital twins can enable autonomy in physical systems and transformation in operations, such as Industry 4.0. Digital twins may also be utilized to better secure cyber-physical systems. Currently, digital twins are found to be compelling, but the definitions of digital twin lack standardization and concerns related to the costs and quality of digital twins create reluctance of adoption. While the emerging technologies being utilized within digital twins increase capabilities, they also generate concerns towards cost and quality. For example, IoT is utilized to inform digital twins, yet many IoT devices have low power constraints and do not have robust security mechanisms. There are also concerns about the interoperability of IoT devices and the replacement and upgrade costs throughout the digital twin's life cycle. Augmented reality is an emerging technology and has been suggested as a user interface for digital twins. Augmented reality enables digital models and scenes to be annotated onto the physical landscape. Fusing physical and virtual worlds is common to both augmented reality and digital twin technologies. However, scaling and sharing immersive experiences is problematic. While research in areas such as IoT, augmented reality, and digital twins are frequent, there are still questions about evaluating the quality of these technologies, the composition and security of digital twins, and their maturity. The main goal of this research is to establish the quality characteristics that digital twin applications should exhibit. We also provide a framework for the construction of application programming interfaces for digital twins and a novel approach for testing augmented reality applications utilizing computer vision. Further outcomes of our research include a benchmarking scorecard for IoT cybersecurity communications and a maturity model for digital twins. Social media analytics are used to reveal

the voice of stakeholders, further defining trends in the realm of digital twins. The social media analytics has shown a dearth of conversations regarding cybersecurity concerns. Organizations must protect their investments in digital twins and utilizing the frameworks, quality and maturity models, and scorecard within this research will reduce implementation risks.

.

ACKNOWLEDGEMENTS

Three individuals that I hold in high regard and with much respect wrote letters of recommendation to the graduate school in support of my acceptance. Rita Carney, Tom Brooks, and Ian Finley are role models and mentors. I appreciate their support, leadership, and the time and investment they made in me.

My advisor, Dr. Yashwant Malaiya, has demonstrated levels of patience for which I did not know could exist, and I now must exemplify. It is with absolute wisdom that Dr. Malaiya has been my advisor and offered guidance. I credit him with skills in ideation, knowledge creation, and managing research efforts that have helped me publish five times at the time of this writing. Dr. Malaiya has always pushed me to ask better questions and to provide better answers. Much of my growth in the last four years is because of his patience and ability to coach individuals to become more successful.

I have utilized many open-source libraries in my research. Indeed, too many to thank all the contributors individually. To those who contribute their software solutions to be used by researchers, thank you for your passion and contributions.

To my dissertation panel, Dr. Erika Miller, Dr. Hilla Skiba, and Dr. Thomas Bradley, I thank each of you for your leadership at our Colorado State University and for sharing your knowledge and guidance with me and with so many CSU Rams. I appreciate each of you for your investment in education, in my development, and in my future. I also acknowledge the support from the graduate school's administration staff, and specifically, Ingrid Bridge who has helped coordinate many student activities and procedures.

I apologize to my wife and children for the nights and weekends that I missed being dad and husband. I appreciate and thank my wife Carly for her forgiveness, patience, and support.

I love my family and thank them for their love and reassurance. I'm sorry for talking about my research so often and thank you all for the encouragement along the way.

TABLE OF CONTENTS

CHAPTER 1: API DEVELOPMENT MODEL FOR DIGITAL TWINS

Digital twins are a conduit between physical entities, virtual representations, and humans, geared towards using monitoring, modeling, and prediction to improve life [1]. The digital twin is becoming more important, and not just in understanding their physical counterparts, but also in controlling them [2]. Within the engineering process for digital twins, API designs and their execution and operation are critical [2]. These APIs as components are more important than ever in modern software engineering [3]. Here an approach is proposed that quickly produces a development model to meet growing expectations.

Software Development Model of Digital Twins 1

An optimal digital twin is virtual and can generate any behavior and control any state that would be accomplished by engaging its physical twin [4]. Such abstraction allows maintenance and reuse of both the physical devices and the software twin over time. The software acts as a proxy, allowing monitoring or controlling of the physical device.

Digital Twins Needs Analysis 1.1

Engineering a digital twin should be done in a modeled approach by performing a needs analysis from the physical device and its operating environment. From these needs, we can visualize the software proxy for the remote operations. Defining the system's functions and components begins with an objectives tree which then progresses into a contextual diagram. The branches of the objectives tree align to monitor and control and later translate into GET and POST type API transactions.

The physical device exists within an operational environment which is a critical design factor. The contextual diagram is key to understanding the relationship of the physical device to the environmental conditions and constraints, and for building the digital twin to facilitate these relationships and limitations. These concerns will be implemented as objects, functions, and components utilizing design tools and specifications such as OpenAPI, visual modeling and design tools such as Node-RED, or implementation and runtime tools such as Azure Digital Twins. OpenAPI designs list the various objects, paths, and operations of the API [5], creating documentation and hence helping to specify the contract between design and implementation.

Digital Twin Code Design 1.2

The proposed development model focuses on quality code design and testing early, or shift-left, through Test-Driven Development (TDD). The TDD practice detects defects early and reduces the cost of software development [6]. Software defects that leak into the production digital twin can be disruptive to the physical counterpart, and expensive to remediate.

By writing our test cases first, we create a design contract between what the successful implementation necessitates and the code that is later written. By creating and deploying the tests first, we create code that accomplishes two things. First, it focuses on quality by sufficing the test. Second, code that is easily unit tested is more reliably integrated with other code, improving the API usability. Going forward, during refactoring or other product enhancements, compliance with the design contract built through TDD will be achieved.

Performance Engineering 1.3

In a digital twin ecosystem, performance could become a limiting factor. Since the digital twin will monitor and control the physical twin, if performance is degraded, the physical device could suffer from untimely decisions in extreme environmental conditions. This could result in critical losses including loss of life, in manufacturing or medical systems. For this reason, performance testing is a requirement, and within our development model, we test performance iteratively, early, and often.

While HTTP is common in APIs, IoT introduces new protocols into the API and Digital Twin ecosystem. A performance engineering approach is required that can be applied to the various implementations.

API Test Framework 1.4

Beyond the previously established unit and performance testing, there exist vast areas where an API can experience failures and degradations. To limit these and control experience, we implement test contracts in the form of automated functional and nonfunctional tests. We interrogate how the API should behave, with a focus on service level agreements, security, and entitlement. Our functional tests assure that the operation of the API satisfies functional specifications and that our user has the correct permissions. The time to defect discovery and product costs can be reduced over time by introducing test automation [7]. Implementing test automation into the development model reduces some of the tedious and error-prone tasks, such as manual testing, defect analysis, and reporting.

The International Organization of Standards 25010 quality model is composed of eight attributes. One of these eight is the product's usability. This is frequently considered as a part of user interface testing. Because digital twins may be informed by many integrating IoT, we must assure the usability of the API for the maintainability and interoperability costs of the digital twin overtime. Table 1 outlines focus areas for API usability.

Table 1: API Usability Characteristics

Concepts and Artifacts	**Explanation of applicability and importance**
API Documentation	When starting a digital twin initiative, there may be many APIs from different providers. Proper documentation enhances usability and promotes initial use.
API Example Code	API Authentication examples are common.
Consistent Operation Naming	Method naming standards accelerate learning how to use the API by providing a repeatable verb and noun association
Consistent Return Types	Developers will expect an API method to return a consistent data type, even if a failure occurs.

Limited Arguments	Going above six arguments being passed into an API method will begin to deteriorate the usability of the API

API Mediation 1.5

Thirty-eight percent of respondents in a Gartner survey reported that APIs are "very important" to their digital platform [8]. If the API service is important, it should be secured, managed, and measured, and mediation is the solution for each of those concerns. API mediation is not new, but it is critical to the digital twin development model and is the deployment method of choice. Inner APIs attach directly to the digital twin. Outer APIs are exposed through mediation. Outer APIs can be composite endpoints created by joining two or more inner APIs. Using this mediation management layer allows us to better secure and maintain the integration experience through an additional layer of management. This mediation enables metering and measuring usage and performance.

Conclusion and Future Research 2

Our model begins with creating objective trees and context diagrams. These supply the starting points to creating architecture and specifications around the interfaces and performance characteristics. Our model then moves into Test-Driven Development and OpenAPI to design and code to a contract (Fig 1). Performance engineering is our next major initiative. Early feedback is critical as performance issues can stem from poor design, and design changes are more costly to correct than feature changes. After automated functional testing, we cover the deployment and operations of the API through mediation. This allows us to maintain the API and the digital twin, even to the point of swapping them out. This development model takes into consideration requirements, design, build, test, and operations, with a focus on automation and quality.

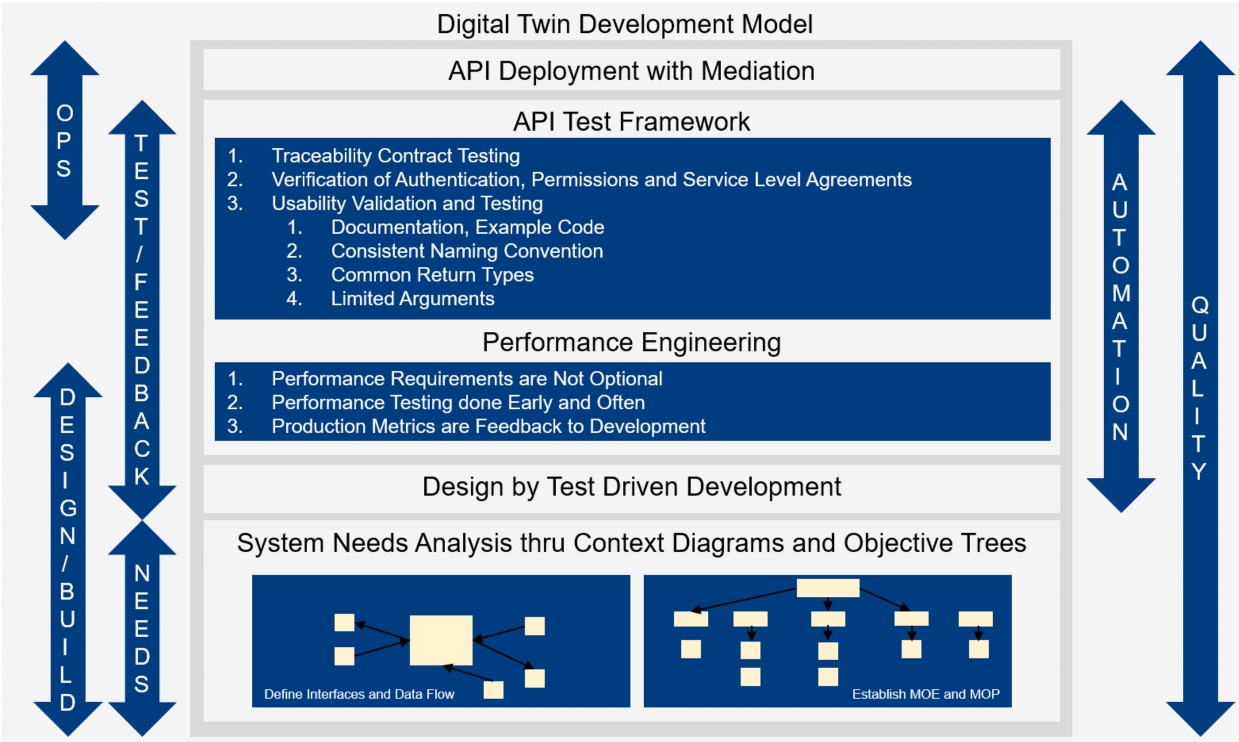

Figure 1: The Digital Twin API Development Model

This paper proposes a development model of a digital twin, from requirements and design, with a shift-left approach to testing early and often, up to the API mediation layer. The UI of the digital twin would then utilize the mediation layer. Further work is required to develop an effective user interface for the digital twin. The following chapter discusses the development and quality of augmented reality interfaces and applications.

REFERENCES

[1] El Saddik, A., (2018). Digital twins: the convergence of multimedia technologies." *IEEE MultiMedia*, 25, 87-92.

[2] Cearley, D., Velosa, A., Kerremans, M., Walker, M., & Burke, B. (2018). Top 10 strategic technology trends for 2018: digital twins. *Gartner IT Glossary*. March 08, 2018. Accessed October 07, 2018. https://www.gartner.com/doc/3867164

[3] Burns, C., Ferreira, J., Hellmann, T.D., & Maurer, F. (2012). Usable results from the field of API usability: a systematic mapping and further analysis. *IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*, https://doi:10.1109/vlhcc.2012.6344511

[4] Grieves, M., & Vickers, J. (2017). Digital twin: mitigating unpredictable, undesirable emergent behavior in complex systems. https://doi:10.1007/978-3-319-38756-7_4

[5] Di Martino, B., Rak, M., Ficco, M., Esposito, A., Nacchia, S., & Maisto, S.A., (2018). Internet of things reference architectures, security and interoperability: a survey. *Internet of Things,* 1-2:99-112. https://doi:10.1016/j.iot.2018.08.008

[6] Damm, L., Lundberg, L., & Olsson, D. (2005). Introducing test automation and test-driven development: an experience report. *Electronic Notes in Theoretical Computer Science,* 116(2005): 3-15. https://doi:10.1016/j.entcs.2004.02.090

[7] Kumar, D., & Mishra, K.K. (2016). The impacts of test automation on software cost, quality and time to market. *Procedia Computer Science*, 79;8-15. https://doi:10.1016/j.procs.2016.03.003

[8] O'Neill, M., Wong, J., Golluscio, E., & Resnick, M. (2018) API mediation is the key to your multiexperience strategy. Hype Cycle Research Methodology Gartner Inc. July 23, 2018. https://www.gartner.com/doc/3883663.

CHAPTER 2: QUALITY MODEL FOR TESTING AUGMENTED REALITY APPLICATIONS

Augmented Reality applications have the capability of merging virtual objects into a physical setting, or they can wrap physical objects within a virtual scene. Augmented reality applications are like virtual reality applications, the visualizations are computer-generated, but augmented reality apps also must contain a view of the physical world. Augmented reality applications are being utilized in service, manufacturing, product areas, as well as gaming. Mobile devices are becoming common runtime environments for augmented reality applications and mobile device proliferation is enabling a wave of AR applications. Due to the combined nature of digital and physical objects, as well as the environmental and contextual constraints, a traditional test plan is not sufficient. A new quality model is proposed that takes these issues into account, and an example of how machine learning can assist with aspects of the model is discussed.

Introduction 2.1

Augmented reality is accomplished by creating an experience that merges both physical surroundings as well virtual objects [1]. The term augmented reality was first used by Caudell and Mizell in 1992 [2]. The development of augmented reality applications has very much been democratized within the last decade, due to the rise of libraries that make building the applications easier, and the increase in compute power of devices such as mobile phones. Mobile phones have emerged into ubiquitous platforms as they fit and have been embraced into our lives [3]. Similarly, augmented reality has allowed us to keep a physical realm as it is, and yet decorate upon it digital objects. AR is becoming popular because it allows change to what we see and at the same time it allows the physical world to stay the same. According to one Gartner survey conducted in 2018 (Figure 2), ”According to a Gartner Research Circle study, 27% and 17% of respondents were either using or evaluating AR and MR technologies “[4].

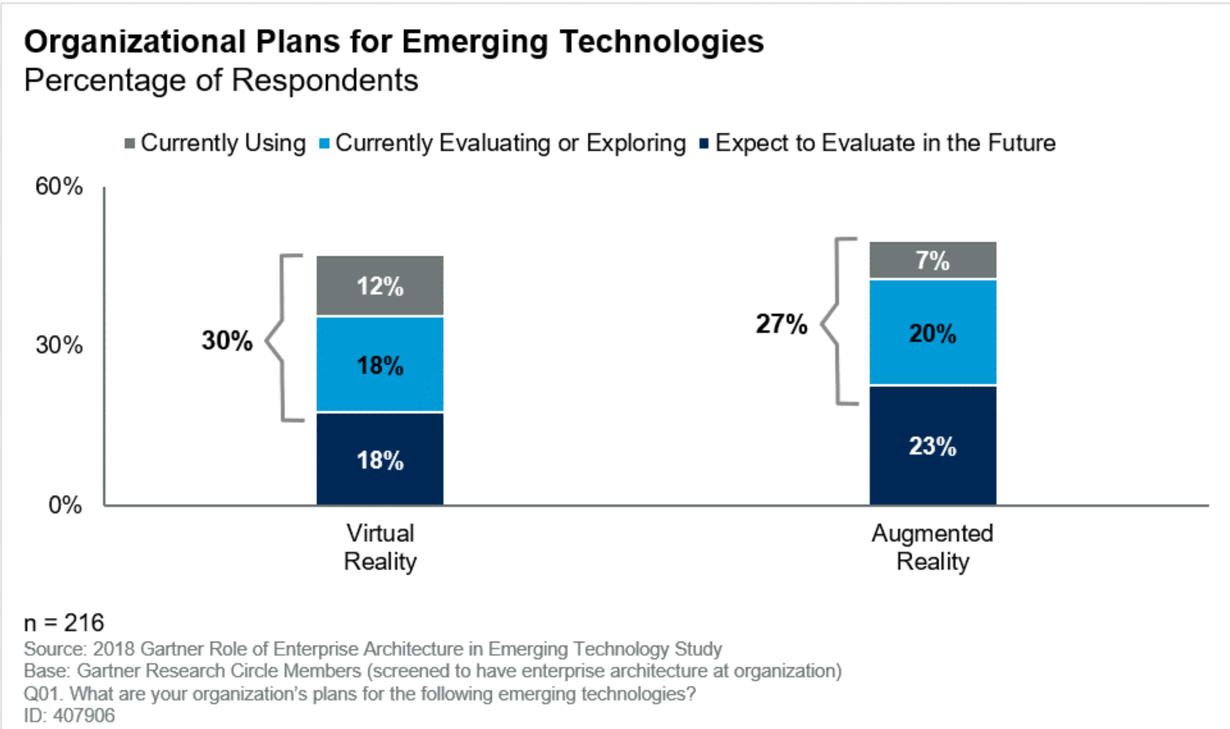

Figure 2: Gartner Emerging Technology Analysis: Augmented and Mixed Reality Opportunity for 3D Design Software and Vertical ISVs

An additional Gartner survey focused on multi-experience application development. The analysts stated, “It is, however, surprising to see VR apps identified as the second most impactful type of multiexperience app (20%), as AR has more potential use cases and device support. But only 14% of the respondents thought that AR apps would be the most impactful, despite AR app development tools being more widely available. “[5] (Figure 3).

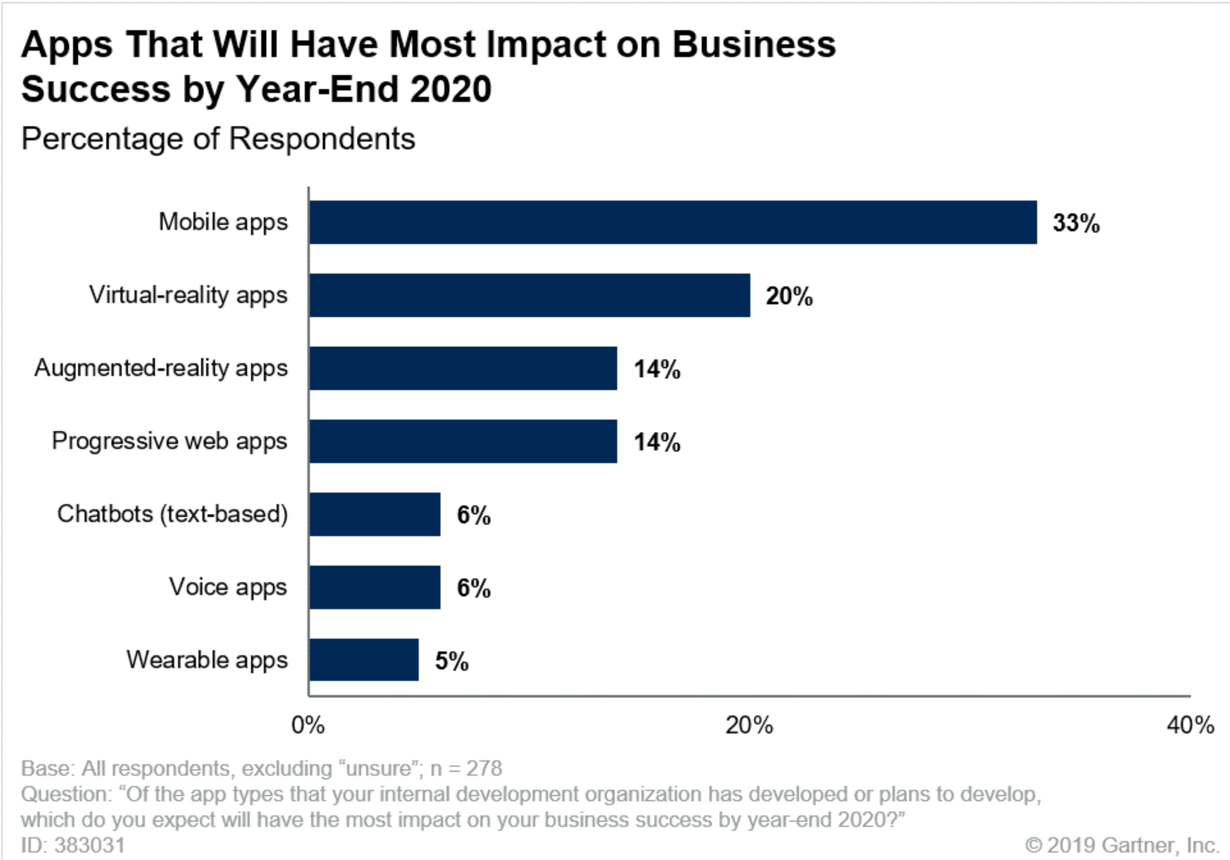

Figure 3: Gartner Survey Analysis: Insights to Kick-Start an Enterprise Multiexperience Development Strategy

There are several development tools to assist developers in the creation of AR applications. The ARToolkit was developed by Hirokazu Kato in 1999 which is now available for Windows, Mac, Linux, and Android platforms. ARCore, an SDK for Android and iOS phones, was developed by Google and released in 2018 [6]. Apple has its native tooling, ARKit, which was released with iOS 11 in 2017 [7]. Facebook released its Camera Effects Platform in 2017. In October of 2018, Facebook rebranded the platform to Spark AR [8]. Qualcomm released the QCAR SDK in 2011 [9], it then sold to PTC in 2015 for $65 million [10]. Maxst started as a company in 2010, released an AR mobile game in 2011, 5 years before the release of Pokemon GO, and released their AR SDK in 2012 [11]. Wikitude's JavaScript SDK was released in 2012 [12].

One driving factor of the popularity of mobile augmented reality applications, are the many SDKs available and the ease of use. Easy application development can also lead to applications having poor user experiences [13]. There are several choices of development tools, open-source, and commercial, which span across runtime environments. Table 2 lists several common examples of these types of augmented reality application development tools.

Table 2: Example Augmented Reality Development Tools

Tool name	Platform
Vuforia Studio [14]	Android, iOS, UWP, Unity
Spark AR [15]	Android, iOS
ARKit [16]	iOS
ARCore [17]	Android, iOS
Wikitude [18]	Android, iOS, Microsoft Tablet, smart glasses
Maxst AR [19]	Android, iOS, Windows, Mac, Unity

Industry usage of augmented reality technology is now becoming widespread. AR applications have been utilized in manufacturing, the services industry, and product design. AR has aided in the layout and design of factories [20]. Military machinery service and maintenance have also been enabled thru AR technology [21]. AR has also been utilized to assist in training new assembly skills [22]. In the automotive industry, Mercedes Benz utilized augmented reality to create a consumer application allowing potential buyers to customize a new car as it would appear in their driveway. In the first three months of this application being available to iOS users, there were over thirty-four thousand downloads and an average application rating of 4.3 out of 5 stars [23].

The development tools and the application of augmented reality are becoming more abundant. However, AR-specific testing tools and practices are lagging. One famous virtual reality headset, the “Sword of Damocles”, was invented in 1968 by Ivan Sutherland [24]. Yet, while VR technology has existed for a longer period than AR, there are no common test automation practices for application interactions in VR [25]. Figure 4 illustrates the continuum between AR and VR relationship and similarity [26].

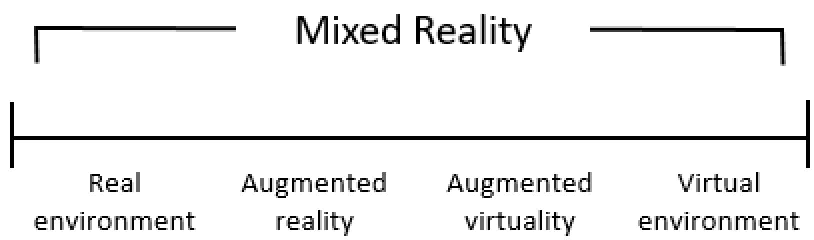

Figure 4: The Real-Virtual Continuum

While previous research has identified a potential void in testing practices for novel interfaces, that is not to say that industry users won't search them out or rely upon them. Test automation solutions for web and mobile are available from both open-source communities and commercial vendors. Test automation tools specifically for AR apps are not as easily found. Figure 5 identifies Gartner survey respondents' answers to a question about technology adoption specifically in support enterprise multiexperience app development. Gartner analysts in their analysis of the survey have stated, “The respondents had fairly equal adoption rates across the technology categories we asked about, but Gartner advises application leaders to emphasize adoption of test automation tools and DevOps practices.” [5].

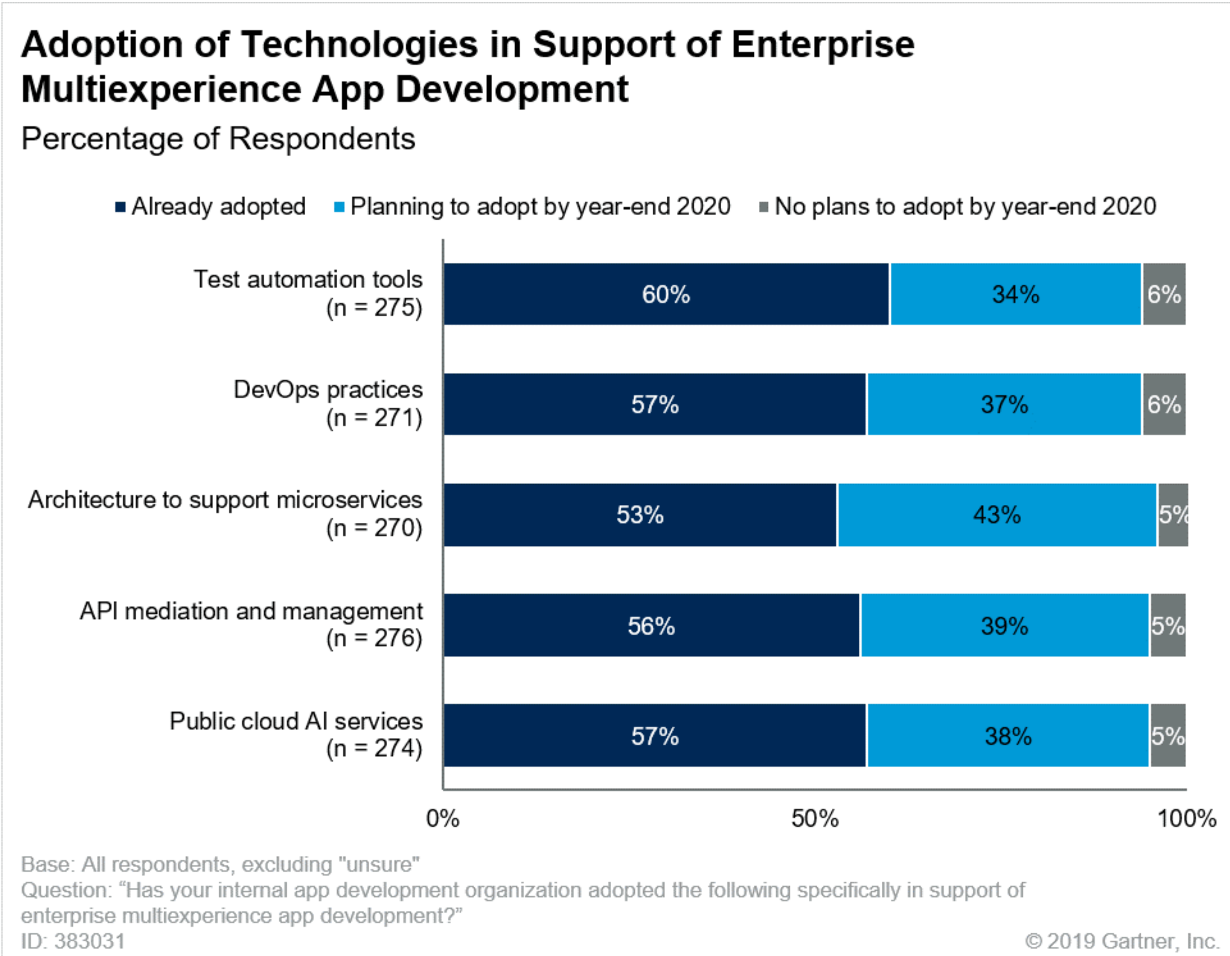

Figure 5: Gartner Survey Analysis: Insights to Kick-Start an Enterprise Multiexperience Development Strategy Findings on Adoption of Technologies in Support of Enterprise Multiexperience App Development

Assessing the quality of an AR application is more complex than the quality of a typical traditional software product. Early adopters of mobile AR applications have experienced bugs in production versions. Reports of quality issues found in Amazon's application ARView include poor experience in anchoring digital models to shiny surfaces, poor model detail, unnatural models, and incorrect proportion of models compared to the environments they are deployed into [27]. Ikea's AR application has had similar quality issues, mostly within low-light conditions, where digital models would float away from where they were deployed to [28].

Evaluating AR systems is difficult because we do not yet fully understand the expectations of end-users [29]. In addition, traditional testing methods are not up to the task of detecting defects in applications having novel interfaces [30]. Nielsen defined usability as software attributes of learnability, efficiency, memorability, satisfaction, and error [31]. AR usability requires both the physical and virtual components to work together. Thus, if either is degraded, such as in a low-light physical scenario, the AR application quickly loses usability. There is also the concern of color sameness and collision between models and physical environments. Presence and collaboration in an augmented environment are also new quality concerns compared to traditional testing procedures [32]. The application must be intelligent in these cases to instruct the user to improve the environmental settings or move to a more traditional application interface to solve their needs. This implication of an augmented reality application having such intelligence must be addressed at design.

Here an approach is proposed that produces a quality model to meet the growing expectations from the AR products. This model uses features from the ISO 25010 model.

Software Quality Model of Augmented Reality Applications 2.2

A well-designed augmented reality application blends the virtual and physical worlds into an experience that facilitates perspective, presence, interaction, and immersion [33]. Perspective is necessary for engaging virtual objects realistically within dynamic physical environments. Presence is the user's ability to be within their physical environment and engage in the virtual model or scene. This engagement is the interaction that occurs, as we manipulate the model, or as it changes concerning the perspective of our environment. This all works to create immersion, our ability to be in both physical and virtual presence at the same time. However, to scale AR, the application must also have persistence. This allows the augmented models and scenes to be experienced by other application users, scaling the augmented reality experience. These characteristics of augmented reality applications require the developers to go beyond traditional testing concerns.

Environmental Context as it Pertains to Functionality 2.2.1

Augmented reality applications must be tested for at least two distinct characteristics, occlusion, and collision [34]. Users of AR applications need realistic interactions between the physical and virtual objects. This means objects should maintain proper shadows, depth, and should not occupy the same space as physical objects at the same time.

Occlusion requires that if a virtual object is rendered further away compared to closer physical objects, the physical objects should hide or occlude the virtual object in a realistic manner. Occlusion is difficult because AR engines must properly represent the physical world, even as our relationship to it changes based upon physical movement. The AR engine must reconstruct the 3D model according to the properties of the physical world and render the dynamic user interface appropriately. Test cases must be created that mimic the possible multitudes of physical realms the application will run in, and then load and engage the virtual objects to test propensity to fail occlusion or create an experience that does not mimic reality.

A collision occurs when a virtual object approaches a physical object and attempts to occupy the same 3D space at a time. Many AR applications will allow the virtual object to be manipulated as it appears in the physical plane. However, the virtual object should not occupy the same space as the physical objects, despite the manipulation. Collision detection needs to be tested in an implementation. Test cases should also include how color and light change the ability to detect and prevent a collision.

Environmental Context as it Pertains to Visual Distinction 2.2.2

AR applications can lose visual distinction based upon the physical world they are operated within. Physical environments that make gradient distinction difficult, will reduce the usability of the applications. Engaging virtual objects is difficult and prone to fail in areas of intense light, where the color white is dominant, when the model and the physical scene seem chameleon in color, or when surfaces are smooth and reflective in general. Distinction testing must have multiple test cases that alter the physical environment to these types of extremes, to understand the application's behavior, and to potentially present the user with options outside of the AR user interface, such as a traditional menu-based interface, in the case failure to render virtual objects such that they are distinct.

Environmental Context as it Pertains to Portability 2.2.3

Another environmental concern to AR applications is the variations among the platforms (runtime OS and the computational and display/interaction hardware). There are many physical form factors, and versions of device operating systems, thus an AR application will need to be tested for correct execution across devices and platforms. This is where utilizing a device farm, as we would for typical web and mobile development, will help assure the breadth of device compatibility needed. For those utilizing AR applications internally to an organization, and having a policy on acceptable hardware, the concern for variations in the mobile runtime platform may be diminished. There are many mobile device farm vendors to select from, some having both on-premise and cloud options. Example vendors having both on-premise and cloud offerings include Experitest, Mobile Labs, and SmartBear, as well as others. Crowd testing could also be utilized to establish a large test bed of diverse devices. Crowd testing can be facilitated internally or externally to an organization, through crowd testing platforms. Crowd testing can accomplish quality goals such as user experience, functional or regression testing, and it offers diversity beyond devices, but also in geography and the skills and experience of end-users. Crowd testing platforms, as examples, could include RainforestQA, Applause, or testIO.

Performance Engineering 2.2.4

Three-dimensional rendering while operating camera hardware and merging the image into a user interface is computationally expensive. Early releases of Google's ARCore came with warnings in the release notes that to protect device resources and performance, digital models should be detached or deleted when not in use [35].

Augmented reality applications are often implemented using a distributed architecture, using both local clients and some cloud or centralized services. In these cases, we must do performance analysis on the centralized resources, with the understanding that the computational capacities of those assets may be stretched when many clients engage the services. The client devices themselves may also have performance concerns, including battery depletion over long runtimes. Traditional performance testing will solve much of this analysis, and those utilizing cloud resources will need to add validation of the complexity and cost of automatically scaled resources to test plans. Cloud performance can be degraded when there is high network traffic to the cloud's data center [36]. Augmented reality applications that scale and serve distributed experiences may generate higher than normal network traffic.

Scaling Usability with Persistence 2.2.5

Scaling augmented reality experiences across multiple users requires centralizing and persisting physical maps, virtual objects, and meta-data of both. By providing data persistence, multiple users on distributed systems can access the centralized store and then share the experience and immersion. ARKit and ARCore technologies both offer features to aid in the persistence of digital objects [37]. Persistence is a requirement of scaling users' ability to share augmented experiences. Persistence also plays into quality for single-user scenarios. A user can place the virtual object and then days later expect to see it embedded into the same physical location. Test cases will require an amount of complexity in the numbers of testers that engage in one shared scenario and experience, as well as the duration of time the test case covers.

Creating the Quality Model 2.2.6

Test planning for an augmented reality application release has many environmentally contextual factors. The virtual object needs to behave within acceptable limits of the physical world. The device runtime environment and form factor are variables as well. Depending upon the nature of the application, cloud technologies such as shared anchor points may be required, which can also lead to having more complex functional and non-functional quality requirements.

User experience is a very important and influential attribute of application quality [38]. Table 3 attempts to modernize aspects of the ISO 25010 model to reflect the user experience expectations of augmented reality applications. The model contains the quality characteristics of perspective, presence, interaction, immersion, and persistence. One previous research has indicated three characteristics of user involvement, the interaction between user, product, and other agents, and what can be observed or measured [39]. While another work found five characteristics of aesthetics, appeal, joy, usability, and utility [40]. The model represented in Table 3 is meant to reflect the ISO 25010 model and relate to the original eight quality characteristics as possible.

Table 3: Characteristics of Augmented Reality Software Quality

User Experience	**Description**
Functionality / Presence	Occlusion and Collision characteristics will make the AR application more realistic giving more functional uses. This characteristic allows users to be present in both physical and virtual worlds.
Visual distinction / Perspective	The environment, such as highly reflective and bright lighting may affect the visual distinction of the application's ability to inject virtual into physical consistently. Perspective is assured by testing for realistic rendering of virtual objects concerning their physical counterparts so that height, width, and depth are accurately maintained.
Performance / Interaction	Multiple anchors achieved by having many virtual objects in one physical layout can impact device performance. This can limit the user's ability to interact if interaction requires frequent or multiple virtual object creations, instances, or mutations. The performance also depends upon cloud anchor mechanisms so that multiple users can share augmented reality experiences.

User Experience	Description
Portability / Immersion	The many mobile form factors and operating systems create a complex runtime environment to be tested cross-device. For many users of an augmented reality application to experience immersion, the application must be compatible across many device types and versions.
Usability / Persistence	Scaling to achieve a shared augmented reality experience across multiple simultaneous users requires persistence. Persistence utilizes cloud technologies to enable distributed users to share anchor points and other object and scene metadata so that they can become immersed in each other's virtual and physical worlds.

Applying Automated Testing to AR Applications using Machine Learning 2.2.7

Automation helps accelerate testing so that we can consider production release sooner and so that we may find more bugs in a shorter amount of time. Like how traditional test practices do not consider novel user interfaces such as AR, testing tools are also lacking in this area. One case study by SmartBear is available to discuss their approach [41]. While they describe the physical mobile lab and physical robotics that reposition the devices allowing their cameras to adequately zoom and reset for incorporating the 3D model into the physical world, the case study does not portray how they tested the functionality and user experience as described in our quality model, or specifically how they tested for aspects such as collision or occlusion.

We present a framework that includes test automation, common tools and practices, and the use of machine learning to build an approach to testing that provides the probability of accurate model presentation in the physical world, to accelerate manual testers and developers so that they may quickly home in on problematic areas.

This approach begins with the models, their files, and the integrated development environment. Figure 6 is a partial screenshot taken from an IDE, Android Studio, and illustrates one example of how model files can be stored and organized during AR application development. These become a source of comparison to be utilized in the proposed automation framework.

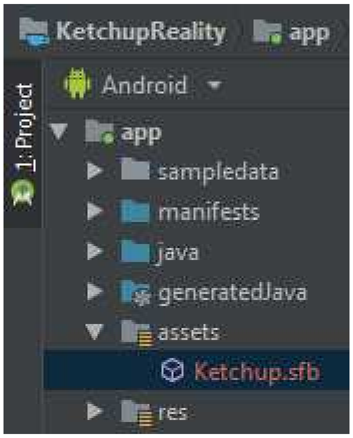

Figure 6: Android Studio Folders and sfb Model File

What is important about these models is the return value when they are sent into common ML image recognition engines. Our framework uses a common open-source test automation tool, Appium, to take screenshots during the application under test runtime. This suggested framework has access to the model within the development environment, as well as the screenshot of the model when utilized within the mobile app. Figure 7 is an image of the model within the development environment.

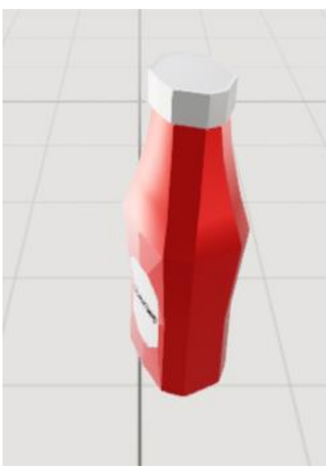

Figure 7: Model file of Ketchup Bottle

As a direct example of this framework, we can submit the model of the ketchup bottle as pictured in Figure 7 to the AWS Rekognition service. As indicated in Figure 8, the service returns a probability of 97.6% certainty that it is both Food as classification and it is identifiable as a Ketchup bottle.


```
"Labels": [
   {
      "Name": "Ketchup",
      "Confidence": 97.69482421875,
      "Instances": [],
      "Parents": [
         {
            "Name": "Food"
         }
      ]
   },
   {
      "Name": "Food",
      "Confidence": 97.69482421875,
      "Instances": [],
      "Parents": []
   }
```


Figure 8: Return values from calling the Rekognition Service after posting the image in Figure 7

Figure 9 indicates what failed occlusion looks like, and Figure 10 gives the AWS Rekognition return types that would help indicate a defect. In Figure 9, the bottle is anchored to the back corner of the room yet is also floating and occluding the guitar that is nearest the mobile device.

Figure 9: Screenshot taken of mobile AR app

Figure 9 is a screenshot of an AR app running where the model is now occluding. The two images are sent to an ML-driven image recognition engine. When the two resultants are within an acceptable value over time, we have accelerated through the testing effort by assuring the model is being detected with confidence when being displayed in the physical world. This is one aspect of how machine learning and test automation can assist in the entire software development lifecycle of augmented reality applications. The same type of harness can be created and run autonomously while the application is in production, again assuring that within the changing environments, the models are being adequately presented in the novel user interface of AR.

However, when we utilize a screenshot of the same model file while it experiences occlusion during application runtime, we see the ML service returns concerning results. The probabilities from the same AWS recognition engine have changed dramatically. The AWS Rekognition engine no longer identifies a bottle or food. Rather we have detected Furniture, Leisure Activities, and a Guitar (illustrated in Figure 10). This would indicate potential defects and manual testing could then home in on this test case.


```
"Labels": [
   {
      "Name": "Furniture",
      "Confidence": 99.3681869506836,
      "Instances": [],
      "Parents": []
   },
   {
      "Name": "Leisure Activities",
      "Confidence": 88.02964782714844,
      "Instances": [],
      "Parents": []
   },
   {
      "Name": "Guitar",
      "Confidence": 74.00513458251953,
      "Instances": [
         {
            "BoundingBox": {
               "Width": 0.5425460338592529,
               "Height": 0.39460861682891846,
               "Left": 0.30224451422691345,
               "Top": 0.5120397210121155
            },
            "Confidence": 74.00513458251953
         }
      ],
      "Parents": [
         {
            "Name": "Leisure Activities"
         },
         {
            "Name": "Musical Instrument"
         }
      ]
   },
```


Figure 10: Return Values from Rekognition Engine for Figure 9

Conclusions 2.3

Augmented reality applications merge virtual models and the physical world. These applications are becoming more popular in many verticals, and several SDKs are available to assist in creation. However further development of testing methodologies and tooling is needed. The novel features of an AR user interface are not adequately addressed by traditional testing methods, or of the ISO 25010 model. This research suggests creating test cases that focus on characteristics of perspective, presence, interaction, immersion, and persistence. Automation and machine learning of image detection features also need to be leveraged to assist in the detection of potential defects in AR applications. The revised quality model and the ML-enabled automation framework seek to expand current capabilities and methods to enhance defect detection in AR applications.

REFERENCES

[1] Wang, X., Kotranza, A., Quarles, J., Lok, B., Allen, B.D. (2005). A pipeline for rapidly incorporating real objects into a mixed environment. *In The 4th IEEE/ACM International Symposium on Mixed and Augmented Reality*, 170-173.

[2] Rabbi, I., & Ullah, S. (2013). A survey on augmented reality challenges and tracking. *Acta Graphica: J. Print. Sci. Graphic Commun*, 24(1-2), 29-46.

[3] Gulliksen, J., Göransson, B., Boivie, I., Blomkvist, S., Persson, J., & Cajander, A. Key principles for user-centered systems design. *Behaviour and Information Technology*. 22, 397-409.

[4] Nguyen, T., (2019). Gartner emerging technology analysis: augmented and mixed reality opportunity for 3D design software and vertical ISVs. 23 August 2019.

[5] Wong, J., & Leow, A., (2019). Gartner survey analysis: insights to kick-start an enterprise multiexperience development strategy. 12 February 2019.

[6] GitHub. (2018). Google-ar/arcore-android-sdk. February 23, 2018. Accessed April 28, 2019. https://github.com/google-ar/arcore-android-sdk

[7] Bohon, C. (2019). Apple's ARKit: cheat sheet. TechRepublic, June 7, 2019. Accessed September 28, 2019. https://www.techrepublic.com/article/apples-arkit-everything-the-pros-need-to-know/

[8] Williams, R., (2018). Facebook rebrands camera effects platform, expands program to Instagram. October 8, 2018. Accessed September 28, 2019. https://www.mobilemarketer.com/news/facebook-rebrands-camera-effects-platform-expands-program-to-instagram/539066/

[9] Announcing the 1.0 release of the QCAR SDK. April 26, 2011. Accessed September 28, 2019. https://developer.vuforia.com/forum/news-and-announcements/announcing-10-release-qcar-sdk

[10] Fried, I. (2015). Qualcomm sells its vuforia augmented reality business to PTC for $65 million. October 12, 2015. Accessed September 28, 2019. https://www.vox.com/2015/10/12/11619500/qualcomm-sells-its-vuforia-augmented-reality-business-to-ptc-for-65

[11] See the Future Create the Augmented World, 2019. Accessed September 28, 2019. http://maxst.com/#/en/aboutus

[12] Release Notes Wikitude SDK, 2012. Accessed September 28, 2019. https://www.wikitude.com/external/doc/documentation/latest/android/changelog.html

[13] Arifin, Y., Sastria, T. G., & Barlian, E. (2018). User experience metric for augmented reality application: a review.” *Procedia Computer Science,* 135, 648–56.

[14] Use Vuforia for computer vision and augmented reality apps, 2019. Accessed September 28, 2019. https://www.ptc.com/en/academic-program/academic-products/free-software/vuforia

[15] Downloads, 2019. Accessed September 28, 2019. https://developers.facebook.com/docs/ar-studio/downloads/

[16] Get ready for the latest advances in augmented reality, 2019 Accessed September 28, 2019. https://developer.apple.com/augmented-reality/

[17] ARCore overview, February 28, 2019. Accessed September 28, 2019. https://developers.google.com/ar/discover/

[18] Wikitude Augmented Reality SDK, 2019. Accessed September 28, 2019. https://www.wikitude.com/products/wikitude-sdk/

[19] AR SDK, 2019. Accessed September 28, 2019. http://maxst.com/#/en/arsdk

[20] Yun, S.P., Yap, H.J., Singh, R., Chang, R., & Cheong, K.L.R. (2014). Augmented reality assisted factory layout planning and analysis for a flexible manufacturing cell. *In International Conference on Computer Science and Computational Mathematics*, 106-111.

[21] Henderson, S. J., & Feiner, S. (2009). Evaluating the benefits of augmented reality for task localization in maintenance of an armored personnel carrier turret. *In 8th IEEE International Symposium on Mixed and Augmented Reality*, 135-144.

[22] Westerfield, G., Mitrovic, A., Billinghurst, M. (2015) Intelligent augmented reality training for motherboard assembly. *Int. J. Artif. Intell. Educ.* 25(1), 157-172.

[23] Mercedes cAR app. Car parking in the living room, 2018. Accessed September 28, 2019. https://developer.mercedes-benz.com/inspire/cARapp

[24] Boas, Y. A. G. V. (2013). Overview of virtual reality technologies. *In Interactive Multimedia Conference*.

[25] Bierbaum, A., Hartling, P., Patrick, & Cruz-Neira, C. (2003). Automated testing of virtual reality application interfaces. *Eurographics Workshop on Virtual Environments ACM*, 107–114.

[26] Azuma, R.T., Baillot, Y., Behringer, R., Feiner, S. K., Julier, S., & Macintyre, B. (2001) Recent advances in augmented reality. *Computer Graphics and Applications* 21(6), 34–47.

[27] Tracy, P. (2017). Amazon's AR tool is a useful way to see products before you buy them. November 2, 2017. Accessed September 28, 2019. https://www.dailydot.com/debug/amazon-ar-view-app/

[28] Diaz, J. (2019). Avoid the dreaded IKEA fight with this new app. September 22, 2017. Accessed September 28, 2019. https://www.fastcompany.com/90143908/avoid-killing-your-relationship-at-ikea-with-this-new-app

[29] Dünser, A., & Billinghurst, M. (2011). Evaluating augmented reality systems. *In: Furht B. (eds) Handbook of Augmented Reality*. 289-307.

[30] Sutcliffe, A., & Kaur, K. (2001). Evaluating the usability of virtual reality user interfaces. *Behaviour and Information Technology*, 19.

[31] Nielsen, J. (1993). *Usability engineering*. Morgan Kaufmann.

[32] Stanney, K. M., Mollaghasemi, M., Reeves, L., Breaux, R., & Graeber, D. A. (2003). Usability engineering of virtual environments (VEs): identifying multiple criteria that drive effective VE system design. *Int. J. Hum.-Comput. Stud.*, 58, 447-481.

[33] Tao, F., Zhang, M., & Nee, A.Y.C. (2019). Digital twin driven smart manufacturing. *Elsevier Academic Press*, 229.

[34] Breen, D. E., Rose, E., & Whitaker, R. T. (1995). Interactive occlusion and collision of real and virtual objects in augmented reality. *Technical report ECRC-95-02, European Computer Industry Research Center*.

[35] Fundamental Concepts. Google Inc. ARC Core, February 28, 2019. Accessed April 28, 2019. https://developers.google.com/ar/discover/concepts#anchors_and_trackables.

[36] Asadi, A. N., Azgomia, M. A., & Entezari-Malekib, R. (2019). Unified power and performance analysis of cloud computing infrastructure using stochastic reward nets. *Computer Communications*, 138, 67-80.

[37] Persistence in AR - Why Is It so Important? AR Critic. June 16, 2018. Accessed May 01, 2019. https://arcritic.com/2151/persistence-in-ar-why-is-it-so-important/

[38] York, J., & Pendharkar, P. C. (2004). Human–computer interaction issues for mobile computing in a variable work context. *International Journal of Human- Computer Studies*, 60, 771-797.

[39] Albert, W., & Tullis, T. (2013). Measuring the user experience: collecting, analyzing, and presenting usability metrics. *Newnes*; 2013 May 23.

[40] Hassenzahl, M. & Tractinsky, N. (2006). User experience - a research agenda. *Behaviour & Information Technology*, 25, 91-97.

[41] Kaasila, J. (2018). Google augmented reality - how bitbar helped automate their app testing. Bitbar, April 17, 2018. Accessed September 28, 2019. https://bitbar.com/blog/how-bitbar-helped-google-to-automate-their-augmented-reality-testing/

CHAPTER 3: CONTEXTUALIZATION OF THE AUGMENTED REALITY QUALITY MODEL THROUGH SOCIAL MEDIA ANALYTICS

Augmented Reality applications are gaining popularity while maintaining novelty. Many industries are utilizing the user interface type, and use cases are becoming repeat patterns of problem solutions. Despite this rising popularity, quality has not matured nor has the technology become mainstream. Novelty must be approached as risk, and risk must be evaluated for and tested to assure adequate levels of quality. Quality itself can also be vague and have a contextual definition. For these reasons, a quality model for augmented reality was created. This work analyzes over two hundred thousand tweets, collected during 2019 and 2020, relating to augmented reality technology, and contextualizes various data points to the established AR Quality Model. The education industry had the highest mentions among the tweets within the scope of this research while the tweets labeled to the transportation industry had the highest sentiment. Furthermore, the tweets were shown to illustrate the needs of testing against the characteristics within the quality model: presence, perspective, interaction, portability, and persistence.

Introduction 3.1

This chapter uses analysis from social media posts to extend the previous chapter and the Augmented Reality (AR) quality model [1]. Augmented reality was first coined as a term by Caudell and Mizell in 1990 [2]. Augmented reality is accomplished through an application's user interface by merging physical and digital worlds [3].

Research of AR applications covers many different use cases and industries. One study has uncovered findings such as children seem to understand how to use the augmented reality application as soon as they began to interact with the technology. Another research effort found learning English utilizing AR can help students in learning material and motivated students to learn the language. Another study proposes AR mobile applications in the household to educate parents and caregivers of potentially dangerous objects in the home that may endanger infants. It is easy to see how innovative the space is becoming, and how useful augmented reality applications are for humans. Gartner, Inc. has predicted that “by 2021, at least one-third of enterprises will have deployed a multiexperience development platform (MXDP) to support mobile, web, conversational and augmented reality development” [4].

The Augmented Reality Application Quality Model 3.1.1

The importance of this technology extends into many industries. AR has been utilized in applications for factory design [5], military equipment maintenance [6], and for teaching electronics assembly skills [7]. The effectiveness of the applications requires them to have levels of quality. The ISO 25010 model identifies eight characteristics of quality software: Functional, Performance, Compatibility, Usability, Reliability, Security, Maintainability, and Portability [8]. However, the expectations of AR application users are not well known [9]. Because the AR interface is a novel one, traditional software testing technologies and practices are not up to the task [10]. For these reasons, a quality model was established that encapsulates the characteristics for AR applications broadly as presence, perspective, interaction, immersion, and persistence [1].

Beyond the characteristics of quality is the necessity to utilize context in testing. Emerging technologies and designs such as digital twins must still begin with system needs analysis through known methods such as context diagrams and objective trees [11]. And the same is true for established technologies. According to Gartner, “your strategy for API quality needs to be built with the business and environment context in mind” [12]. These contextual concerns drive much of this extended research. Utilizing over 200,000 tweets towards AR and VR, this research contextualizes quality regarding AR and various industries using text analysis with content-based, time series, and sentiment analysis.

Methodology 3.2

Collection 3.2.1

Beyond reviewing and incorporating popular and academic papers, this chapter's research methodology focuses on social media analytics. Data acquisition is completed through a small program written in R and executed daily beginning in August of 2019. The program uses the twitteR library as an API to Twitter's social media platform. The program populates a MySQL database hosted in AWS RDS with Twitter data when the tweet's message is relevant to augmented reality topics. Figure 11 omits various keys and passwords yet implements the basics of importing the twitteR library and establishing connections to both Twitter and the MySQL database before searching on '#AR' and storing the search results.

```
1  library(twitteR)
2  library(RCurl)
3  library(ggplot2)
4  library(reshape)
5  library(sentimentr)
6  library(RMySQL)
7
8  consumer_key <- ''
9  consumer_secret<- ''
10 access_token <- ''
11 access_secret <- ''
12
13
14 register_mysql_backend("", "", "", "")
15 setup_twitter_oauth(consumer_key, consumer_secret, access_token, access_secret)
16
17 dbartweets <- search_twitter_and_store('#AR', retryOnRateLimit = 20, lang = "en")
```

Figure 11: Screenshot of example usage of the twitteR library within R Studio

Many hashtags and Twitter profiles names were utilized in the searches. The next table identifies some of the additional search items, extending the code found on line 17 of Figure 11.

Table 4: Search Terms used in the twitteR API Calls

Hashtags	Accounts
'#AR'	'@AR_Maxst'
'#VR'	'@wikitude'
'#augmentedreality'	'@Vuforia'
'# omnichannel'	'@fbplatform'
'#multiexperience'	@GoogleARCore

Hashtags	Accounts
'#virtualreality'	@ArcoreGoogle
'#arkit'	'@AR_Maxst'
'#arkitnews'	
'#sparkar'	

The tweets are stored with other certain meta data, made available through the twitteR library and API. Beyond the tweet's text, Twitter data collected and utilized in this research include discrete data fields towards the number of times a tweet has been marked as a favorite by a Twitter user, as well as the number of times a tweet has been re-tweeted. A date field within the table indicates the date the tweet was created on the Twitter platform.

Data Preparation and Processing 3.2.2

Social media analytics is not without data quality and research concerns. Tweets lack standards, such as in hashtag usage. A tweet that embeds #AR may be about Arkansas, a state within the United States, and not about augmented reality technology. Before formulating findings and information from the collected data, much data preparation and cleaning are required. Data preparation activities may account for eighty percent of the time invested in data science efforts [16]. The amount of data dropped during preparation from the data set can be large. This is to be expected, as with the example of #AR as Arkansas, and not augmented reality. The search algorithm sought out technology vendors, whose tweets may have been towards other products, processes, marketing, or communication needs. Tweets are fixed in their length of 280 characters [17]. A short message may come from a targeted technology provider, but also fail to cite key search criteria words. Within this research set, counting only the tweets that mention specifically the words "augmented" or "virtual" (not case sensitive), the original 637,384 tweets were reduced by two-thirds to 211,269 tweets. The breakdown of augmented and virtual tweet counts are discussed further.

Data preparation and cleaning convert raw data into usable information [16]. In this research, we create useful subsets without disrupting the raw database contents, as new hypotheses and meaning may yet be executed and found across the raw data.

To create useful datasets, we must also create new fields, such as categorical values. Tweets have some indication to business and industry, so categorizing the tweets by industry lends toward understanding a contextual need for quality, such as HIPAA requirements and testing concerns towards health or pharmaceutical implementations of AR. A small program in R reads the tweet texts from the accumulated corpus and maps to common industries, counting the industry mentions, including if multiple industries are mentioned in a single tweet.

Results 3.3

The first finding is that among all tweets that mention either “augmented” or “virtual”, augmented reality mentions (135,319) are more frequent than virtual reality mentions (80,540), illustrated in Figure 12.

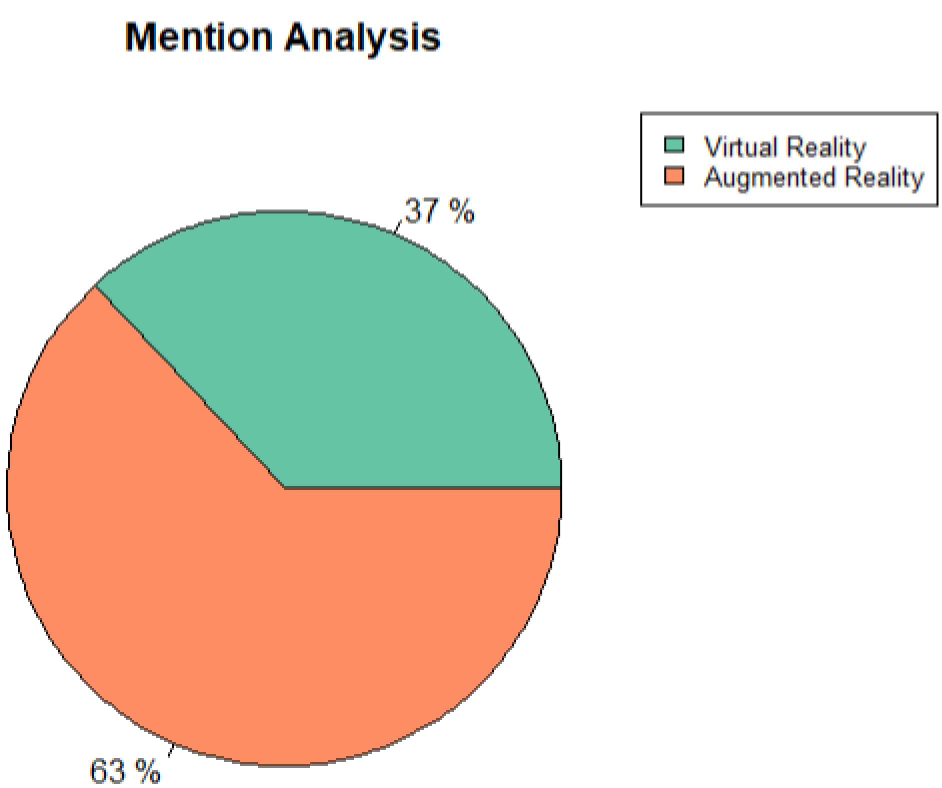

Figure 12: Percentage of "Augmented" versus "Virtual" mention in tweets

This first finding seems to contradict a 2018 survey by Gartner, where survey recipients responded towards virtual-reality apps (20%) more than augmented-reality apps (14%), in terms of apps that will have the most impact on business success by year-end 2020. Although, the Gartner report analysts stated, “It is, however, surprising to see VR apps identified as the second most impactful type of multiexperience app (20%), as AR has more potential use cases and device support. But only 14% of the respondents thought that AR apps would be the most impactful, despite AR app development tools being more widely available. “[18]. Figure 13 shows where virtual-reality apps were more frequent in survey recipient response compared to augmented-reality apps for business success.

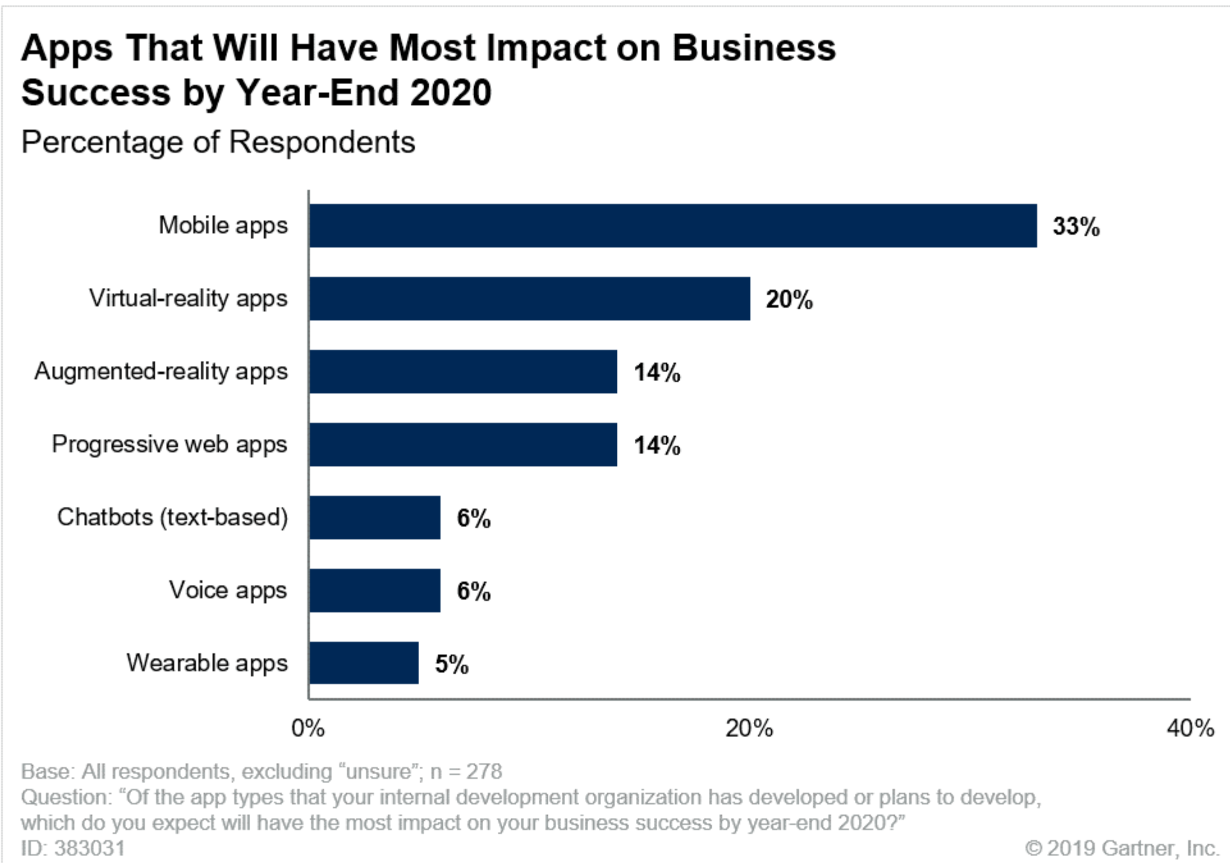

Figure 13: Gartner: Apps expected to have most impact on business success by year-end 2020

Most Favored and Most Retweeted Tweets about Augmented Reality 3.3.1

Tweets can be liked, or favorited, as well as re-tweeted. These mechanisms, favorite and retweet, are utilized to propagate, share, show appreciation of another user posting, and the number of favorites indicates the count of unique user accounts that like or agree with the content of the tweet [19]. Tweets frequently have meaning, and in this research, the gathered tweets are towards specific technologies.

Within the scope of this research, the most favorite tweet is towards AR hardware, wearable glasses, and contains an embedded video with a link to a Kickstarter site to raise investment through crowdfunding (tweet image found in Figure 14). The glasses, by TiltFive, are marketed to be the future of tabletop games and a new take on traditional board games.

Figure 14: Most favored tweet within the research scope

The digital boardgame approach gives the characters dimension, allows sharing the digital and augmented experience with other players, as well the ability to save and persist the immersive experience. These AR application quality characteristics are mentioned and defined within the quality model as presence, perspective, and persistence. Section four provides more examples of how the characteristics within the AR quality model are showcased in popular AR tweets.

The following tweet, posted in 2017, had at the time of this research been viewed over 212,100 times and was re-tweeted 4,247 times. The tweet embeds a demonstration video to augment the physical environment and makes this statement, “This is ARKit’s A-ha moment”. ARKit is a development technology for building AR applications. Apple released ARKit with iOS 11 in 2017 [20]. More development tool analysis is found in section 4.5. Within the scope of this research, Figure 15 illustrates the most re-tweeted tweet.

Figure 15: Most re-tweeted tweet within the scope of this research

Content-based, Sentiment and Time Series Analysis 3.3.2

Content-based analysis can identify the frequency of theme and text occurrence. Research shows augmented reality applications are being utilized across many industries; content-based analysis from social media provides an additional perspective. Education, Entertainment, and Commerce are the top three mentioned industries. This is determined by first creating a list of known industries from a source [21]. Then the algorithm reads the tweet texts and will count industry mention, including if multiple industries are mentioned in a single tweet (illustrated in Figure 16).

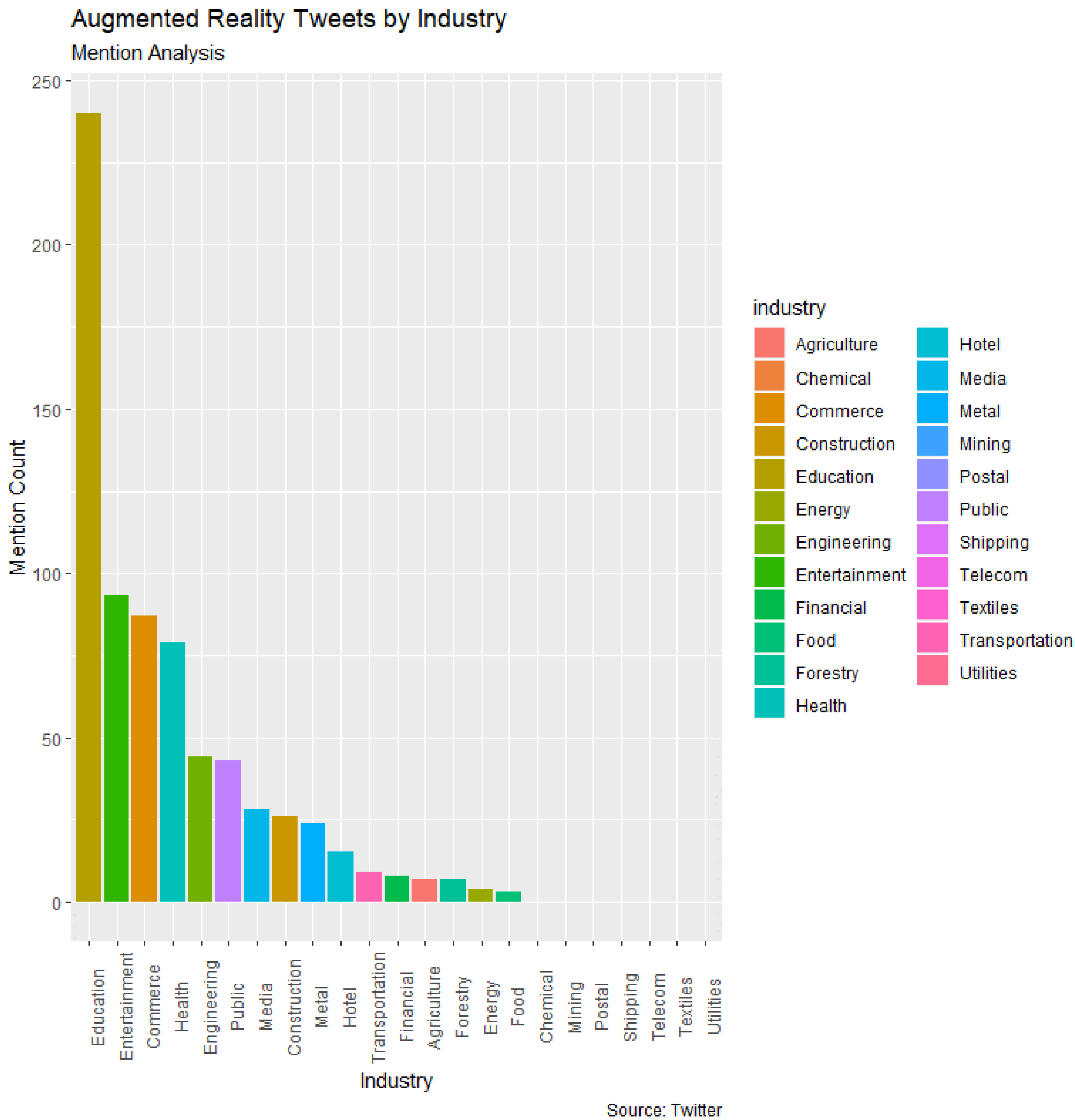

Figure 16: Industry mention analysis: education, entertainment, and commerce are the top three mentioned industries

Further evaluation of the top mentioned industry, education, is provided utilizing a time series chart. Figure 17 indicates a slightly positive trend in the smoothed line and a peak of tweets in November of 2019.

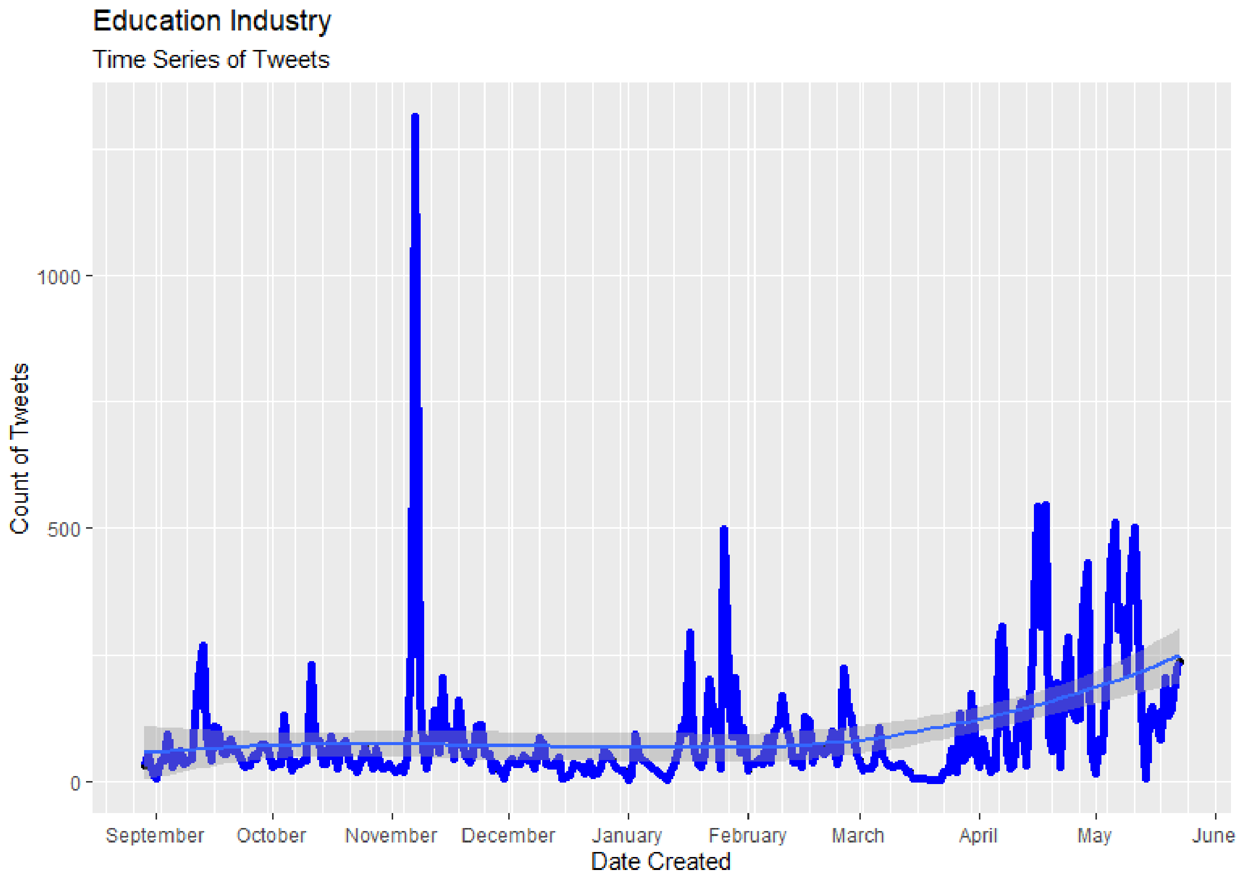

Figure 17: Time series chart of tweets towards Education industry

The following tweet was retweeted over 1,000 times and was the largest contributing factor to the November peak of tweets having mention of the education industry. Figure 18 is a tweet containing an embedded AR demo video of a skeletal dinosaur walking on a physical plane as an educational and awareness instrument.

Figure 18: Highly favored tweet was a large contribution to the November 2019 peak of education industry tweets

Tweets are short in character length but offer a volume of opinion. This volume of data can be mined as user-driven data and classified in terms of sentiment [22]. Opinion mining and sentiment analysis techniques can be utilized to support or drive product and organizational communication plans [23]. Sentiment analysis has also been combined with other analytic methods to produce stock market prediction models with accuracy greater than 60% [24]. Tweets are opinionated text. Opinionated text is the primary resource for shoppers when making purchases [25]. We can utilize user-driven data for measuring system quality. Indeed, the quality of the system is also important in procurement, investment, and communication planning.

The industry that has the highest average sentiment score within the scope of this research is Transportation. Figure 19 illustrates average sentiment scores by the aggregated tweets within industry categories. A tweet that references multiple industries would affect the sentiment score towards each referenced industry.

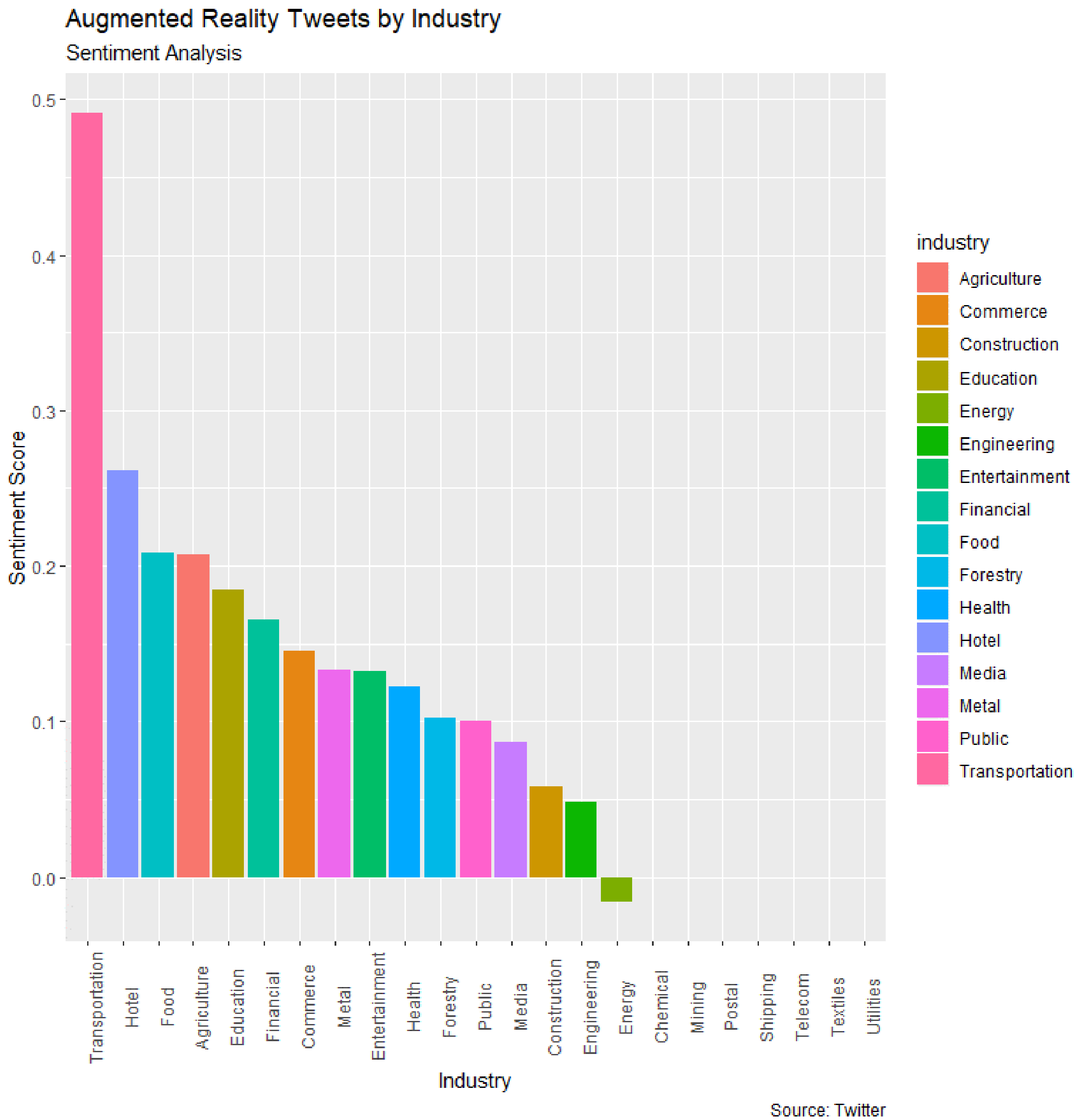

Figure 19: Sentiment scores by aggregated tweet sentiment and grouped by industry

Further analysis of Transportation industry tweets indicates a spike in March of 2020. Figure 20 is a chart indicating the slope of the time series data as well as the spike.

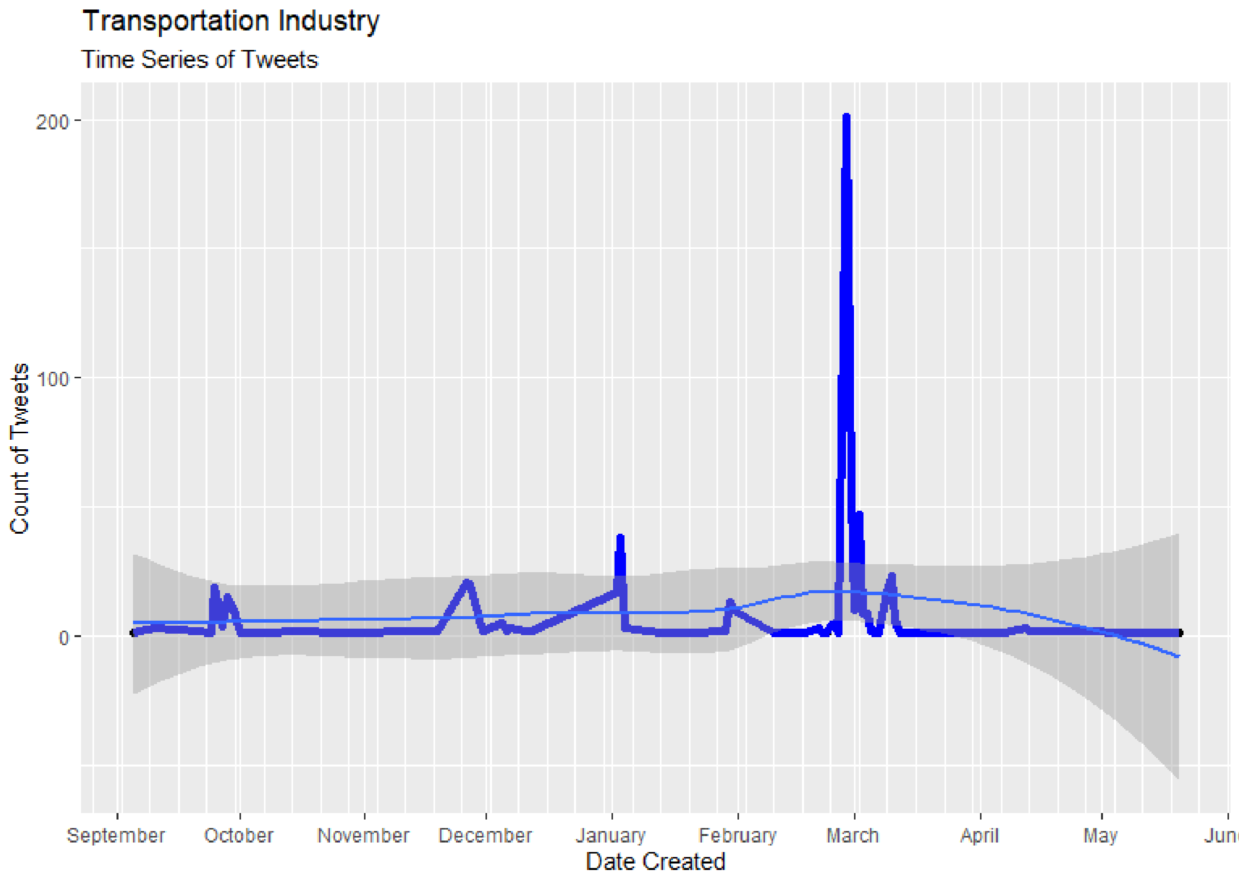

Figure 20: Time series chart of tweets towards the transportation industry

The most retweeted tweet during the peak period of transportation-related tweets references the value of AR for business logistics. The tweet is illustrated in Figure 21.

Figure 21: Highly leveraged tweet and main contributor to the March spike of transportation industry tweets

The Sentimentr library is utilized within the R code of this research to provide sentiment scores of the collected tweets. The sentimentr library utilizes valence shifters, words that lead and may reverse or alter the sentiment of following terms, to increase the libraries accuracy [26]. The highest sentiment-scored tweet, a score of 1.540922, is found in Figure 22 (sentimentr library scores range between -2, 2). The tweet references an easier and more fun life due to augmented reality technology and utilizes the hashtag #CX (customer experience). While application users may be customers, they may also be internal employees. This research refers to user experience as a more generalized approach.

Figure 22: Tweet within research scope having highest sentiment score

Application of Analysis to the Established AR Quality Model 3.4

Augmented reality applications are novel and require levels of quality to be established to be effective for the user experience. Establishing the quality of an AR application goes beyond traditional testing practices and standards. A new model has been established for test case focus and categorization, to build system confidence and inform system readiness and release decisions. These are critical decisions when dealing with novel technology.

Gartner's 2020 Digital Friction Survey found that among a survey n size of 4,582 employees surveyed, only 31% experienced High-Quality UX that was productive, empowering, and easy [27]. Why might the system need to have this level of quality? The same survey found that "employees with a high-quality UX are 1.8 times more likely to have a high intent to stay as indicated by a lack of interest in leaving their current employer or actively search for a new job." The Gartner research report further explained, "Employees with a high-quality UX are 1.5 times more likely to have high discretionary effort, which includes employees' willingness to help colleagues, take on additional responsibilities, put in extra effort and find better ways of working."

Augmented reality applications are a novel approach to a system's user experience. The user experience has an opportunity to benefit users and organizations alike. The technology implementation requires quality for these benefits to be realized.

Contextual Approaches to Quality 3.4.1

The characteristics of the quality model are quickly identified in the following bulleted text. Further sections then develop the characteristics and apply the findings of the social media analysis to the previously established quality model.

- Presence
- Perspective
- Interaction
- Portability
- Persistence

Twitter Analysis Applied to the Characteristic of Presence 3.4.2

Accurately executed occlusion and collision behavior will make the AR application more realistic [1]. Together they form the characteristic of presence. Presence allows AR application users to experience both the physical and digital worlds accurately. Presence is exemplified when digital models have correct shadowing and obey gravity such as physical objects would. Presence must be tested for. Figure 18, a digital model of a dinosaur overlaid onto a physical landscape (the original tweet is a video of the skeletal dinosaur walking) maintains the existence of shadow beneath the digital model, an example of presence in AR.

According to Gartner, "realism, as well as usability, will be enhanced or undermined by the digital model's ability to have accurate presence within a physical landscape. There should therefore be a test for occlusion which occurs when a far object hides a near object, for example, and which also sometimes occurs when objects are floating above and closer than an original anchor point. Test for occlusion by placing multiple models or scenes and moving throughout the physical environment." [28]. Collision in AR applications exists when a digital object and a physical object appear to occupy the same physical place at the same time. Of the 211,269 tweets related to augmented and virtual reality in this research, only 70 tweets reference occlusion or collision. The average sentiment for an AR tweet is .1176; however, for tweets mentioning collision the sentiment turns negative at -0.0157, and ventures more negative when mentioning occlusion -0.0475.

Twitter Analysis Applied to the Characteristic of Perspective 3.4.3

Perspective frequently refers to someone's point of view. In augmented reality, perspective must encompass the physical and digital objects, combined behavior, and how well visual distinction exists among them.

Perspective may be affected by the physical environment, the digital model's attributes, and the interaction between them. This would include the inability to scale a digital model to the correct size when embedded into the physical view. Perspective may fail when rendering white text or digital models within a physical environment that is brightly lit or has a predominately white background. The perspective within AR applications should provide visual distinction and improve the quality of the user experience. Figure 22, the tweet within the scope of this research having the highest sentiment score, adequately illustrates the characteristic of perspective. This example uses a neon green annotation text color over a mostly grayscale physical scene. The perspective enables quickly drawing attention to a physical device through digital annotation.

Twitter Analysis Applied to the Characteristic of Interaction 3.4.4

The characteristic of interaction is directly related to performance. Interaction directly affects the device, the costs, and the execution environment. The tweet text found in Figure 21 refers to AR applications used to model warehouses which may assist in finding specific freight.

Warehouses can be quite large, requiring longer application runtime, as well as the rendering of many possible digital annotations, such as directions to a warehouse location, before arriving at the correct position. If the AR application is utilizing commercial-grade devices, we must concern about how the interaction of such an application will consume the WIFI network, the battery of the device, and whether other applications will run on the device at the same time. The greater the interaction, the more important the performance.

Twitter Analysis Applied to the Characteristic of Portability 3.4.5

There are many software development kits available to build AR applications. Table 5 references many examples of these technologies, and the runtime platforms they support. Not every AR application will run on your commercial or consumer system. It is not uncommon for some popular mobile AR apps to exhibited very limited portability and only run on a single mobile operating system [29]. The value and availability of your application will depend upon the portability characteristic, which is a quality concern that must be considered and established during design and development.

Table 5 Augmented reality development tools and supported runtime platform/s

Tool name	**Supported Runtime Platform/s**
Vuforia Studio [30]	Android, iOS, UWP, Unity
Spark AR [31]	Android, iOS
ARKit [32]	iOS
ARCore [33]	Android, iOS, Unity, Unreal
Wikitude [34]	Android, iOS, Microsoft Tablet
Maxst AR [35]	Android, iOS, Unity

Table 4 provides the search terms that were utilized to scope the collection of tweets for this research. Many of those search terms are Twitter handles or the names of SDK technologies. The top three mentioned development vendors are Apple (966), Unity (713), and Google (601). These figures are illustrated in Figure 23. The popularity found in the content-based analysis is not conducive to portability, as Apple is one of the more limited SDKs when concerned with runtime environment support as indicated within Table 5.

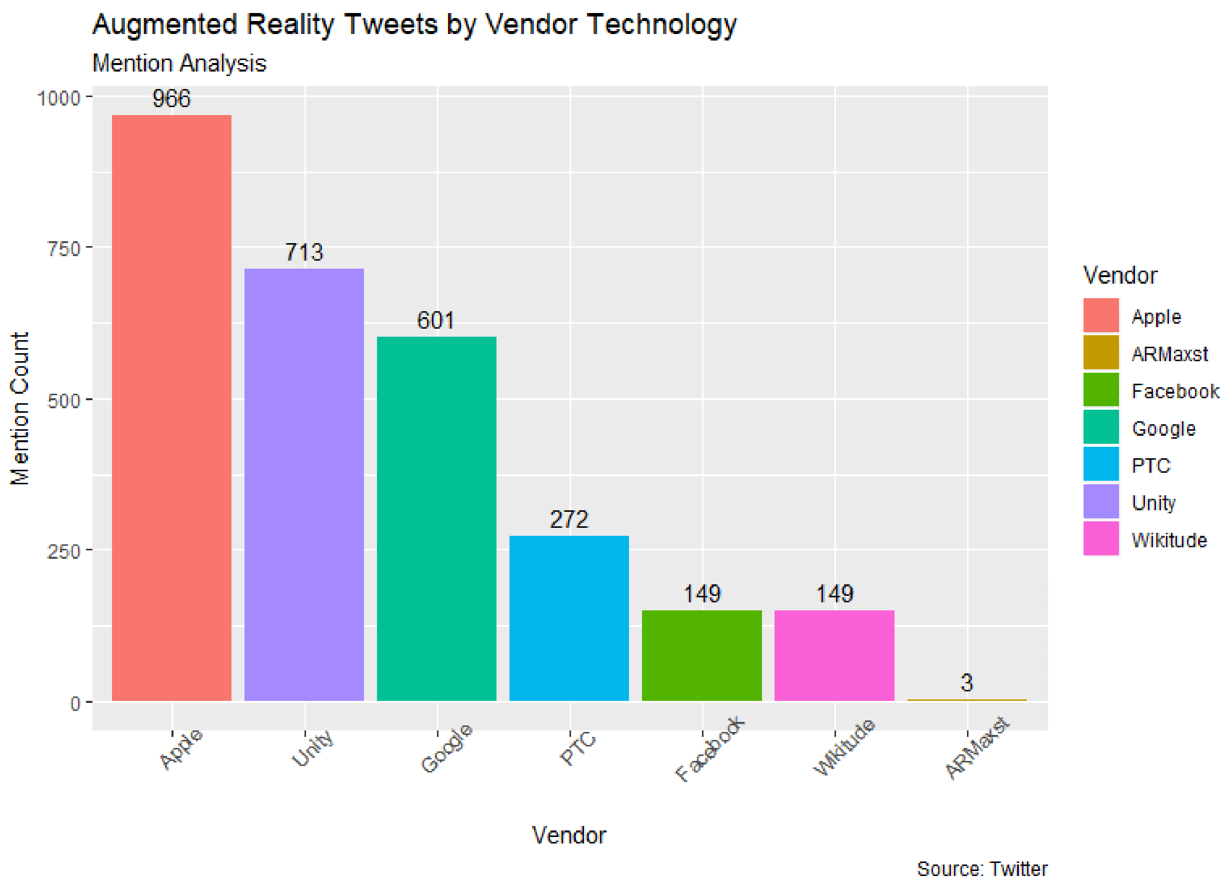

Figure 23: Example augmented reality development technology ordered by mention count

Twitter Analysis Applied to the Characteristic of Persistence 4.4.6

Scaling augmented reality experiences across multiple users requires centralizing physical maps and virtual object anchors so that multiple users can share immersion [1]. Persistence allows an AR user experience to be shared among other application users from separate devices or application instances during runtime. Additionally, persistence allows the AR experience to be saved and resumed later by the original, or potentially other, application users.

Figure 14 is the most favored tweet within the scope of this research, and it adequately exemplifies many of the characteristics of this quality model. The tweet references the future of tabletop games, a physical board game with digital models rendered above. The persistence of the digital models allows multiplayer action in the game, such as competing players may experience the avatars that are being controlled by other's play. The ability of the application to run on and that users may leverage whichever phone device or headset they have, would encompass the portability characteristic. If the game is to be played for many hours, the interaction level will be high, perhaps enough to warrant plugging an AR hardware device into a power source. Perspective is required so that the digital scenes are relevant to the location, size, and position of the physical game board as seen from other's points of view. The presence of the digital models, the prevention of collision, will also provide an experience worth dedicating many hours towards.

While the previously mentioned and exemplified tweet represents the entertainment industry, certainly implementation across healthcare, manufacturing, real estate, education, and many other industries would prioritize quality potential during investment in novel technology. Hence the AR quality model, which provides a basis for test case creation and categorization that promotes confidence in the application's readiness and informs decisions towards the application's release.

Conclusions and Future Works 3.5

Conclusions 3.5.1

The findings reported in this research study represent essential considerations for the investment and realization of value through augmented reality applications. Augmented reality applications have characteristics that are beyond traditional interfaces. Despite a novel technology, quality still matters. For these reasons, the AR quality model has been established and extended with input from social media and in some cases exemplified through such uses.

Among all tweets that mention either "augmented" or "virtual", augmented reality mentions (135,319) are much more frequent than mentions of virtual reality (80,540). The top mentioned industries within the tweet collection scope were Education, Entertainment, and Commerce. The top average tweet sentiment scores by mentioned industries within the study scope were Transportation, Hotel, and Food.

AR quality model characteristics of Presence, Perspective, Interaction, Persistence, and Portability all had representation among some of the highest retweeted or liked tweets, or among the tweets that drove industry spikes in the time series charts. Still, limitations exist. Only one social media platform was utilized in this research, Twitter. Facebook, another such social media platform, is also a technology provider of AR software development. However, that platform was not utilized in this research. The AR quality model requires further research beyond social media and into specific teams and implementations within the industry. Research towards failed AR implementations would also support the model or indicate where omissions of key characteristics exist.

Future Works 3.5.2

Among the quality model and where it must extend, the quality characteristic of security. 475 of the tweets within this research mentioned terms towards privacy or security. That is less than one-quarter of one percent of tweets within the scope of this research. While the percentage is small, it is critical to consider that augmented reality applications require access to the device camera. The device camera may capture personal or organizational information, potentially confidential or private, that may become vulnerable to threat actors if exploited. Future work should engage in various threat and vulnerability research of AR applications and methods and technology which may help mitigate such risk or exposure. Future works should also include case studies from many of the named industries within the paper with indication towards whether the AR quality model adequately supports the user's experience.

REFERENCES

[1] Scheibmeir, J., & Malaiya, Y. K. (2019). Quality model for testing augmented reality applications. *In IEEE 10th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON),* 0219-0226, https://doi:10.1109/UEMCON47517.2019.8992974.

[2] Rabbi, I., Ullah, S. (2013). A survey on augmented reality challenges and tracking. *Acta Graphica: J. Print. Sci. Graphic Commun.* 24(1-2), 29-46.

[3] Wang, X., Kotranza, A., Quarles, J., Lok, B., & Allen, B.D. (2005). A pipeline for rapidly incorporating real objects into a mixed environment. *In The 4th IEEE/ACM International Symposium on Mixed and Augmented Reality*, 170-173.

[4] Iquira-Becerra, D., Flores-Conislla, M., Carlos-Chullo, J.D., Sotelo-Castro, B. Payalich-Quispe, C. & Corrales-Delgado, C. (2020). A critical analysis of usability and learning methods on an augmented reality application for zoology education. *Advances in Science, Technology and Engineering Systems Journal*, 5(2), 384-392. https:// DOI:10.25046/aj050250

[5] Syarifudin, Z., Suharjito. (2020). Mobile based for basic english learning assessment with augmented reality. *Advances in Science, Technology and Engineering Systems Journal*, 5(2), 774-780. https://doi:10.25046/aj050297

[6] Nishizaki, M. (2018). Visualizing affordances of everyday objects using mobile augmented reality to promote safer and more flexible home environments for infants, *Advances in Science, Technology and Engineering Systems Journal*, 3(6), 74-82. https:// doi:10.25046/aj030607

[7] Cearly, D. et al. (2020). Top 10 strategic technology trends for 2020: multiexperience. Gartner, 10 Mar. 2020, www.gartner.com/doc/3981950.

[8] Yun, S.P., Yap, H.J., Singh, R., Chang, S.W., & Cheong, K.L.R. (2014). Augmented reality assisted factory layout planning and analysis for a flexible manufacturing cell. *In International Conference on Computer Science and Computational Mathematics*, 106-111.

[9] Henderson, S. J., Feiner, S. (2009). Evaluating the benefits of augmented reality for task localization in maintenance of an armored personnel carrier turret. *In 8th IEEE International Symposium on Mixed and Augmented Reality*, 135-144.

[10] Westerfield, G., Mitrovic, A., Billinghurst, M. (2015). Intelligent augmented reality training for motherboard assembly, Int. J. Artif. Intell. Educ., 25(1), 157-172.

[11] ISO/IEC 25010. ISO 25000. Accessed May 22, 2020. https://iso25000.com/index.php/en/iso-25000-standards/iso-25010.

[12] Dünser, A., & Billinghurst, M. (2011), Evaluating augmented reality systems. *Handbook of Augmented Reality. Springer*, 289-307.

[13] Sutcliffe, A. G. & Deol Kaur, K. (2000). Evaluating the usability of virtual reality user interfaces, *Behaviour & Information Technology*, 19(6), 415-426, https://doi:10.1080/014492900750052679.

[14] Scheibmeir, J., & Malaiya, Y. (2019). An API development model for digital twins. *In IEEE 19th International Conference on Software Quality, Reliability and Security Companion (QRS-C)*, 518-519, https://doi:10.1109/QRS-C.2019.00103.

[15] Scheibmeir, J., Murphy, T., & Herschmann, J. (2019). Promote continuous quality for APIs to support digital business. Gartner, February 21, 2019. https://www.gartner.com/doc/3902567.

[16] Cirillo, A. (2016). *RStudio for R Statistical Computing Cookbook*. Birmingham, Packt Publishing, 2016.

[17] "How to Tweet." Twitter. Accessed May 23, 2020. https://help.twitter.com/en/using-twitter/how-to-tweet.

[18] Wong, J., Leow, A. (2019). Gartner survey analysis: insights to kick-start an enterprise multiexperience development strategy, 12 February 2019. https://www.gartner.com/document/3901479

[19] Perdana, R.S., & Pinandito, A. (2018). Combining likes-retweet analysis and naive bayes classifier within twitter for sentiment analysis. *Journal of Telecommunication, Electronic and Computer Engineering*, 10(1-8), 41-46.

[20] Apple's ARKit: Cheat sheet, Cory Bohon, TechRepublic, June 7, 2019. Accessed September 28, 2019. https://www.techrepublic.com/article/apples-arkit-everything-the-pros-need-to-know/

[21] International Labor Organization. "Industries and Sectors." Industries and sectors, n.d. https://www.ilo.org/global/industries-and-sectors/lang--en/index.htm.

[22] Ghiassi, M., Skinner, J., Zimbra, D. (2013). Twitter brand sentiment analysis: a hybrid system using n-gram analysis and dynamic artificial neural network. *Expert Systems with Applications*, 40(16), 6266-6282. https://doi:10.1016/j.eswa.2013.05.057

[23] Purohit, H., Hampton, A. Shalin, V.L., Sheth, A.P., Flach, J., & Bhatt, S. (2013). What kind of #conversation is twitter? mining #psycholinguistic cues for emergency coordination. *Computers in Human Behavior*, 29(6), 2438-2447. https://doi:10.1016/j.chb.2013.05.007

[24] Tanulia, Y., & Girsang, A.S. (2019). Sentiment analysis on twitter for predicting stock exchange movement. *Advances in Science, Technology and Engineering Systems Journal*, 4(3), 244-250.

[25] Hung, C. (2017). Word of mouth quality classification based on contextual sentiment lexicons. *Information Processing & Management*, 53(4), 751–63.

[26] Rinker, T. "Sentimentr." README, n.d. https://cran.r-project.org/web/packages/sentimentr/readme/README.html

[27] Application Research Team. "Proven Design Principles to Deliver a High-Value Employee User Experience." Gartner, May 12, 2020. https://www.gartner.com/doc/3985028.

[28] Scheibmeir, J., Herschmann, J., & Murphy, T. (2020). Quality is the key to avoiding digital distortion with your augmented reality strategy. Gartner, April 16, 2020. https://www.gartner.com/doc/3983581.

[29] Jansen, M. (2019). “Escape Reality with the Best Augmented Reality Apps for Android and IOS.” Digital Trends. Digital Trends, June 9, 2019. https://www.digitaltrends.com/mobile/best-augmented-reality-apps/.

[30] Use Vuforia for computer vision and augmented reality apps, 2019, https://www.ptc.com/en/products/augmented-reality/vuforia

[31] Downloads, 2019, https://developers.facebook.com/docs/arstudio/downloads/

[32] Get ready for the latest advances in augmented reality, 2019, https://developer.apple.com/augmented-reality/

[33] ARCore overview, February 28, 2019, https://developers.google.com/ar/discover/

[34] Wikitude Augmented Reality SDK, 2019, https://www.wikitude.com/products/wikitude-sdk/

[35] AR SDK, 2019, Maxst. “Setup Guide.” 3.4.x | MAXST Developer Site, n.d. https://developer.maxst.com/MD/doc/unity/setup.

CHAPTER 4: MULTI-MODEL SECURITY AND SOCIAL MEDIA ANALYTICS OF THE DIGITAL TWIN

Digital twins act through application programming interfaces to their physical counterparts to monitor, model, and control them. Beyond these traditional functions of digital twins, they must also act to secure their physical counterparts. A multi-model scheme is presented to help digital twins towards the task of securing the physical system. Additionally, this work includes an analysis of more than four hundred thousand tweets each relating to digital twin technology and cybersecurity which were collected during June and July in 2020. Of the first corpus of tweets collected by searching for #digitaltwin during the research period, only a small population of 10% reference security concepts. In the second and larger corpus of collected tweets, the top mentioned industries or sectors were health, education, and public. A naïve Bayes model reached a 70.3% accuracy at differentiating tweets that were either related to cybersecurity or the internet of things. The study also indicates that cybersecurity tweets are consistently more negative in many areas of sentiment when compared to tweets about the internet of things. The sentiment findings of cybersecurity tweets will reinforce the need to address culture in cybersecurity posture while the security multi-model schema contributes to the state of the art.

Introduction 4.1

An API development model for digital twins has been proposed [1]. This chapter augments that work to address the cybersecurity concerns of the digital twin within the context of the various implementations and industries where the twin may be utilized. Furthermore, the adoption of current security models, such as state machine, lattice, non-interference, and information flow models are proposed for digital twins.

In chapter one, a development model for the APIs of the digital twin was introduced. The model was proposed to establish the context of the environment, the system, relationships between the two, and the functional and non-functional requirements. This chapter extends chapter one’s concepts by including the mining of Twitter tweets using appropriate search criteria for digital twins, the internet of things (IoT), and cybersecurity. The analysis includes content-based analysis of industries, exploration of the sentiment, natural language processing (NLP) towards the classification of tweets, and modeling to predict the twitter user behavior.

Three sets of tweets were collected for this study with the latter two sets being joined into a combined corpus. Each set was collected through the execution of scripts written in R. The first collection is smaller and has 3,102 tweets about digital twins. The second and third collections are larger. In total, 422,963 industrial internet of things (IIoT) tweets and 497,174 cybersecurity tweets were collected. These larger sets were combined into one corpus for analysis. The search criteria for the tweet collections are listed in Table 6.

Table 6: Tweet collection search criteria

Script Topic	**Criteria used in actual R code**
Digital Twin	#digitaltwin
IIoT	#industry40 or #IIOT or #IOT
Cybersecurity	#cybersecurity or #infosec or #hacking

This study is organized as follows. Section 2 reviews the literature and the background information on the subject. Section 3 describes our research method, including the collection of tweets as well as the processing. Sections 4 and 5 present the results from social media analysis. Section 6 proposes a multi-model security scheme for digital twins. Finally, section 7 presents the conclusion of this study.

Background and related work 4.2

Digital twins are virtual representations of their physical counterparts that aid in monitoring and controlling functions [2]. The design of the digital twin and IoT interfaces may utilize structured standards such as OpenAPI to list the various objects, paths, and operations [3]. Digital twins can be used during system design or during system operations to assist with information security [4]. Digital twins may be informed by Internet of Things (IoT) devices and IoT systems of systems [5].

Wireless Sensor Networks (WSN) and attached IoT sensor components can be effective monitoring approaches [6]. WSNs are frequently resource constrained and thus more vulnerable to security threats [7]. To maintain the integrity of the sensor data and overall system state, authorization and authentication are required [8].

Vulnerable digital twins offer hackers a blueprint of the physical counterpart and its integrations [9]. The digital twin can help secure the physical systems or be yet another vulnerability as IoT growth has increased the cyber-attack surface [10]. IoT systems may exhibit security vulnerabilities such as authorization, authentication, privacy, and control loss [11]. In a 2019 study, of the 220 security leaders in industrial and manufacturing who took part in the research survey, 79% indicated they had experienced an IoT cyberattack within that past year [9]. While digital twins may present such risks, they are still getting popular. The smart factory market, an economy of integrated and automated manufacturing components, is predicted to be valued at approximately $205 billion by 2022 [9]. IoT and digital twins-based systems can be attractive targets for both valid and malicious actors.

Hearn and Rix identify advantages, such as the prevention of downtime and monitoring attacks against cyber-physical systems, as potential benefits of digital twins [9]. Risks of digital twins in cyber-physical systems (CPS) include the intellectual property incorporated in the digital twins as well as critical information about the CPS itself. For such reasons, the digital twin itself must go through software hardening routines, such as fuzz and penetration testing [9]. Security of the digital twin must start from ideation and continue through its lifecycle into retirement. Security concerns must be incorporated with the organization’s culture [9].

IoT systems and CPS are being deployed in the energy industry, and there have been instances of cyber-attacks on them. A few well-known cyber-physical system attacks include the stuxnet attack on an Iranian uranium plant [10] and more recently, a ransomware attack upon Colonial Pipeline that resulted in the gas pipeline being shut down for six days and a near $5 million payout to the hackers [30]. Attacks such as those can be facilitated by common problems; low-power IoT devices are resource constrained and are not resilient to denial-of-service attacks. A digital twin may create yet another vulnerability of an industrial or energy system.

Atalay and Angin propose that the digital twin can help to secure a cyber-physical system by establishing a security framework [10]. The framework should contain an extensible digital twin that represents the physical counterpart, a cyber-threat database containing the applicable attack vectors, attack simulation tools, and an analysis and reporting module [10].

The framework suggested by Atalay & Angin starts with the creation of the digital twin through specification and reuse of other models. The framework separates the modeling of the CPS into two phases, identification of the elements and their relationships, and building the network model using simulation. By creating the digital twin by using models, it is possible to achieve standardization, information reuse, and a level of security [10].

Digital twins may aid in physical system security by implementing rule-based and intrusion detection algorithms [4]. Eckhart and Ekelhart [4] identify a wide range of definitions for digital twins. Some definitions omit simulation and focus on visualization, such as implementing augmented reality (AR). AR applications have their security and quality concerns. Eckhart and Ekelhart also explain how a digital twin be used to implement intrusion detection for an industrial control system (ICS) or CPS. They identify concerns beyond intrusion detection, such as:

- the definition of digital twins has evolved from early intentions
- digital twins may have fidelity issues which may limit their use
- digital twins have a lifecycle that follows the lifecycle of their physical counterpart
- proper retirement of a digital twin system is imperative as it may contain private and confidential data

• digital twins may be used for security personnel training or in a cyber range as the attack target

• cost of digital twin creation can still be a prohibitive factor

Eckhart and Ekelhart have suggested reducing implementation costs by using previously established specifications of the physical system in the creation of its virtual counterpart [12]. AutoML and Mininet were suggested technologies to aid in the cost reduction of digital twin creation [12]. Eckhart and Ekelhart further proposed that digital twins can assist in intrusion detection by offering a baseline of past behavior that can be analyzed and compared to future states to detect malicious activity [12].

Methodology 4.3

Collection 4.3.1

This chapter focuses on the analysis of collections of Twitter tweets, as well as academic papers. Scripts have been written in R and executed throughout the research period. More R code has been utilized to clean the collected tweets, transform new classification data fields, and analyze them for n-grams, content-based, and sentiment analysis. The scripts use various libraries, such as the rtweet library as an API to the Twitter social media platform. Figure 24 below omits various keys and passwords yet illustrates the basics of library import and the implementation of a search and data frame store of tweets.

```r
library(rtweet)
library(ggplot2)
library(dplyr)
library(tidyr)

c_key <- ''
c_secret<- ''
access_token <- ''
access_secret <- ''
appdt <- ''

twitter_token <- create_token(
  app = appdt,
  consumer_key = c_key,
  consumer_secret = c_secret,
  access_token = access_token,
  access_secret = access_secret)

setup_twitter_oauth(consumer_key, consumer_secret, access_token, access_secret)

digitaltwin_tweets <- search_tweets(q = "#digitaltwin",since = '2020-06-08', n = 50000)
```

Figure 24: Screenshot of example usage of rtweet library within R Studio

Three scripts were written to collect tweets for this research. The first used a single search criterion, “#digitaltwin”, which has collected 3,102 tweets within 21 days of June 2020. The second and third scripts ran for the entire month of July 2020. 422,963 industry internet of things (IIoT) tweets and 497,174 cybersecurity tweets were collected. The tweets are stored with metadata made available via the rtweet API, but not all data has been retained. Additional data has been generated and stored such as the industry a tweet may mention and the sentiment score (positive or negative) of the text within the tweet. Table 7 lists the acquired and generated data.

Table 7: Tweet data fields

Field Name	**Short Description**
created_at	The date the tweet was originally posted to the Twitter platform
text	Character contents of the tweet
favorite_count	An integer representing the number of times a tweet has been marked as a favorite by Twitter users
retweet_count	An integer representing the number of times a tweet has been retweeted by Twitter users
country	Frequently NA within the dataset and sometimes providing the country of origin of the Twitter user posting the tweet
retweet_location	Frequently NA within the dataset and sometimes providing the country of origin of the Twitter user retweeting a tweet

Field Name	Short Description
Industry and count	A list of industries and the count of mentions from the gathered tweets
Sentiment	The individual sentiment score of a tweet's text

Data Preparation and Processing 4.3.2

Data preparation activities may account for 80 percent of the time invested in data science efforts [13]. In this study, the collection took many days, but the Twitter data was usable practically from the beginning. This is due to the limited search criteria used and the nature of the required data (such as we did not need the country or retweet_location fields for this research, both having many empty or not applicable entries). The longest time in preparation was determining industry mentions and labeling the tweets. Industry and identification keywords for cross-reference and classification efforts were collected from the International Labor Organization's website [14]. Preprocessing was done cross-referencing the ILO keywords to the tweet's text to determine the tweet's topics. This enabled mention analysis by the labor industry.

In the second corpus of tweets, the labeling of tweets allowed the naïve Bayes models to be trained and then tested. Tweets were labeled based upon the search criteria that were used to collect the tweets. This label would consist of either IoT or cybersecurity. Eighty percent of the data was used to train the models and twenty percent was utilized to test the model accuracy. Various models and parameters were used to predict the correct tweet label. Other preprocessing for the model input parameters included sentiment analysis and word lemmatization. The package textstem was utilized for word lemmatization and the packages tidytext and sentimentr were utilized for the sentiment analysis.

Results of First Corpus Analysis 4.4

The first corpus contained 3,102 tweets found by using the search criteria #digitaltwin. Common stop words were removed from the corpus before generating the word cloud. A word cloud was created to show the most common bigrams found in the corpus of tweets. While concerning the word ‘security’ is not paired with another common word frequently enough to make this bi-gram chart, we do see those conversations on Twitter around defense, risk, and the industrial use of IoT are occurring, each having relevance towards this research. The word cloud is illustrated in Figure 25.

Figure 25: Word cloud of bigrams within the tweet collection text

The word cloud used in Figure 25 omits counts of the actual instances of the bigrams. Table 8 outlines a few common security terms and the number of instances they occur within our first corpus of 3,102 tweets collected by searching #digitaltwin. Such terms may occur in one tweet multiple times. Overall, terms such as “encrypt”, or “vulnerability” have almost no mention in such tweets and do not appear as a part of the social media conversation which is a concerning observation.

Table 8: Common security terms and their occurrence

Terms	**Count of occurrences**
“security”|”secure”	88
“risk”	212

“defense”|”defensive”	179
“iiot”|”industrial iot”	750
“encrypt”|”decrypt”|”cipher”	1
“vulnerability”|”threat”	1

We observe a dearth of conversation within the social media analysis towards securing digital twins, which could be used to implement the security models needed. The percentage of tweets containing common security terms makes up 10% of the total volume of tweets collected using the #digitaltwin search criteria within the specified time range, as illustrated in Figure 26.

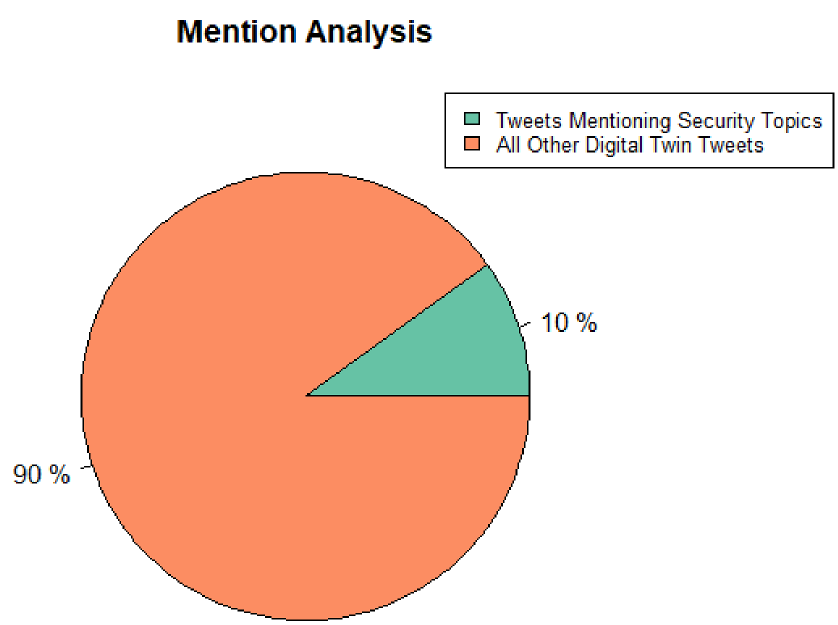

Figure 26: Most IoT tweets do not mention security topics

Most Favored and Most Retweeted Tweets about Digital Twins 4.4.1

Tweets can be favored as well as re-tweeted. The number of favorites indicates the count of unique user accounts that like or agree with the content of the tweet [15]. In the first corpus, the most favored tweet is also the most retweeted. It will be discussed later in this section. To provide more breadth of the findings, the second most favored tweet [16] is also illustrated in this chapter (Figure 27).

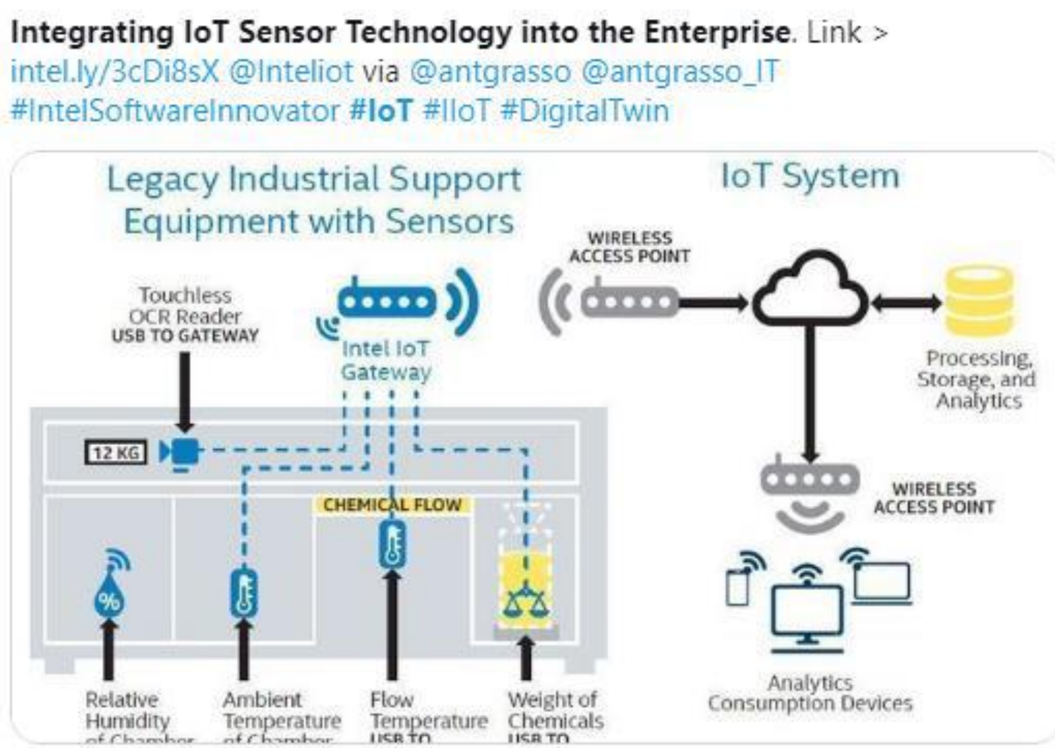

Figure 27: Second most favored tweet within the scope of this research [16]

The tweet in Figure 27 refers to Intel technology and links to a 2015 white paper where Intel published best practices to help with IoT sensor technology implementation against the challenges cost, security, and scalability [17].

Within the scope of this first corpus, the most retweeted tweet is about the use of digital twin used in contact tracing to help prevent the spread of Covid-19. The digital twin was specifically built and used to help contain the spread of the virus in Dharavi, a location within Mumbai that houses nearly one million residents. Many of these residents rely on daily wages. Living and working conditions may reduce or prevent social distancing. During the time of the tweet creation, epidemiologists had credited the Dharavi community for their efforts towards containment of the Covid-19 virus [18]. Figure 28 is the most favorited tweet within the first corpus [19].

Figure 28: Most retweeted tweet within the research scope [19]

While we cannot confirm the accuracy of the tweet illustrated in Figure 28, we can recognize from the tweet contents the broad application of digital twin usage to monitor, model, and control a pandemic. The nature of such a system and the associated data also provide support for why such the system requires a mature security posture, to protect individuals and their privacy rights.

Mention and Sentiment Analysis 4.4.2

Studies show digital twins being utilized across industries; mention analysis from social media provides an additional perspective. To analyze industry mention among the first tweet collection, we crossed referenced the tweets' text with a list of known industry categories [14]. The simple algorithm reads the tweet texts and counts industry mentions, including when multiple industries are mentioned within a single tweet's text. Figure 29 illustrates mention analysis results with health, education, and public as the top three industries or sectors with the most mentions within the tweets' text.

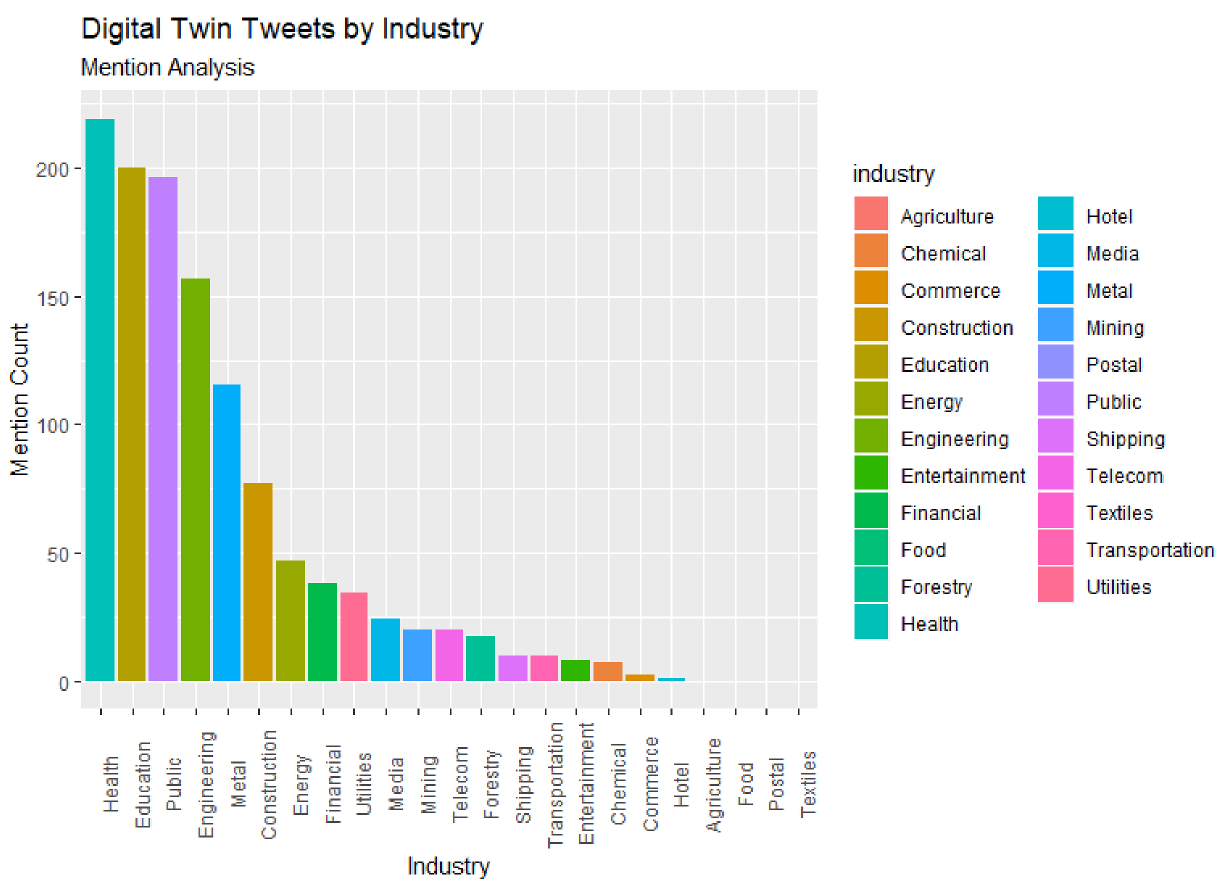

Figure 29: Industry mention analysis: health, education, and public are the top three mentioned industries

Utilizing the sentimentr package, each industry group has the individual tweet text sentiment scores averaged and placed into a chart. The sentiment scores from the sentimentr package may range from a negative 2.0 to a positive 2.0. While the health and public industries were high in the mention-analysis, the sentiment scores for those industry groups are negative. The three industry groups having the highest average sentiment scores are engineering, commerce, and energy, illustrated in Figure 30.

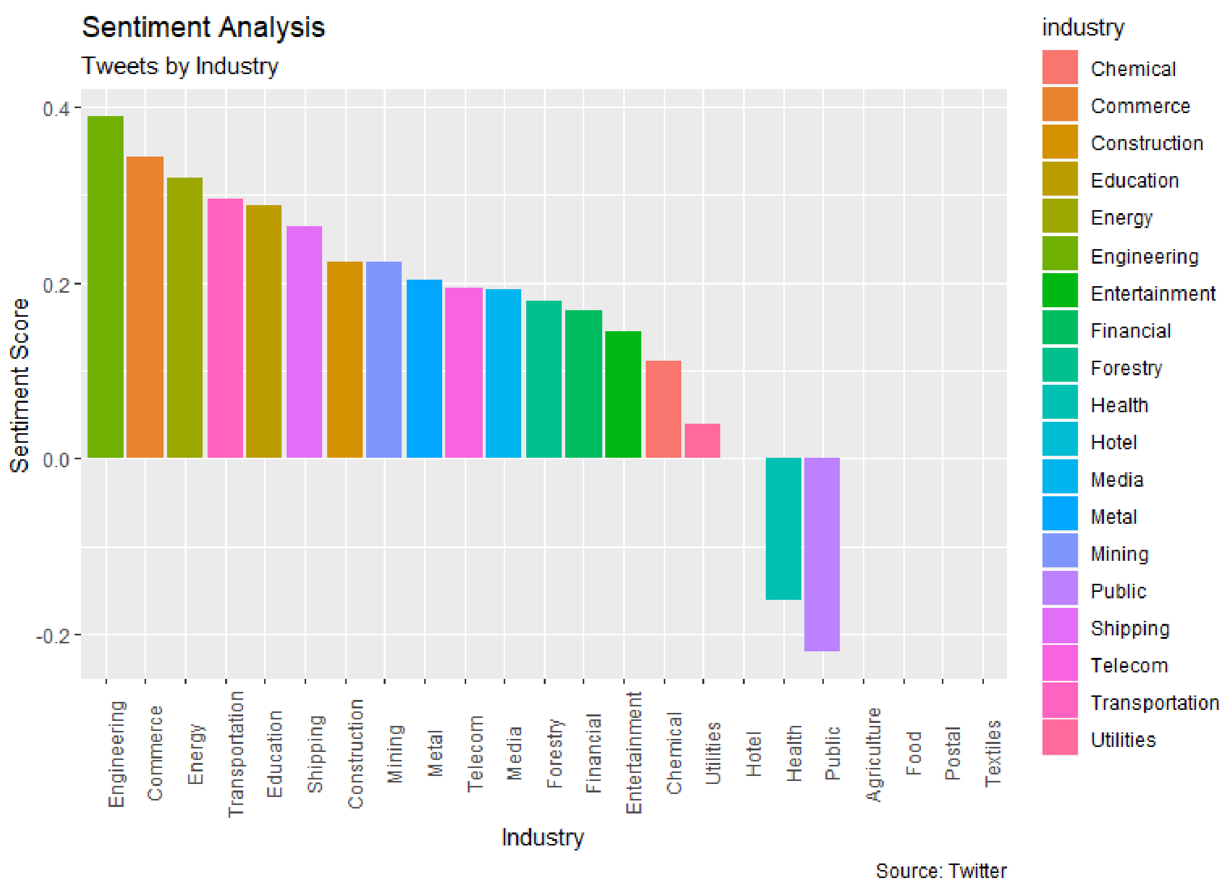

Figure 30: Aggregated sentiment scores grouped by industry

The tweet in the engineering group having the highest sentiment score is given in Figure 31 [20]. This tweet indicates the usage of digital twins to assist in the simulation of electric vehicles for optimal energy management.

Figure 31: Tweet having highest sentiment score within the engineering group [31]

The electric vehicle market is high stakes. The current pandemic threatens the automotive supply chain [21], marketing of a Tesla truck towing a competitor's vehicle (F-150) up a hill made headlines [22], and Tesla competitor Nikola Motors saw a one-day stock jump of 104% [23]. Security posture is required around the intellectual property in this dynamic and competitive market. An elevated security posture must be applied when implementing digital twins in this field and others.

Results of Social Media Analytics of IIoT and Cybersecurity Tweets 4.5

Multiple prediction models were created to further inspect the tweets. This included attempts to predict whether a given tweet was towards the topic of cybersecurity or IIoT. Methods such as word stemming, lemmatization, and term frequency-inverse document frequency (TFIDF) were utilized in naïve Bayes and general linear models.

A naïve Bayes model that attempted to predict the label, cybersecurity or IIoT, of a given tweet using word lemmatization and sentiment analysis could predict the label correctly with a 70.3% accuracy. While this prediction accuracy is not high, perhaps the sentiment coefficient correlations are more telling. As previously stated, security must start with the organization's culture [9]. An organization's culture consists of the beliefs and values of the employees.

Communication incorporating the sentiment of the topic may be more influential towards influencing behavior or belief. The sentiment level of the cybersecurity tweets, when compared to the IIoT tweets, was higher in the areas of anger, disgust, fear, negative, and sadness, whereas the IIoT tweets sentiment level, when compared to cybersecurity tweets, was higher in anticipation, joy, positive, surprise, and trust. Table 9 lists the average sentiment scores for each classification of sentiment by the cybersecurity and IIoT labels.

Table 9: Conditional probabilities of IIoT and cybersecurity tweet sentiment

Sentiment	**Cybersecurity**	**IIoT**
Anger	0.07145	0.03765
Anticipation	0.11663	0.14300
Disgust	0.01868	0.01400
Fear	0.11321	0.06652
Joy	0.05369	0.07048
Negative	0.14541	0.07757
Positive	0.25503	0.35884
Sadness	0.04404	0.02833

Given a tweet is about Cybersecurity, there is a 14.5% probability that the tweet is negative, an 11.6% probability that the tweet provokes anticipation, and an 11.3% probability that the tweet has the sentiment of fear. This information may be key for communicating cybersecurity strategies and plans within an organization. We should consider the sentiment of our security posture messaging, and how it fits into the culture, behaviors, and values of the communication's target audience.

The Term Frequency-Inverse Document Frequency (TFIDF) score utilization within a naïve Bayes model had less accuracy, 56.3%, compared with the model using both word lemmatization and sentiment analysis. TFIDF does allow us to see important words in a corpus, we found that “bugbountytip” had the highest TFIDF score within the cybersecurity corpus.

Application of a Digital Twin as a Multi-Model Security Architecture 4.6

Digital twins are commonly used to help monitor, model, and control cyber-physical systems. Engineering a digital twin should be done using a systematic approach by performing a needs analysis from the physical device and context of the operating environment [1]. Needs analysis and environmental context must include non-functional requirements, including cybersecurity. Securing the information of these cyber-physical systems is a growing concern [24]. Beyond the common monitor, model, and control mechanisms, digital twins may also be utilized for incorporating multiple security models.

Several security control models exist. The information flow model is a common mechanism to maintain information confidentiality [25]. We outline in Table 10 how the information flow model may also work to maintain the integrity of a digital twin and its physical counterpart. Multiple models will need to be implemented in the system engineering effort to establish a secure posture. Each model has a strength, such as Bell-LaPadula for protecting the confidentiality of information [26]. Furthermore, in Table 10, we indicate how the lattice model, incorporating Biba, can work to ensure information integrity.

The digital twin may also incorporate concepts and techniques such as automated policy exchange (APEX) to prevent data leakage. APEX technology works to keep system users working within the guidance of security best practices and policies even when policy is complex or unknown to operators [27]. Using APEX, the digital twin would monitor labeled files throughout their lifecycle and use control mechanisms, such as deployed agents within the cyber-physical system, to prevent or warn on file activities by users.

Another possible avenue would be for the monitoring portion of the digital twin to utilize a System Call Intercept (SCI) framework. SCI works to reduce data leakage by comparing policy to user action and stopping some user-based activity within the system from completing [28]. Using the SCI framework, the monitor and modeling functions of the digital twin would use policies to review specific subjects or objects (perhaps all users and system components within a zero-trust approach) and trap some actions for review before the OS or another system component completing the action. An inventory of objects, subjects, and system calls that could be trapped would be required to implement such a framework, which will improve data protection. This inventory could be complemented with behavior analysis by machine learning models to identify malicious activity. Machine learning can be used for countering cybersecurity threats and vulnerability mitigation using regression, prediction, and classification techniques.

Table 10 provides high-level examples of how a digital twin implementation may satisfy or utilize common security models such as State Machine, Lattice, Interference, and Information Flow models.

Table 10: Application examples of a digital twin as a multi-model security schema

Security Model	**Example Digital Twin Application**
State Machine	Digital twin control input is first executed within the model. Validation occurs within the model for an appropriate output state. The input may be rejected before entering the digital twin control mechanism for actual processing on the cyber-physical system.
Non-Interference Model	Monitoring and control mechanisms must obfuscate users reading (monitoring) and writing (controlling) across domains. This reduces the probability of users being influenced by the actions of subjects with greater/lesser clearance. This approach may utilize common discretionary access control (DAC) across file systems, mandatory access control (MAC) through rule-based systems, and role-based control. DAC and MAC limit the availability of information objects to certain users and processes to retain confidentiality [29].
Lattice Model	Biba protects the integrity of information by preventing a subject from reading untrusted information [29]. Using the Biba Model approach, the digital twin modeling process accesses only known secured sensors (having a level of integrity through controls). While we may monitor many sensors, decision dashboards and predictive analytics should utilize only trusted and secured sensors. A matrix of such processes and secure sensors would instruct inputs for such high integrity algorithms.

Information Flow Model	Sensor information flow is one way. The Digital twin may read sensor data for monitoring or modeling purposes. No data is transmitted to a sensor to be later sent back, as sensor-originated data. This prevents levels of spoofing sensor data, by not allowing writes to the sensor memory. Any PUT type of action to sensors APIs would be rejected and logged as malicious intent.

Figure 32 illustrates a situation where a population, Bluetooth sensors from mobile devices, network paths, and system logic are used for the creation of a population's digital twin. In this case, the digital twin of the population would be used by public health officials and the application subscribers for contact tracing. The diagram's purpose is to provide examples illustrating the concepts given in Table 10 applied to a hypothetical contact tracing system.

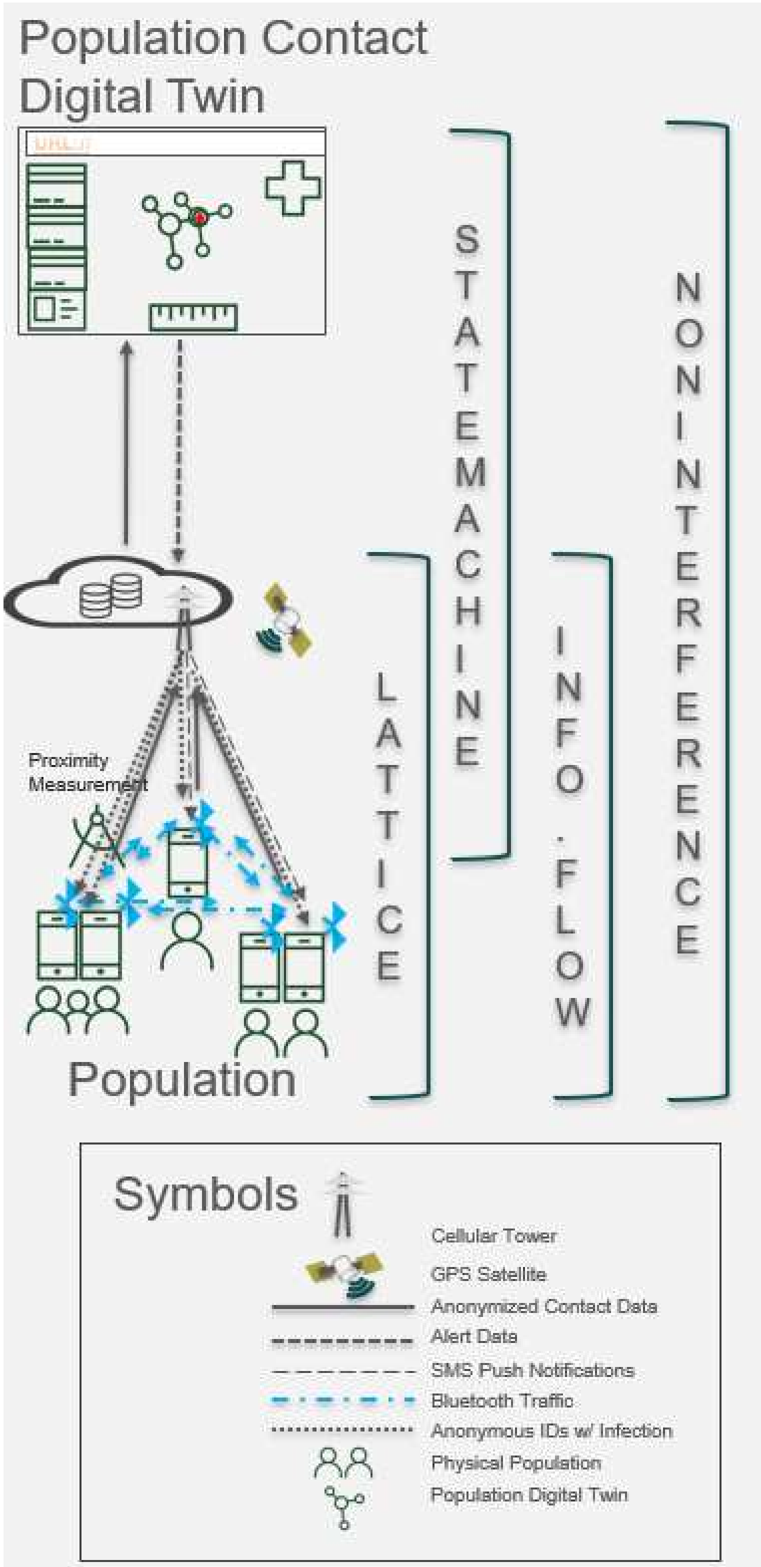

Figure 32: Applying the multi-model in digital twin security

If contact tracing systems were injected with malicious data, it could alter the behaviors of a population. Applying the lattice model helps assure that a system having integrity does not read from a system or sensor with low integrity. The lattice model can prevent a malicious actor's device farm from spoofing the contact system. To validate the integrity of a sensor or mobile device in this case, the system may validate a history of GPS or use human interaction to assure the device is not a simulator or otherwise driven by a malicious bot.

Public health officials may instruct the system to deploy warning notifications to participants who have opted in to use the feature. However, before deploying the notification, the digital twin should utilize the state machine model. Again, a notification could be a malicious message being sent to alarm a population. The logic should first model that notification within the digital twin. The modeling should evaluate the message accuracy and may even predict how the population may behave to help public health officials craft response plans. If the notification messages were not related to a contact tracing scheme, the message may be malicious and should be stopped.

The information flow model is utilized to protect the population and their privacy. The mobile device's application logic allows Bluetooth connections to write to ledgers indicating contacts. This data is anonymized and does not indicate a person's identity. Writing to this ledger is only allowed by the local mobile device which is in the possession of the user. The centralized system does not allow writing to this contact ledger. The mobile device application will pull anonymized alert data from the central system. This data is written to a separate ledger. The only mechanism for the centralized system to send data to the host devices is through the SMS push notification mechanism, which has levels of security such as users being required to opt-in and the state machine model.

Users of the Population Contact Digital Twin dashboard will utilize roles, authentication, and authorization to access the digital twin of the population. Low entitlement users would not have the ability to interfere with the population by spying on them using inference techniques.

Entitled users, such as medical staff, may be granted access to information having more fidelity to help protect the population's health. This protects a population using the non-interference model with access controls.

Conclusions and Future Work 4.7

Conclusions 4.7.1

Digital twins operate from and create new cyber-physical system information. A proper design approach for a digital twin should model its cyber-physical system, including its future states. In addition to monitoring, modeling, and controlling the cyber-physical system, a digital twin must also provide security. This fourth element as applied to the prior three digital twin requirements allows for a layered approach to securing cyber-physical systems by implementing multiple security models. While past researchers have focused on intrusion detection, this article proposes that common security control models such as state machine, non-interference, lattice, and information flow models can be utilized in digital twin implementation to secure itself the associated cyber-physical system.

In the first corpus of tweets collected in this study by searching for #digitaltwin, those referencing security concepts represented a small population of 10%. In the second and larger corpus of the collected tweets, the top mentioned industries or business sectors were health, education, and public services in that order. The top three industries having the highest tweet sentiment were engineering, commerce, and energy. While naïve Bayes models reached only a 70.3% accuracy at differentiating a tweet that was about cybersecurity versus a tweet about digital twins, the sentiment findings of the messages leaned towards negative messaging from the cybersecurity tweets.

Future Work 4.7.2

Most of our research utilized social media, which can be time-consuming to prepare such data and has limits in what it can contribute. Future research should include detailed case studies of direct implementation which would strengthen the approaches proposed in this work. The sentiment findings of cybersecurity tweets deserve further research as other researchers have stated that security must start with an organization’s culture. The implication of negative messaging on an organization’s cybersecurity posture and culture needs to be studied.

REFERENCES

[1] Scheibmeir, J., & Malaiya, K. (2019). An API development model for digital twins, *In 2019 IEEE 19th International Conference on Software Quality, Reliability and Security Companion (QRS-C)*, 518-519. https://doi:10.1109/QRS-C.2019.00103.

[2] Saddik, E. A. (2018). Digital twins: the convergence of multimedia technologies. *IEEE Multimedia Mag.*, 25, 87-92.

[3] Martino, B., Di, M., Rak, M., Ficco, A., Esposito, S., & Nacchia, S. (2018). Internet of things reference architectures, security and interoperability: a survey. *Internet of Things*, 1-2, 99-112. https://doi:10.1016/j.iot.2018.08.008

[4] Eckhart, M., & Ekelhart, A. (2019). Digital twins for cyber-physical systems security: state of the art and outlook. *Security and Quality in Cyber-Physical Systems Engineering*. Springer, https://doi:10.1007/978-3-030-25312-7_14

[5] Canedo, A. (2016). Industrial IoT lifecycle via digital twins. *Proceedings of the Eleventh IEEE/ACM/IFIP International Conference on Hardware/software Codesign and System Synthesis. ACM*, 1-1. https://doi:10.1145/2968456.2974007

[6] Yadav, D., Jayanthu, S., Das, S., Chinara, S. & Mishra, P. (2019). Critical review on slope monitoring systems with a vision of unifying WSN and IoT. *IET Wireless Sensor Systems*, https://doi:10.1049/iet-wss.2018.5197

[7] Wang, Y., Garhan, A., & Byrav, R. (2006). A survey of security issues in wireless sensor networks. *IEEE Communications Surveys & Tutorials*, 8(2), 2-23. https://doi:10.1109/comst.2006.315852

[8] Porambage, P., Schmitt, C., Kumar, P., Gurtov, A. & Ylianttila, M. (2014). Two-phase authentication protocol for wireless sensor networks in distributed IoT applications. *IEEE Wireless Communications and Networking Conference*, https://doi:10.1109/WCNC.2014.6952860

[9] Hearn, M., & Rix, S. (2019). Cybersecurity considerations for digital twin implementations report. *IIC Journal of Innovation*

[10] Atalay, M., & Angin, P. (2020). A digital twins approach to smart grid security testing and standardization. *IEEE International Workshop on Metrology for Industry 4.0 & IoT*, 435-440, https://doi:10.1109/MetroInd4.0IoT48571.2020.9138264

[11] Girma, A. (2018). Analysis of security vulnerability and analytics of internet of things (IoT) platform. *In Information Technology - New Generations. Advances in Intelligent Systems and Computing*, 738, https://doi:10.1007/978-3-319-77028-4_16

[12] Eckhart, M., & Ekelhart, A. (2018). Towards security-aware virtual environments for digital twins. *In Proceedings of the 4th ACM Workshop on Cyber-Physical System Security, ACM*, 61–72, https://doi.org/10.1145/3198458.3198464

[13] Cirillo, A. (2016). *RStudio for R Statistical Computing Cookbook*. Birmingham Packt Publishing.

[14] International Labor Organization. “Industries and Sectors.” Industries and sectors, https://www.ilo.org/global/industries-and-sectors/lang--en/index.htm.

[15] Perdana, R.S., & Pinandito, A. (2018). Combining likes-retweet analysis and naive bayes classifier within twitter for sentiment analysis. *Journal of Telecommunication, Electronic, and Computer Engineering*, 10(1-8), 41-46.

[16] Grasso, A. [@antgrasso]. (2020, June 8). Integrating IoT Sensor Technology into the Enterprise. [Twitter moment]. Retrieved from: https://twitter.com/antgrasso/status/1269962528530071552

[17] Gaiser, K. H., Harrington, V. P., Loughrin, G. T., Meyer, S. J., & Sartini, J. M. (2015). Integrating IoT sensor technology into the enterprise. December 2015. https://www.intel.com/content/www/us/en/it-management/intel-it-best-practices/integrating-iot-sensor-technology-into-the-enterprise-paper.html

[18] Parth, M.N., & Bengali, S. (2020). "How an Indian Slum Became an Unlikely COVID-19

[19] Success Story." Los Angeles Times. June 24, 2020. https://www.latimes.com/world-nation/story/2020-06-24/dharavi-slum-in-mumbai-india-contained-covid-19.

[20] Zuppa [@ZuppaGeoNavTech]. (2020, June 17). "Offence Is the Best Defense" [Twitter moment]. Retrieved from: https://twitter.com/ZuppaGeoNavTech/status/1273449554797408257

[21] Simcenter [@Simcenter_] (2020, June 12). Learn more about #simulation for optimal vehicle energy management in our new on-demand webinar series. [Twitter moment] Retrieved from: https://twitter.com/Simcenter_/status/1271442013586624512

[22] Harbour, L. (2020). "The Coronavirus' Impact on The Global Automotive Supply Chain." Forbes. March 14, 2020. https://www.forbes.com/sites/laurieharbour1/2020/03/13/the-coronavirus-impact-on-the-global-automotive-supply-chain/#17fd6b9f444e.

[23] Wayland, M., & Kolodny, L. (2019). "Tesla Cybertruck Tug-of-war with a Ford F-150 Is High Stakes and Highly Unlikely." CNBC. November 27, 2019. https://www.cnbc.com/2019/11/26/tesla-cybertruck-tug-of-war-with-a-ford-f-150-is-high-stakes-unlikely.html.

[24] Fox, M. (2020). "Tesla Competitor Nikola Motors Skyrockets 104% after Setting Reservation Date for Its Electric Truck (NKLA, TSLA) | Markets Insider." Business Insider. June 8, 2020. https://markets.businessinsider.com/news/stocks/nikola-motors-stock-price-soars-setting-truck-reservation-date-competitor-2020-6-1029290291.

[25] Humayed, A., Lin, J., Li, F., & Luo, B. (2017). Cyber-physical systems security—a survey. *IEEE Internet Things J.*, 4, 1802-1831. https://doi:10.1109/JIOT.2017.2703172

[26] Sabelfeld, A. & Myers, A. C. (2003). Language-based information-flow security. *IEEE J. Sel. Areas Commun.*, 21(1), 5-19. https://doi:10.1109/JSAC.2002.806121

[27] Ma, X., Huang, Y., & Li, D. (2010). A security model based on lattice. *International Conference on Electrical and Control Engineering,* 4958-4961. https://doi:10.1109/iCECE.2010.1199

[28] Simske, S.J., & Balinsky, H. (2010). Apex. *Proceedings of the 10th ACM Symposium on Document Engineering*, 10, https://doi:10.1145/1860559.1860587

[29] Balinsky, H., Perez, D. S., & Simske, S. J. (2011). System call interception framework for data leak prevention. *In 2011 IEEE 15th International Enterprise Distributed Object Computing Conference*, https://doi:10.1109/edoc.2011.19

[30] Schou, C., & Hernandez, S. (2015). *Information assurance handbook: effective computer security and risk management strategies*. New York: McGraw-Hill Education.

[31] Turton, W., Riley, M., & Jacobs, J. (2021). “Colonial Pipeline Paid Hackers Nearly $5 Million in Ransom”. Retrieved May 14, 2021, from https://www.bloomberg.com/news/articles/2021-05-13/colonial-pipeline-paid-hackers-nearly-5-million-in-ransom

CHAPTER 5: SOCIAL MEDIA ANALYTICS OF THE INTERNET OF THINGS

The Internet of Things technology offers convenience and innovation in areas such as smart homes and smart cities. Internet of Things solutions require careful management of devices and the risk mitigation of potential vulnerabilities within cyber-physical systems. The Internet of Things concept, its implementations, and applications are frequently discussed on social media platforms. This research illuminates the public view of the Internet of Things through a content-based and network analysis of contemporary conversations occurring on the Twitter platform. Tweets can be analyzed with machine learning methods to converge the volume and variety of conversations into predictive and descriptive models. We have reviewed 684,503 tweets collected in a two-week period. Using supervised and unsupervised machine learning methods, we have identified trends within the realm of IoT and their interconnecting relationships between the most mentioned industries. We have identified characteristics of language sentiment which can help to predict the popularity of IoT conversation topics. We found the healthcare industry as the leading use case industry for IoT implementations. This is not surprising as the current Covid-19 pandemic is driving significant social media discussions. There was an alarming dearth of conversations towards cybersecurity. Recent breaches and ransomware events denote that organizations should spend more time communicating about risks and mitigations. Only 12% of the tweets relating to the Internet of Things contained any mention of topics such as encryption, vulnerabilities, or risk, among other cybersecurity-related terms. We propose a cybersecurity communication scorecard to help organizations benchmark the density and sentiment of their corporate communications regarding security against their specific industry.

Introduction 5.1

The Internet of Things (IoT) is an appealing technology that has eased the management of homes through smart appliances and has enticed industries such as automotive, transportation, and agriculture [1]. IoT was first introduced in 1999 as a concept for solving problems within the logistics domain [2]. IoT can decrease the distance between cloud technology, people, and things [3]. Today, consumers of data are also producers of data. Twitter users tweet nearly 277,000 times every single minute [4]. The action of liking or retweeting a tweet is yet another data point.

We have collected 684,503 tweets within a two-week period from May 1st, 2021, through May 14th, 2021. Twitter data has been utilized in several recent research investigations [5-7]. Social media platforms have been found to support access to information, discuss and solve engineering problems, identify new trends, and communicate science to a public audience [8-12] We extend the collected Twitter data with metadata using hierarchical clustering techniques and content-based analysis. The clustering algorithm is informed of proper cluster distribution by the within-cluster sum of squares (WSS) and average silhouette methods. A mention analysis is then performed to identify the number of industries, trends, and technology vendors having a presence in the tweets. Sentiment analysis is carried out for tweets classified towards the industries and vendor technologies. Factors such as the trend labels, industry labels, and sentiment scores are then used in naïve Bayes prediction models. We illustrate the relationships, or lack of, between the trends, industries, and technology providers utilizing network graphs. Section two contains a brief background on the topics of IoT and social media. The research and analysis methodology is described in detail within section three. Finally, the fourth and fifth sections offer discussion and conclusion to the research. The main contributions of this research work include:

- Using advanced statistical and machine learning (ML) methods including naïve Bayes, hierarchical clustering, and natural language processing with sentiment analysis, we evaluate 684,503 contemporary tweets on the topic of the Internet of Things to shed light on public opinion, technology trends, popular industry usage and the popularity and sentiment of technology providers in this space.

- We uncover the substantial problem of a lack of cybersecurity discussion within the IoT tweets. No cybersecurity concepts were identified in the top ten trends. Organizations must increase their cybersecurity communication cadence to meet the risks.

- We analyzed tweets to identify industries where IoT concepts and technology are being discussed. We found healthcare to be the leading industry of mention.

- We propose a new IoT Cybersecurity Communications scorecard. The score uses a combined index of mention density and sentiment analysis to provide a benchmark of cybersecurity communication posture scores by industry.

- The top three trends identified within the IoT tweets were data science, machine learning, and big data. We performed a network analysis to identify relationships between trends and industries, such as what industries have the greatest or least inclusion of trending concepts and technology.

- We evaluate commercial vendors by the sentiment of messages where they are discussed, as well as the volume of mentions. We provide a positional rank of a selection of IoT commercial technologies based upon this analysis.

This chapter benefits cybersecurity experts, IoT practitioners, and commercial firms. Cybersecurity practitioners and organizational leaders can utilize our findings and scorecard to benchmark areas of their internal behavior. Practitioners, such as developers and engineers of IoT systems, can utilize this research to identify trends within the realm of IoT. Marketing experts of commercial firms benefit from the sentiment analysis and predictive models that shed light on Twitter user behavior regarding the communication of IoT systems. Our contributions are further discussed.

Background work 5.2

Use of social media in research 5.2.1

The Twitter data has been utilized in several recent research investigations [5-7]. The public availability of the tweets allows researchers to extract valuable conclusions from them [13]. It has been found that the geotagging of twitter users' tweets can complement surveys as well as enhance a sampling profile [14]. The same study found that their survey showed bias towards elderly participants while the Twitter data was biased towards a younger population. The researchers utilized these conflicting biases to balance their findings. A study of 640 university students found that the leading factor for using social media was to search for and access information [8]. An earlier study by Bougie et al. [9] followed software engineering groups on Twitter to determine how they utilized the platform. This study found that 23% of the groups' tweets were towards software engineering topics. Of that 23% of their total tweets which regarded software engineering, 62% were towards solving software engineering problems. Another study sought to answer if software engineering practitioners use and cite scientific research in their blogs; they do not [10]. Rather, software engineering practitioners utilize social media to become up to date on technology trends [11]. Another research article states that microblogging serves by linking to web resources, connecting users, and directing users' attention, as well as offering another channel for the public communication of science [14].

Related works on the Internet of Things 5.2.2

Implementing an IoT system requires storage, networks, load-balancing, and analysis tools. According to Atalay and Angin [15], an IoT solution should utilize network partitions in private clouds which provide partitioning for enhancing security. Such network partitions could encapsulate the concerns of actuators and sensors, the model of system states, and the business and program logic. An encryption key management system would be utilized to support encryption across network enclaves and an intrusion detection system (IDS) could be implemented to identify malicious activity.

The growing interest in IoT and the implementation of the systems have resulted in large cyber-attack surfaces [15]. A few well-known cyber-physical system attacks include the Stuxnet effect on an Iranian uranium enrichment plant [16] and more recently, a ransomware attack upon Colonial Pipeline that resulted in the gas pipeline being shut down for six days and a near $5 million payout to the hackers [17]. There are also known consumer exploits including hacked Smart TVs listening to conversations, personal information being extracted from coffee machines, and security cameras leaking images [18]. Today, IoT implementations may be utilized to carry medicine, medical samples, and to assist with the management of pandemics [19], furthering the need for rigor and security in the implementations. In a 2019 research survey of 220 security leaders in industrial and manufacturing, 79% of respondents indicated they had experienced an IoT cyberattack within that past year [20]. The security aspects of IoT have the attention of legitimate organizations who seek to enhance the defense as well as the hackers.

The cybersecurity concerns of IoT systems are growing in complexity and have insufficient security solutions [15]. The evaluation of cyber-physical system component vulnerabilities is a challenging task due to the sheer number of devices and their varied configurations. Common threats include denial of service (DoS) attacks while a common weakness is insecure wireless networking [15]. The complexity of IoT systems and their emergent behavior also complicate the testing of the systems [21].

To manage the complexity while achieving value and providing security of the system's assets and users, five best practices have been suggested by Shi et al. [3]. Good service management of edge computing and IoT systems include these five considerations:

- Differentiation in device identification to discern specific state metrics such as the health of the specific device instances.

- Extensibility in the system to allow for replacement endpoints being easily swapped in and out.

- Isolation of access and data via roles and other controlling factors.

- Security/Privacy which preserves availability and confidentiality characteristics.

- Optimization of the system and components' attributes such as cost, latency, or bandwidth.

While Shi et al. [3] mention the optimization of system components including latency and bandwidth, Fizza et al. [22] dive deeper into optimization stating that existing definitions of quality of experience (QoE) must be renewed with the autonomous IoT systems in mind. The same research found that if QoE is not considered in autonomous IoT applications, poor quality of decisions and resulting actions may occur. Motta et al. [23] have examined the IoT-related literature to find twenty-nine definitions of the concept. Connectivity, a component of QoE, is among the common concepts within the definitions Motta et al. distilled. From those twenty-nine definitions, they have identified seven key facets. These facets must be considered when engineering an IoT software system. They include:

- Connectivity includes the medium for things to connect to implement the IoT paradigm. Connectivity may be challenged by security concerns or the quality of service.

- Things include the number of heterogeneous tags, sensors, actuators, among other things. There exist challenges of maintaining the identities of these devices as well as managing their behavior.

- IoT systems may exhibit emergent behavior, which are side effects that are difficult to predict and are a result of compositions of parts into components and subsystems. The main causes of emergent behavior is due to the complexity of systems and the human interaction within them [24].

- The smartness of the things within the IoT system relates to how devices are managed, orchestrated, and their allowance and use of autonomous behavior.

- Problem Domain may refer to the industry or specific problem that the IoT software system is built to alleviate.

• Interactivity is not limited to the interaction between things and humans, but also the interaction amongst things within the IoT system.

• The environment is the context in which an IoT system operates and can also be specific to the problem domain or implementation.

Methodology 5.3

Data acquisition and preprocessing 5.3.1

We wrote and utilized an R program to manage the downloading of tweets from Twitter's application programming interface (API). Another R program was created to label the tweets and to perform the content-based analysis. The analysis begins with preprocessing the tweets including the removal of stop words and usage of word stemming and lemmatization. The analysis includes an identification of trends within IoT discussions. The tweets are labeled for the factors of popularity (tweets that were liked or retweeted), industry mention, commercial vendor technology mention, and trend identification. There is an evaluation of sentiment within the labeled tweets. We also analyze the relationships between the factors of industry and trending terms. A naïve Bayes model is created to determine whether our labeled factors can predict the content or popularity of the tweets. Using the factors of being favored, industry type, retweet, and IoT vendor name, we could predict the trend a tweet was referencing with an accuracy of 63.9%. Two R programs carry out seven steps across the lifecycle of this analysis. The R programs and a compressed CSV file of the 684,503 tweets are available for use and evaluation on a publicly available Gitlab site [25]. The following illustration presents our methodology in seven steps.

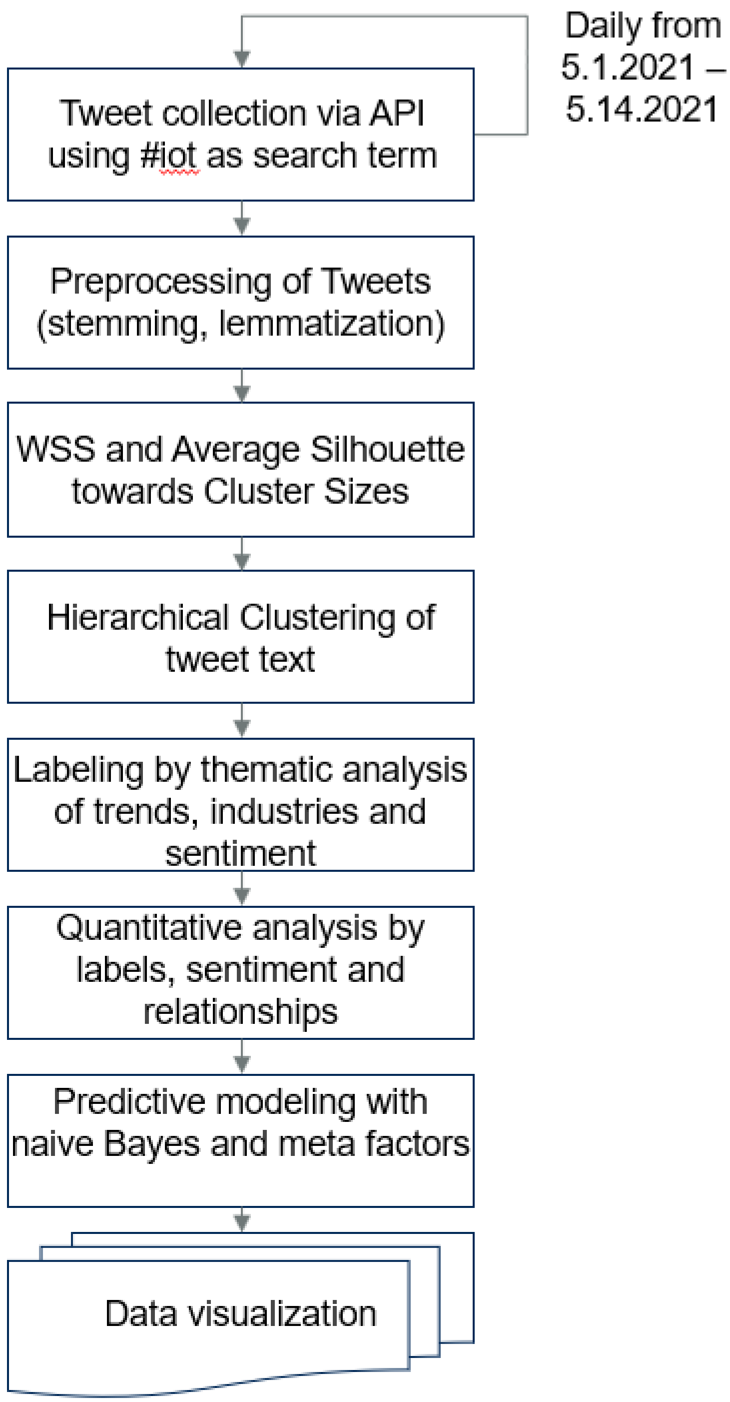

Figure 33: Seven steps make up the methodology starting with an iterative collection of tweets through labeling and analysis until the visualization of the data

To perform the collection of tweets, we first created a programming account on the Twitter platform. This account creation offered the authentication and authorization needed to access the Twitter platform via API. For the first fourteen days of May 2021, we searched for tweets containing #iot and stored up to 50,000 per day. The search limit was required as our AWS EC2 server instance is limited to four cores and 32GB of memory. The impact of the limited server resources will be described later in this section. By the last day of tweet collection, we had successfully captured 684,503 tweets containing #iot.

Number of cluster determination 5.3.2

After data collection, we created a document-term object matrix. The individual words from each tweet were then cast into the matrix and their frequency of appearance recorded. To determine an ideal number of clusters, we utilized within-cluster sum of squares (WSS) and the average silhouette methods. However, due to the size of the term matrix as input into these methods and the restrictions of our compute environment, only samples of the entire tweet corpus were used to generate the term matrix.

The WSS method will iterate through many generations of k-means clusters. During each iteration through k number of clusters, the squared distance between a cluster's observations (within the cluster) and the clusters' centroid are summed and plotted for the given number of clusters. This is done for all clusters and compared for Euclidean distance over the iterations. The ideal number of clusters is frequently determined visually, known as the "elbow method" and identified when the WSS is decreasing and the next increment in cluster generation does not offer much benefit. This is often visually detected by looking for the "elbow" or the "knee" in the line chart where the WSS has dropped and then flattens. The illustration below identifies the knee at four clusters for our dataset of #iot tweets collected over two weeks.

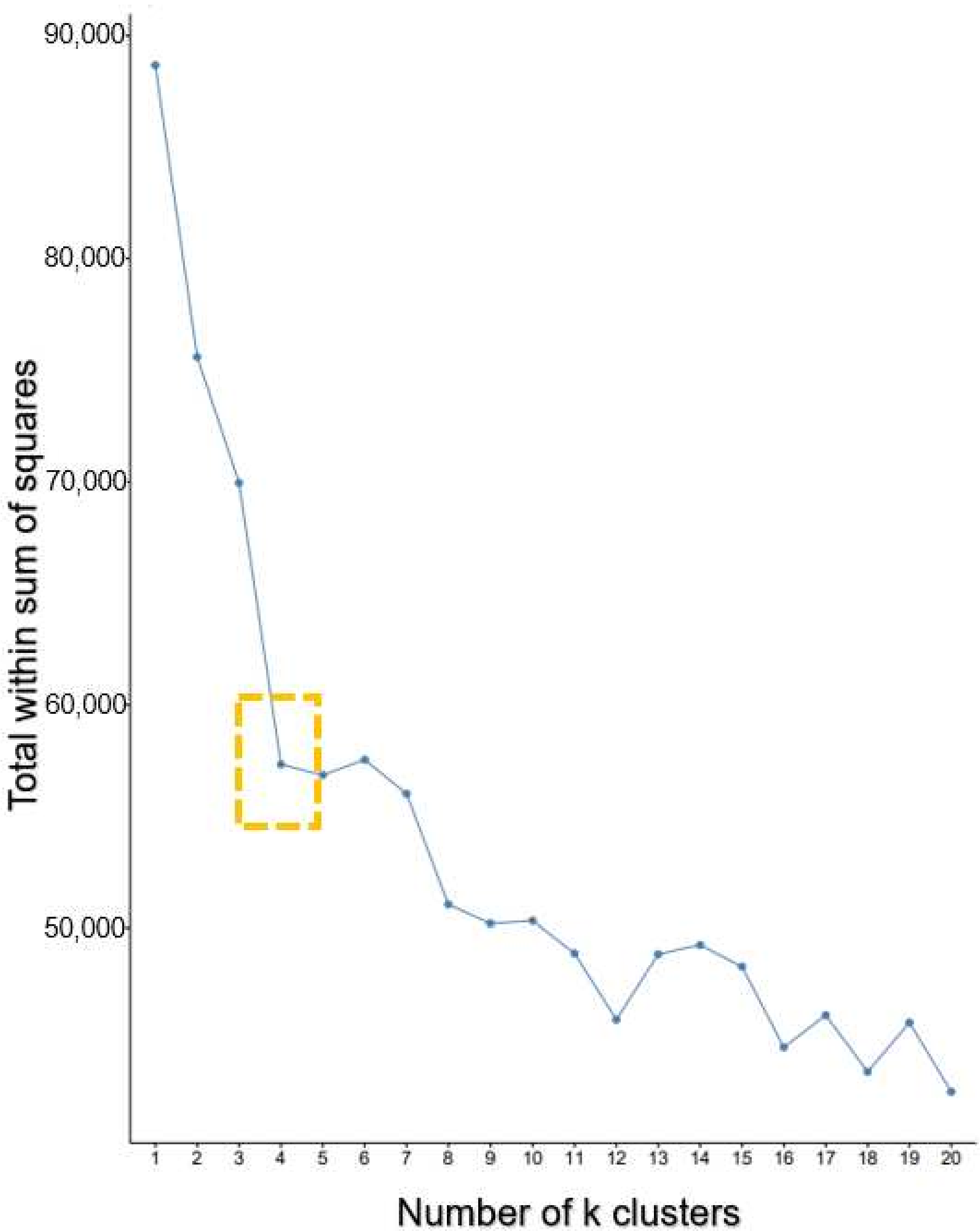

Figure 34: The within-cluster sum of squares method determined that four clusters are the proper number of clusters to be generated

The silhouette method is like the WSS method in that it also generates many iterations of clusters and evaluates them for a proper k size. The average silhouette evaluation is performed by comparing the silhouette width of each cluster within an iteration to cluster widths of succeeding iterations having incrementing numbers of clusters. Overall, when many clusters are found within a small dimension, the width of the clusters (silhouettes) are smaller than if one cluster was occupying the same space. Thus, when having many small clusters in a dimension that could be optimized by having fewer clusters, the average silhouette method will indicate a small average cluster width and an improper number of k clusters.

Additionally, if clusters are generated as tightly grouped neighbors, then one observation in one cluster will be very close in distance to an observation in a neighboring cluster. The closeness of observations belonging to different clusters can indicate that the model suffers too many clusters. A quality number of clusters to generate would be the number of clusters that optimizes the largest average silhouette width. Ribeiro et al. [26] utilized maximum silhouette scores in their graph-clustering algorithm to identify groups of terms and their semantics. Their method, and the inclusion of silhouette scoring, outperformed previous methods. In our research, the silhouette method suggested the proper number of clusters for our dataset of IoT tweets to be five (as shown in Fig. 3), whereas the WSS method suggested the proper number of clusters to be four. To ease the execution of algorithms, we utilized R packages factoextra and NbClust.

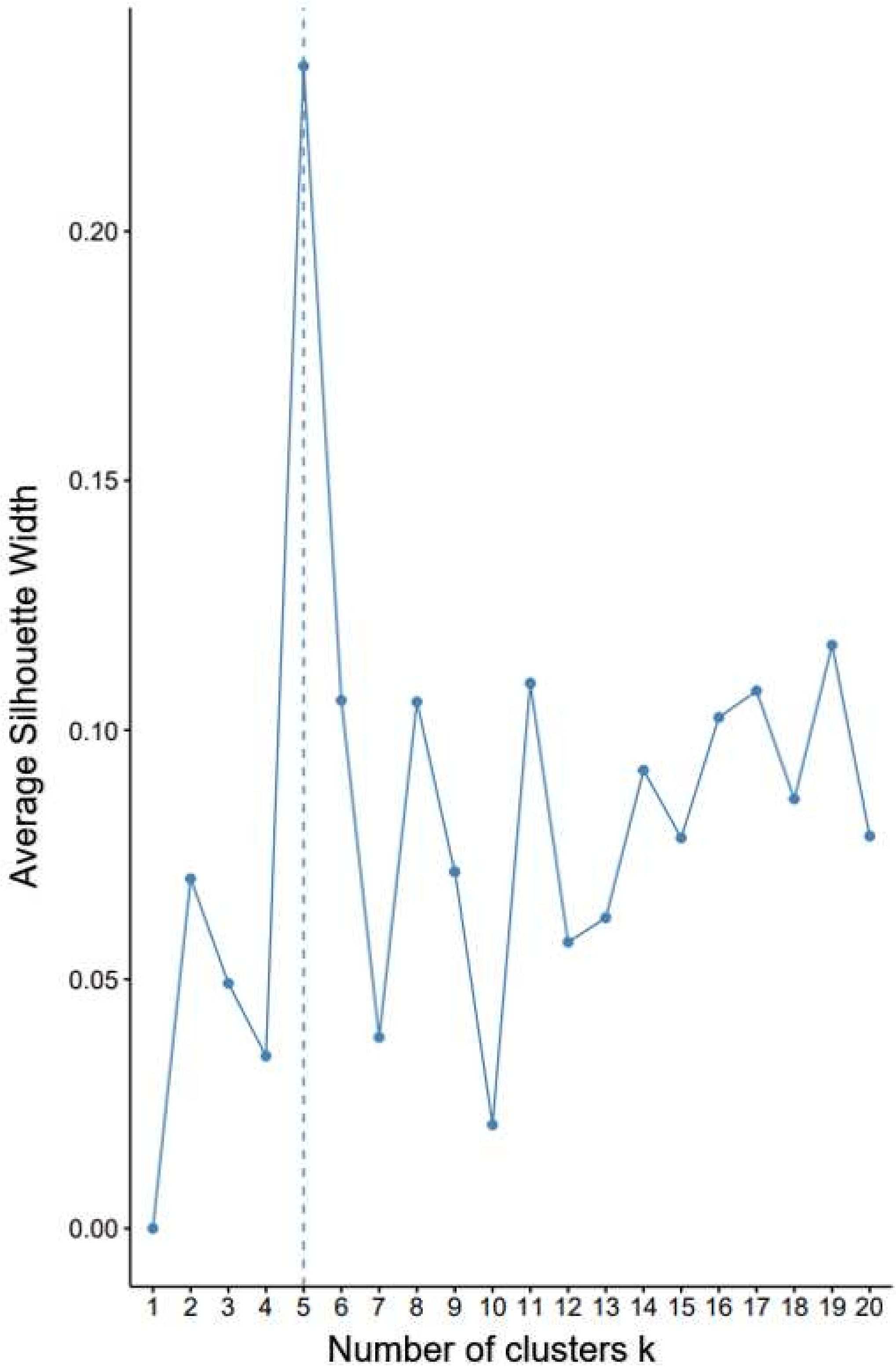

Figure 35: The silhouette method determined five clusters to be the proper number of clusters to be generated

Unsupervised hierarchical clustering was performed for both four and five cluster outcomes. An agglomerative method was used. With agglomerative clustering, each observation initializes as a single cluster and through iterations is joined with nodes being the shortest distance away [27]. The difference of trend identification between the different cluster generations, whether four or five clusters, was not found to be interesting. This is further illustrated within Fig. 4 where the largest clusters of tweets were cast into word clouds. It is seen that the leading terms are still quite similar despite the differing number of clusters generated. What was most concerning, whether four or five clusters were generated, was the lack of any cybersecurity topic as a trending top ten topic. Only 12% of the 684,503 tweets contained any term related to vulnerabilities, hacking, malware, and other cybersecurity-related terms.

The tweets were labeled for having inclusion to industry, trend, and commercial vendor technologies. To determine industry names and search terms, we utilized a list by the International Labor Organization [28]. The tweets were also evaluated for their sentiment by utilizing the NRC lexicon [29]. Our analysis will be further discussed in the following section.

Findings and discussion 5.4

Unsupervised hierarchical clustering and top trends 5.4.1

Because the WSS and average silhouette methods identified the proper number of clusters for our dataset as four and five respectively, we generated clusters of tweets for both findings. However, the leading trends identified did not vary between four and five clusters as illustrated in the word clouds below. Word clouds are a basic and intuitive data visualization technique that allows us to view terms by characteristics such as frequency [30].

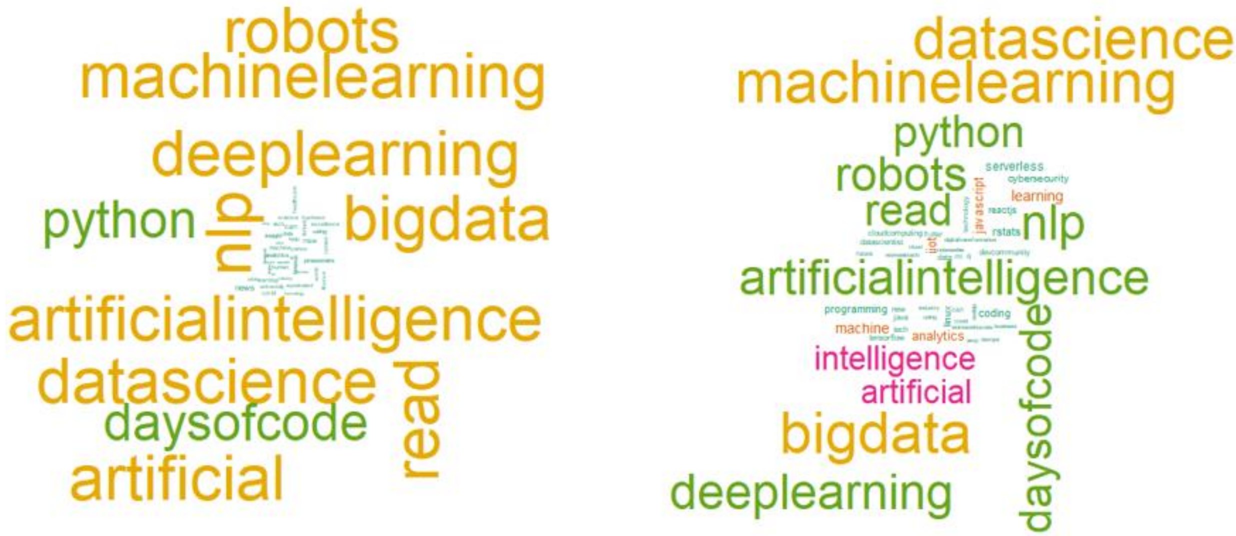

Figure 36: The leading trends do not include cybersecurity terms nor greatly shift whether four or five clusters of our IoT tweets were generated as indicated by word clouds of the largest clusters

The word cloud on the left is the largest cluster when only four clusters were generated. The word cloud on the right is the largest cluster when five were generated. We performed a similar analysis of trend analysis throughout the cluster creation and the leading identified trends did not alter. Regardless of the number of clusters created, the top mentioned term continued to be “data science”. It was closely followed by “machine learning”, and subsequent frequent terms began dropping off in mention at a greater pace than compared to the first and second most mentioned terms. The content-based analysis of trending topics is illustrated in Figure 37.

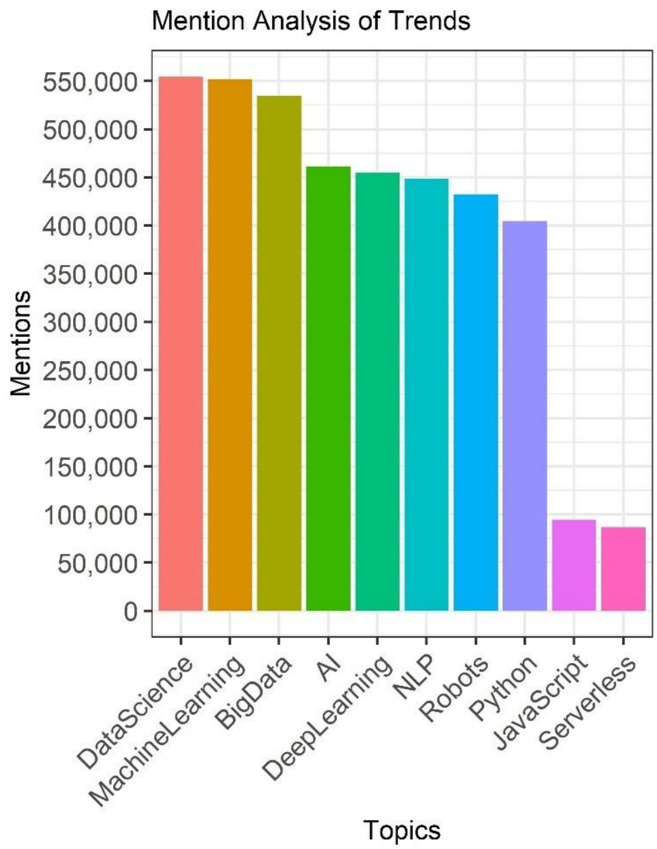

Figure 37: Term frequency is highest for data science, followed closely by machine learning

A Small number of cybersecurity mentions within the IoT tweets 5.4.2

Among the trend analysis, in general, what was most concerning was the lack of cybersecurity topics in the list of top mentioned terms. As illustrated in the following pie chart, only 12% of the 684,503 tweets had any mention of the following stemmed cybersecurity-related terms: cyber, secure, hack, vulnerability, risk, exploit, breach, malware, virus, ransomware, spyware, worm, trojan, encrypt or phishing.

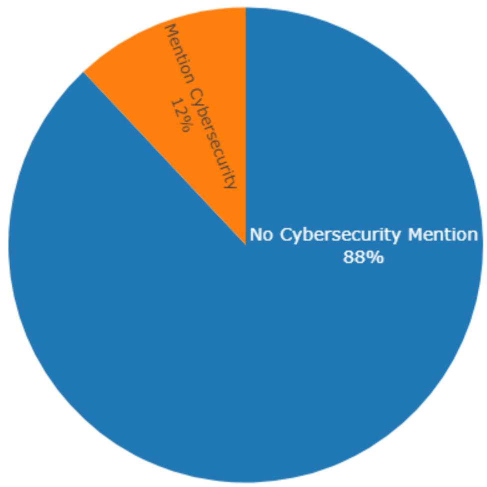

Figure 38: Only 12% of the total collection of IoT tweets had mention of common cybersecurity terms

When tweets did mention cybersecurity terms, the topics of the three most retweeted conversations included an industry roundtable discussion [32], a reference to an opinion article about the risk of AI on military technology [33], and a reference to an article on the risk of AI on national security [34]. Among the most retweeted tweets discussing cybersecurity, the top three are each a technology being touted to secure IoT implementations.

An IoT cybersecurity communications scorecard 5.4.3

The absence of frequent cybersecurity discussion within the collection of IoT tweets motivated examining which industries are communicating about risks the most. To compare the cybersecurity posture of industries based upon the public discussion found within our collection of tweets, we propose a new IoT Cybersecurity Communication Scorecard. The Balanced Scorecard was introduced by Kaplan and Norton in 1992 and evaluates broad performance measurements in four areas [35]:

- customer perspective
- internal perspective
- innovation and learning perspective

- financial perspective

The purpose of a balanced scorecard is to align the organization's resources and activities to its strategy for human resources, information, culture, leadership, and teamwork [36]. A good cybersecurity scorecard helps improve information and communication regarding cybersecurity [37]. Organizations have cybersecurity goals to be within compliance, protect their business, and maintain their employees' and customers' trust. Cybersecurity is not just about technology and systems, but also the people and processes within the systems [38]. Our scorecard allows organizations to compare their communication of cybersecurity knowledge, awareness, and training to a benchmark of public discussion within their industry.

Our IoT Cybersecurity Communication Scorecard assesses posture by comparing the z-scores of density mention and sentiment scores to the relative averages of all collected tweets. Mention density is the percentage of all IoT tweets that mention cybersecurity topics. The mention density and sentiment are each normalized by mean and standard deviations into Z-scores. The z-scores reflect an industry's position in terms of their cybersecurity mention density and the average sentiment of all tweets that reference their industry. The z-scores are found by first determining the average percentage of cybersecurity conversations among all tweets and the average sentiment of all tweets. The standard deviations are also recorded. The z-scores identify the positive and negative distance to the population's mean and are stated in Table 11. The posture score is a combined index of the two z-scores. We gave equal weight in the overall posture score calculation. If an organization placed significant importance on either the volume or the sentiment of the messages, they could apply custom weights.

Organizations should utilize the scorecard as a benchmark to compare their cybersecurity communication volume and sentiment to their industry's scores. For an organization to utilize this scorecard as a benchmark, they must determine their mention density by dividing the number of corporate cybersecurity communications by the total number of corporate communications and compare to their industry within Table 11. A similar comparison can be done to understand the positivity and sentiment of their corporate cybersecurity communications.

The leading industry by posture score within this social media analysis was found to be mechanical. Tweets within the food industry scored the lowest posture. The food sector experiences pressures such as climate change, food price volatility, and food security [29]. We must add cybersecurity risk to this list. Recently JBS USA Holdings, a food manufacturer which supplies the United States with roughly one-fifth of their meat supplies, experienced an expensive and business impacting ransomware attack [39]. Due to the ransomware attack, JBS USA Holdings temporarily shut down operations in nine meat processing plants and eventually paid a ransom of $11 million [40]. Table 11 provides the density of cybersecurity messages and their sentiment by industry. The table is sorted by posture rank. The scorecard research is limited by only comparing the top ten industries by volume.

Table 11: Industry cybersecurity scorecard by mention density and sentiment analysis

Industry	Mention Density	Density Z-score	Sentiment Score	Sentiment Z-score	Posture Score
Mechanical	40.7%	2.187	0.041	-0.033	2.154
Automotive	31.1%	1.384	0.024	-0.194	1.19
Commerce	12.1%	-0.198	0.119	0.699	0.501
Public	12.6%	-0.157	0.077	0.31	0.153
Health	10.3%	-0.353	0.093	0.461	0.108
Financial	5.4%	-0.76	0.128	0.79	0.03
Media	2.3%	-1.018	0.155	1.044	0.026
Transportation	7.6%	-0.519	0.057	0.116	-0.403
Agriculture	12.9%	-0.13	-0.038	-0.783	-0.913
Food	9.9%	-0.38	-0.21	-2.412	-2.792

Content-based analysis of industries within the IoT tweets 5.4.4

What is further concerning by the dearth of cybersecurity-related discussions within the collection of IoT-related tweets is that the top mentioned industry was healthcare. Previous research identified healthcare as one of the lesser influential industries mentioned in research papers on IoT [41]. Our research and this paper are one effort in shifting that claim. The top ten mentioned industries are depicted in Fig. 38.

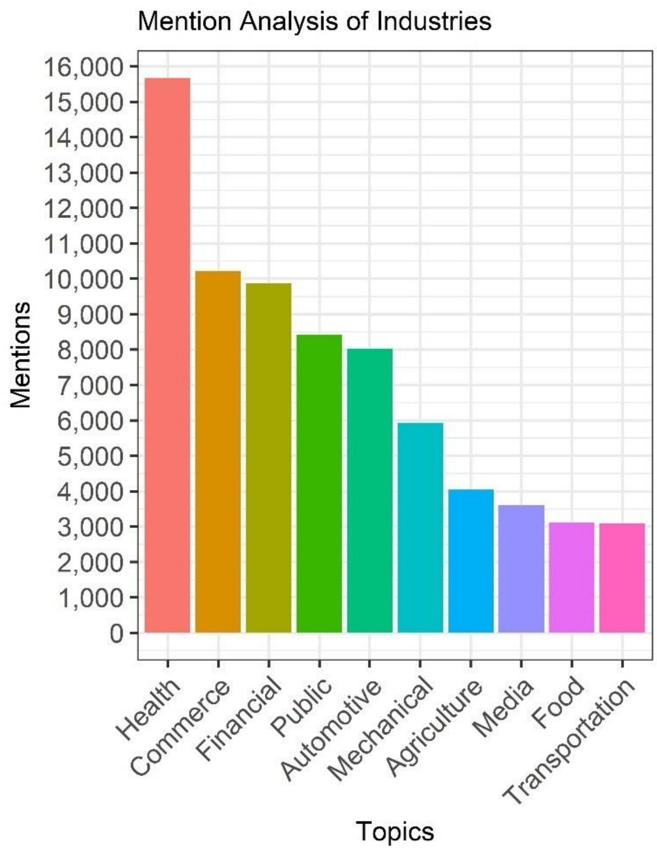

Figure 39: The top ten mentioned industry within the collection of IoT tweets was healthcare followed by commerce and then financial

It is not surprising to see healthcare leading the mentions as many countries are still experiencing the Covid-19 pandemic. While collecting these tweets based upon the inclusion of #iot, 4% of the tweets referenced Covid-19. Recent research has discussed the relationship between digital twins, IoT, and contact tracing technology [42], which could be utilized to help understand the behavior of a pandemic. After healthcare, the second most mentioned industry within the IoT tweets is commerce followed by financial.

Network analysis and relationship identification 5.4.5

A network analysis was also performed on the relationships between trends and industries. Fundamental parameters of a network are its number of nodes and the edges or connecting relationships between them [43]. We are surrounded by naturally occurring networks [44]. Industries and technology trends are no different, as we confirm with this analysis regarding the health industry connections to all the top identified IoT trends

To construct the network graph depicting relationships found in the social media conversations, the tweets' metadata labels were cast as nodes into two tables. The first table listed every industry and the trend terms (nodes) along with a unique identifier. The second table was a large list of the industry nodes, a corresponding trend node, and a weight column that indicated the frequency when a tweet was identified as matching both labels. Utilizing the network and igraph libraries in R, we plotted the node and edge relationships as the data visualization in Figure 40. This figure is a network graph that has the most mentioned industry, healthcare, highlighted as a green network node. Then, red lines which indicate relationships, are drawn to each of the yellow trending terms given both labels co-exist in single tweet metadata that we created during our preprocessing. As the image indicates, all trend terms are found in the network of healthcare tweets. As Figure 37 indicated, serverless was the least mentioned trending term, yet it too has an inner-tweet relationship to those tweets having reference to healthcare.

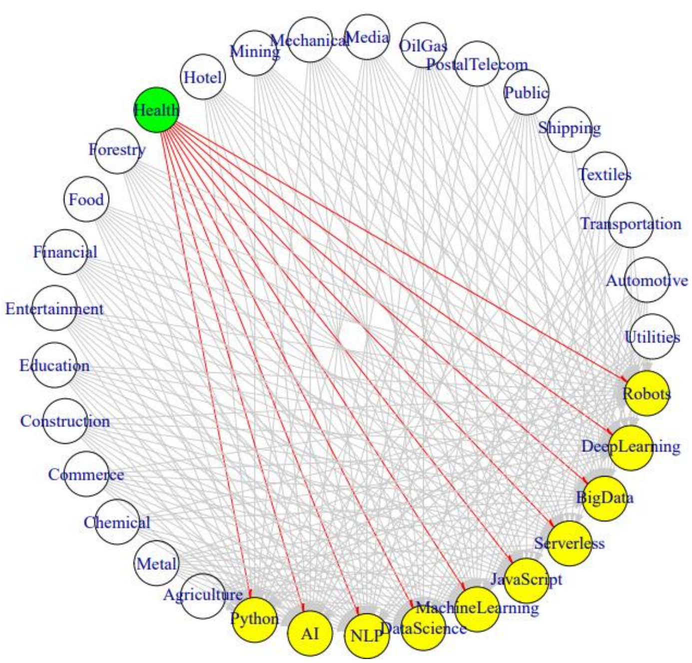

Figure 40: This network graph utilizes red arrows that depict relationships between tweets towards the healthcare industry, highlighted in green near the top of the image, and all the trending terms which are lighted in yellow

Sentiment analysis of commercial technology providers within the IoT tweets 5.4.6

There are many technology providers which have solutions, offer services, or offer platforms to solve IoT opportunities. We performed a content-based analysis of technology vendors within the IoT space. To determine the list of IoT vendors to analyze, we utilized two 2020 research reports by Gartner [45-46]. We utilized the sentimentr library to determine the sentiment scores of industry technology providers.

We plotted the technology provider names into a chart having four sections. The four sections of the chart have an x and y-axis, where the x-axis is the z-score of the tweet sentiments when the vendor is mentioned. The z-score is found by first determining the sentiment of all tweets that mention the commercial technologies, then calculating the average, and the standard deviation. Then, the z-score for a given technology vendor is calculated by dividing the commercial vendor's mentioned tweet sentiment by the number of standard deviations away from the population's average sentiment. The y-axis is measuring the number of times an IoT technology provider is mentioned in our corpus of tweets.

In general, if a vendor is placed on the upper right area of the chart, that implies that they are widely mentioned and the sentiment of the tweets that they are mentioned within is above average sentiment. If a vendor is found on the bottom left side of the chart, they would be both lower in popularity and lower in sentiment positivity within this collection of tweets. Any vendors having less than ten mentions within the tweets were removed from the plot. The dashed blue lines represent the average mentions and average sentiment scores. The average sentiment of all tweets mentioning these IoT solution vendors is slightly positive. Use caution when reviewing the chart as the y-axis is intentionally logarithmic. The logarithmic axis allows the data to pull slightly apart, as though zooming in, for the vendors who have lesser mentions. The vendor placement can be viewed in Figure 41.

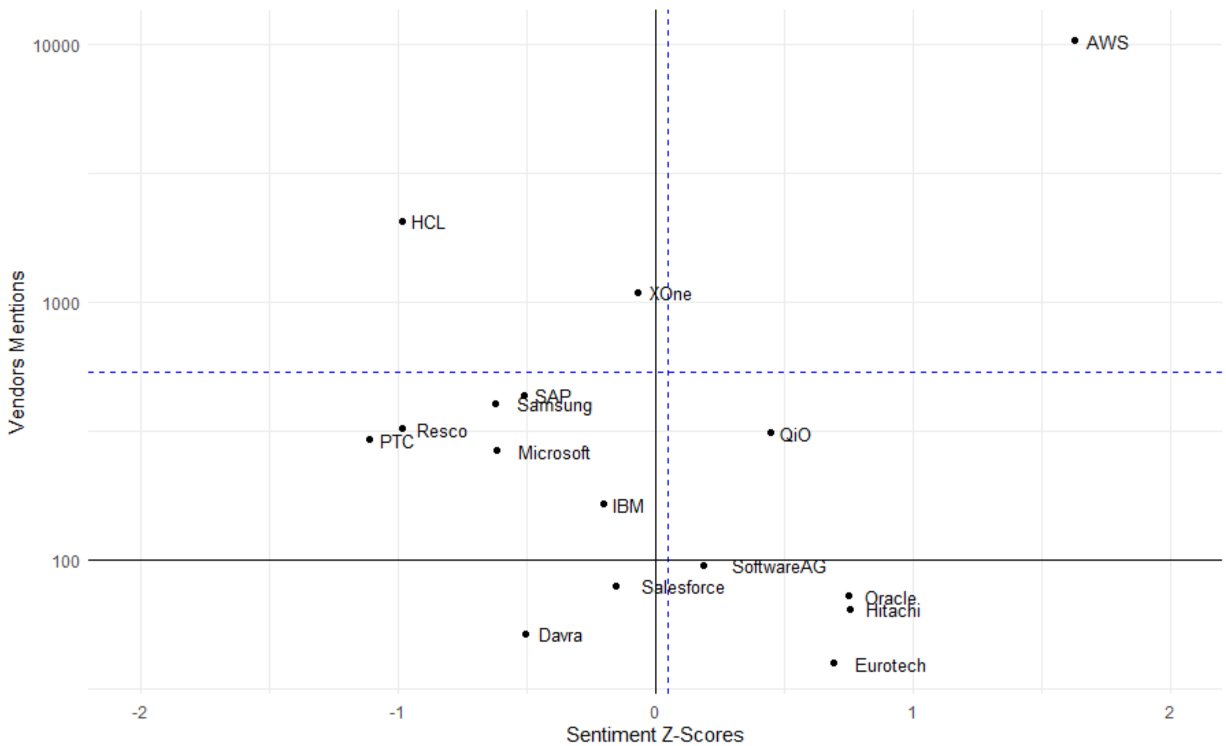

Figure 41: AWS has the most mentions and the highest sentiment among our corpus of IoT tweets while the technology company Davra would have a lesser number of mentions and a sentiment less than average

Amazon's AWS has the most mentions and the most positive sentiment among the vendors being mentioned within the IoT tweets. The AWS IoT Core can connect IoT devices to AWS cloud services and AWS offers an IoT SDK for development in languages such as Java, JavaScript, or Python. JavaScript was identified as one of the top ten trends in our analysis. AWS IoT Core product supports message brokering for these protocols [47]:

- Message Queuing and Telemetry Transport (MQTT)
- MQTT over Websockets Secure (WSS)
- Hypertext Transfer Protocol -Secure (HTTPS)
- Long Range Wide Area Network (LoRaWan)

Davra is within the bottom left area of the plot. They have fewer mentions in the analysis and the tweets that do mention them tend to have a lower sentiment than average across all the analyzed technology vendors. Davra offers an IoT Platform with features for access control of devices and services, service management features for edge, cloud, Kubernetes, and container deployments, as well as supporting many different IoT device protocols [48].

Predictive modeling based upon our IoT tweet metadata factors 5.4.7

Naïve Bayes has been utilized to accurately forecast crime activities [49]. Biology researchers have successfully applied naïve Bayes modeling to determine the presence of links in protein interaction networks [50]. In our research, we utilize naïve Bayes models to understand relationships between the IoT trends, the sentiment of the content, industries, and IoT technology providers.

Using a naïve Bayes model with a dependent factor of trend type and an independent variable of sentiment, we found that given a tweet is labeled as towards the trending topic data science, there is a 66.7% probability that the sentiment of the tweet is positive. Tweets that were labeled as towards the IoT trend of natural language processing (NLP) scored the second highest in positive sentiment probability at 57.1%. The table below notates the conditional probabilities as found by the model.

Table 12: Trending IoT tweet topics having the highest probability of positive sentiment are highlighted in this conditional probability table

Trends (below)	**Anger**	**Anticipation**	**Disgust**	**Fear**	**Joy**	**Negative**	**Positive**	**Sadness**	**Surprise**	**Trust**
AI	0.000	0.214	0.107	0.071	0.000	0.107	0.357	0.036	0.000	0.107
BigData	0.149	0.064	0.000	0.064	0.128	0.106	0.234	0.000	0.064	0.191
DataScience	0.000	0.222	0.000	0.000	0.000	0.000	0.667	0.000	0.111	0.000
DeepLearning	0.100	0.100	0.000	0.000	0.100	0.200	0.300	0.000	0.100	0.100
JavaScript	0.000	0.222	0.000	0.000	0.111	0.000	0.444	0.000	0.000	0.222
MachineLearning	0.045	0.136	0.000	0.091	0.000	0.136	0.455	0.000	0.000	0.136
NLP	0.000	0.143	0.143	0.000	0.000	0.143	0.571	0.000	0.000	0.000
Python	0.000	0.000	0.000	0.500	0.000	0.500	0.000	0.000	0.000	0.000

A second naïve Bayes model was created to help with understanding which factors influence the prediction of tweets being retweeted. The industry and trend factors had little impact on a tweet being retweeted. However, the sentiment did affect the probability of a tweet being retweeted. Given the tweets conveyed either fear or joy would improve the probability of retweet to 13.0% and 12.4% respectively. A third naïve Bayes model was used to predict which trending term an IoT tweet may be about. Using the factors of the industry type, retweet, IoT vendor name, and whether the tweet had been favored, trend prediction was achieved with an accuracy of 63.9%.

Conclusion 5

There are new microblogs on the topic of the Internet of Things each day. From May 1st, 2021, to May 14th, 2021, we collected 684,503 tweets by searching Twitter's API for #iot. While previous research has indicated that healthcare is not a top-three industry influence on the IoT [41], our research determined healthcare the most widely discussed industry in public IoT conversations on the Twitter platform. While the healthcare industry requires secured information systems, only 12% of the tweets within this IoT network analysis referenced cybersecurity concepts. Even less, only 10.3% of the healthcare-related tweets referenced cybersecurity concepts.

From this collection of tweets, the most common trend term was data science. A network analysis graph depicted that every trending term was mentioned within healthcare-related tweets. Whereas for the tweets regarding the shipping industry, only the trends of AI, big data, and machine learning were related. IoT practitioners should utilize network analysis to see how similar organizations are communicating and including technical concepts in their implementation.

No cybersecurity-related terms or concepts, such as encryption, ransomware, zero-trust, or vulnerabilities, were identified as trending terms. In general, there was an alarming dearth of conversations towards cybersecurity as only 12% of the IoT tweets contained any mention of cybersecurity-related topics.

The trending terms having the highest probability of positive sentiment in a referencing tweet were data science followed by natural language processing. We could predict what trend a tweet was referencing with a 63.9% accuracy. To reach that level of accuracy in the model we utilized the factors of whether the tweet had been retweeted, marked as a favorite, and by knowing the industry and vendors being mentioned in the tweet's text. IoT practitioners need to review our identified trends for how these technologies can benefit their implementations and end-users. Future research should include a comparison of the trends we have identified and how they may change over time.

A new IoT Cybersecurity Communication Scorecard was proposed. The posture was scored by the density of cybersecurity conversations and their sentiment. The top ten mentioned industries were ranked by their posture using our IoT cybersecurity communication scorecard. The mechanical industry had the highest-rated posture. The scorecard is limited in that it only ranks based on communication regarding cybersecurity and future research is required to tie the posture score into the many additional areas of securing systems. IT security leaders should utilize this scorecard to benchmark their cybersecurity communication density and sentiment compared to the public discussions referring to their industry.

Amazon AWS had the highest average sentiment among the vendors that were considered in this research. It was also Amazon AWS that was most frequently mentioned in this collection of tweets. Commercial firms can utilize our research and Figure 41 to assess competing organizations and improve their social media presence and marketing messages.

A limitation of this research is that only one microblogging site, Twitter, was utilized for data collection. Another limitation was the available computing power of our systems. Our experience is that 32 GB of memory is not sufficient when analyzing 684,503 tweets and thus forces the use of samples within the collection. Specifically, we turned to use samples when carrying out the unsupervised hierarchical clustering and the naive Bayes models within our methodology. There is a need to study time-dependent trends that will examine if the communication regarding cybersecurity is increasing towards acceptable values. Such research will require periodic data collection for a period spanning several months or a few years.

REFERENCES

[1] Girma, A. (2018). Analysis of security vulnerability and analytics of internet of things (IoT) platform. *In Information Technology - New Generations. Advances in Intelligent Systems and Computing*, 738, https://doi:10.1007/978-3-319-77028-4_16

[2] Ashton, K. (2009). That internet of things thing. *RFiD J.*, 22, 97–114.

[3] Shi, W., Cao, J., Zhang, Q., Li, Y., & Xu, L. (2016). Edge computing: vision and challenges. *IEEE JIOT,* 3, 637-646. https://doi:10.1109/JIOT.2016.2579198

[4] James, J. (2014). Data never sleeps 2.0. Retrieved from https://www.domo.com/blog/data-never-sleeps-2-0/

[5] Cruickshank, I.J., & Carley, K.M. (2020). Characterizing communities of hashtag usage on twitter during the 2020 COVID-19 pandemic by multi-view clustering. *Appl Netw Sci.* https://doi.org/10.1007/s41109-020-00317-8

[6] Guarino, S., Trino, N., Celestini, A. et al. (2020). Characterizing networks of propaganda on twitter: a case study. *Appl Netw Sci.* https://doi.org/10.1007/s41109-020-00286-y

[7] Tien, J.H., Eisenberg, M.C., Cherng, S.T. et al. (2020). Online reactions to the 2017 'Unite the right' rally in Charlottesville: measuring polarization in Twitter networks using media followership. *Appl Netw Sci.* https://doi.org/10.1007/s41109-019-0223-3

[8] Gómez-García, M., Matosas-López, L., & Ruiz-Palmero, J. (2020). Social networks use patterns among university youth: the validity and reliability of an updated measurement instrument. *Sustainability*. https://doi:10.3390/su12093503

[9] Bougie, G., Starke, J., Storey, M., & German, D.M. (2011). Towards understanding twitter use in software engineering: preliminary findings, ongoing challenges and future questions. *In: Web2SE '11. ACM*, 31–36. https://doi:10.1145/1984701.1984707

[10] Williams, A. (2018). Do software engineering practitioners cite research on software testing in their online articles? a preliminary survey. *In: Proceedings of the 22nd International*

Conference on Evaluation and Assessment in Software Engineering, ACM, 151–156. https://doi:10.1145/3210459.3210475

[11] Storey, M. et al. (2010). The impact of social media on software engineering practices and tools. *In: Proceedings of the FSE/SDP workshop on Future of software engineering research. ACM.* 359–364.

[12] Büchi, M. (2017). Microblogging as an extension of science reporting. *Public Understanding of Science,* 953–968. https://doi:10.1177/0963662516657794

[13] D'heer, E., & Verdegem, P. (2015). What social media data mean for audience studies: a multidimensional investigation of twitter use during a current affairs tv programme. *Information, Communication & Society*, 18, 221-234. https://doi:10.1080/1369118X.2014.952318

[14] Martín, Y., Cutter, S., & Li, Z. (2020). Bridging twitter and survey data for evacuation assessment of hurricane matthew and hurricane irma. *Natural Hazards Review*, https://doi.org/10.1061/(ASCE)NH.1527-6996.0000354

[15] Atalay M, & Angin P. (2020). A Digital twins approach to smart grid security testing and standardization. *In: IEEE International Workshop on Metrology for Industry 4.0 & IoT*, 435-440. https://doi:10.1109/MetroInd4.0IoT48571.2020.9138264.

[16] Langner, R. (2011). Stuxnet: dissecting a cyberwarfare weapon. *In IEEE Security & Privacy*, 49-51. https://doi:10.1109/MSP.2011.67.

[17] Turton, W., Riley, M., & Jacobs, J. (2021). Colonial pipeline paid hackers nearly $5 million in ransom. Retrieved May 14, 2021, from https://www.bloomberg.com/news/articles/2021-05-13/colonial-pipeline-paid-hackers-nearly-5-million-in-ransom

[18] Ikeda, S. (2021). Utility regulator says solarwinds backdoor was downloaded by 1/4 of electric utilities on the north american power grid. *In: CPO Magazine*. Retrieved Jul 9, 2021, from https://www.cpomagazine.com/cyber-security/utility-regulator-says-solarwinds-backdoor-was-downloaded-by-1-4-of-electric-utilities-on-the-north-american-power-grid/

[19] Srinivas, R. (2020). 10 IoT security incidents that make you feel less secure. CISO Magazine. Jan. 10, 2020. Available: https://cisomag.eccouncil.org/10-iot-security-incidents-that-make-you-feel-less-secure/

[20] Chen, C.W. (2021). Drones as internet of video things front-end sensors: challenges and opportunities. *Discov Internet Things*. https://doi.org/10.1007/s43926-021-00014-7

[21] Hearn, M, & Rix, S. (2019). Cybersecurity considerations for digital twin implementations report. *Industrial Internet Consortium. Journal of Innovation*. 2019. https://www.iiconsortium.org/news/joi-articles/2019-November-JoI-Cybersecurity-Considerations-for-Digital-Twin-Implementations.pdf

[22] Kim, H., Ahmad, A., Hwang, J., Baqa, H., Le Gall, F., Ortega, M.A.R., & Song, J. (2018). IoT-TaaS: towards a prospective IoT testing framework. *In IEEE Access.* 15480-15493. https://doi:10.1109/ACCESS.2018.2802489

[23] Fizza, K., Banerjee, A., Mitra, K. et al. (2021). QoE in IoT: a vision, survey, and future directions. *Discov Internet Things*. https://doi.org/10.1007/s43926-021-00006-7

[24] Motta, R., de Oliveira, K., & Travassos, G. (2018). On challenges in engineering IoT software systems. *In: Proceedings of the XXXII Brazilian Symposium on Software Engineering. ACM*, 42–51. https://doi:10.1145/3266237.3266263

[25] Grieves, M., & Vickers, J. (2017). Digital twin: mitigating unpredictable, undesirable emergent behavior in complex systems. *Springer International Publishing, Cham*. 85–113.

[26] Scheibmeir, J. (2021) Code and data files. Gitlab. https://gitlab.com/jimscheibmeir/socialmediaanalyticsofiot

[27] Ribeiro, E., Teixeira, A., Ribeiro, R. et al. (2020). Semantic frame induction through the detection of communities of verbs and their arguments. *Appl Netw Sci.* https://doi.org/10.1007/s41109-020-00312-z

[28] Hastie, T., Tibshirani, R., & Friedman, J. (2009). The elements of statistical learning: data mining, inference, and prediction. Springer.

[29] International Labor Organization. Industries and sectors. Accessed: May 16, 2021. Available: https://www.ilo.org/global/industries-and-sectors/lang--en/index.htm

[30] Mohammad, S., & Turney, P. (2010) Emotions evoked by common words and phrases: using mechanical turk to create an emotion lexicon. *In: Proceedings of the NAACL-HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*.

[31] Altuncu, M., Mayer, E., Yaliraki, S. et al. (2019). From free text to clusters of content in health records: an unsupervised graph partitioning approach. *Appl Netw Sci.* https://doi.org/10.1007/s41109-018-0109-9

[32] [@#100DaysOfCode]. Sponsored Roundtable: Artificial Intelligence Tech – SecurityInfoWatch. May 14, 2021. Twitter.

[33] [@HubOfML]. Opinion: Artificial Intelligence's Military Risks, Potential – GovTech May 14, 2021. Twitter.

[34] [@#100DaysOfCode]. Artificial intelligence and the future of national security - ASU Now May 12, 2021. Twitter.

[35] Kaplan, R. S., & Norton, D. P. (1992). The balanced scorecard-measures that drive performance. *Harvard Business Review*, 70, 71–79.

[36] Herath, T., Herath, H., & Bremser, W.G. (2010). Balanced scorecard implementation of security strategies: a framework for IT security performance management. *Information systems management,* 27(1), 72–81.

[37] Whalen D.T., & Bell, G. (2014). Cyber security risk: key areas of focus. *Directors and Boards,* 38(3), 14–.

[38] Bowen, B. M., Devarajan R., & Stolfo, S. (2011). Measuring the human factor of cyber security. *In Proceedings of the 2011 IEEE International Conference on Technologies for Homeland Security*, 230-235. https://doi:10.1109/THS.2011.6107876

[39] Bunge, J. (2021). WSJ News Exclusive | JBS Paid $11 Million to Resolve Ransomware Attack. In: The Wall Street Journal. https://www.wsj.com/articles/jbs-paid-11-million-to-resolve-ransomware-attack-11623280781. Accessed 6 Jul 2021

[40] Lerman, R. 2021. JBS paid $11 million in ransom after hackers shut down meat plants. In: The Washington Post. https://www.washingtonpost.com/technology/2021/06/09/jbs-11-million-ransom/. Accessed 7 Jul 2021

[41] Dachyar, M., Zagloel, T., & Saragih, L.R. (2019). Knowledge growth and development: internet of things (IoT) research. *Heliyon*, 2019(5), 2006-2018. https://doi.org/10.1016/j.heliyon.2019.e02264.

[42] Scheibmeir, J., & Malaiya, Y.K. (2020). Multi-model security and social media analytics of the digital twin, *ASTEJ.,* 5, 323-330.

[43] Smith, K.M., & Escudero, J. (2020). Normalised degree variance. *Appl Netw Sci.*, 5; 32. https://doi.org/10.1007/s41109-020-00273-3

[44] Kolaczyk, ED, & Csárdi, G. (2014) *Statistical Analysis of Network Data with R*, vol. 65. Springer.

[45] Goodness, E., Velosa, A., Friedman, T., Berthelsen, E., Kim, S., Havart-Simkin, P., & Thielemann, K. (2020). Magic quadrant for industrial IoT platforms. *Gartner Inc*. Oct. 19, 2020, Accessed: May 16, 2021. Available: https://www.gartner.com/doc/3991952

[46] Wong, J., Leow, A., & Batchu, A. (2020). Magic quadrant for multiexperience development platforms. *Gartner Inc.*, July 7, 2020. Accessed: May 16, 2021. Available: https://www.gartner.com/doc/3987201

[47] AWS. What is AWS IoT? AWS. Accessed on: June 1, 2021, Available: https://docs.aws.amazon.com/iot/latest/developerguide/what-is-aws-iot.html

[48] Davra. IoT platform: internet of things platform - davra IoT platform. Davra, Accessed: June 1, 2021. Available: https://davra.com/iot-platform/

[49] Niu, X., Elsisy, A., Derzsy, N. et al. (2019). Dynamics of crime activities in the network of city community areas. *Appl Netw Sci.,* 4(127). https://doi.org/10.1007/s41109-019-0239-8

[50] Singh, K.V., & Vig, L. (2017). Improved prediction of missing protein interactome links via anomaly detection. *Appl Netw Sci.,* 2(2). https://doi.org/10.1007/s41109-017-0022-7

CHAPTER 6: SOCIAL MEDIA PERSPECTIVES ON DIGITAL TWIN AND ITS MATURITY MODEL

This chapter offers an analysis of the varying conversations of Digital Twins on social media, specifically the Twitter platform. Social media offers a platform for sharing information that can be analyzed to extract valuable information. Social media records can be analyzed to evaluate the velocity, volume, and variety of data related to a specific topic. Industry mentions, use cases, and sentiment of the associated topics and network graphs are introduced as well as supporting background information. The analysis reviews over 24,000 tweets collected between September of 2019 through July of 2021. We have identified the most mentioned industries with interest in Digital Twins. Among identified trending topics, the top three include the Internet of Things, artificial intelligence, and industrial uses. A maturity model for digital twins is introduced, informed by the identified trends and their popularity. The significance of the findings is discussed.

Defining Digital Twins 6.1

The digital twin concept has many definitions and contributing authors. Jones et al [1] attribute Michael Grieves, along with John Vickers, with the origination of the concept. The origin of the digital twin concept according to Rosen et al. [2], has roots in NASA's Apollo program, twinning a spacecraft for training and mission support purposes. The term "digital twin" was coined by Shafto et al in 2010 [3]. Grieves describes a digital twin as consisting of a physical asset, its virtual representation, and a two-way connection [4]. Eckhart & Ekelhart do not define digital twins as having control over their physical counterparts [5]. Rather they focus the definition and capability of the digital twin towards monitoring, visualization, and prediction. In their research of identifying definitions of the digital twin, Negria et al. [6] found that the digital twin's definition has varied and diverged away from solely modeling a physical system. Implementation definitions also range from digital twins being an augmented reality (AR) application [7] to machine learning models. The amount and timeliness of integration that is required for a virtual instance to be considered as a digital twin have not been agreed upon [5]. Eckhart & Ekelhart do not specify that a digital twin should secure the physical counterpart unless that is a part of the optimization [5]. Digital twins can aid in the security of the physical counterpart using different access models and malicious activity identification techniques [8]. Eckhart & Ekelhart have suggested characterizing the digital twin concept based upon the level of data flow, integration, and autonomy (identified in Table 13).

Table 13: Digital twin types defined by dataflow, integration, and autonomy adapted from Eckhart & Ekelhart [5]

Level of integration	Dataflow and Connectedness	
	Physical -> Digital	Digital -> Physical
Digital model	Manual	Manual
Digital shadow	Autonomous	Manual
Digital twin	Autonomous	Autonomous

The CIRP encyclopedia definition does not include a virtual to physical connection in its description [9]. Grieves is a commonly cited author and includes a bi-directional and virtual to physical connection in his definition [4]. Tao et al extend the original three-component model by Grieves, the physical, virtual, and bi-directional connection between them, to one having five dimensions. The five dimensions from Tao et al include the physical environment (PE), virtual environment (VE), services of both, the data of the digital twin, and the connection [10]. Equation 1 is an adaptation of expression by Tao et al [10].

Equation 1: The Digital Twin Model by Tao et al. has Five Dimensions [10]

$$Model_{DT} = \sum (PE, VE, Ss, DD, CN)$$

Haag et al. characterize digital twins as simulators [11]. A few other explanations and definitions have been given towards the digital twin such as assisting with the operational health of the physical system [12] and optimizing a process [13]. The most common discrepancies of the definitions are whether the digital twin has responsibility for controlling the physical counterpart or only simulates it and whether the digital twin should have a cybersecurity function.

Digital twins increase the digital touchpoints of a cyber-physical system (CPS) and offer hackers knowledge of system integrations [14]. Many digital twin integrations are with the Internet of Things (IoT) technology. IoT has improved the management of homes, businesses, industries, and public sectors [15]. The information security concerns of IoT range from authorization, authentication, privacy, and access control of embedded systems [15]. In general, IoT technology has produced a new cyber-attack surface [16].

A study on consumer IoT, smart speakers, identified enjoyment (34.24%) as the most influential reason for the adoption of the devices [17]. The second and third most influential factors were social norms (15.43%) and usefulness (11.86%) respectfully. While IoT consumer devices may offer a fun experience, social acceptance, and usability, they must also offer security in the solutions.

While enjoyment is a major contributing factor to IoT adoption among consumers, miniaturization, and cost reduction have attracted Industry 4.0 [11]. Industry 4.0 is the convergence of modern manufacturing and modern computing. Smart factories are building smart devices [11]. If, or when, a smart factory is exploited, the supply chain of smart devices may generate exponential security concerns.

To mitigate new threat vectors, a multi-model of security access controls can help the digital twins secure their physical counterparts [8]. Multiple security models within the digital twin act as filters that trap malicious behavior before the physical assets executing the instruction. Control instructions, current, and predicted future states can be compared across the physical and virtual systems. Discrepancies can imply an inaccurate digital twin or indicate malicious acts. However, codified rules and advanced analysis techniques within system operations will not be enough to deter and prevent all risks and exploits. Security must start with the organization's culture in a bottom-up approach (people, processes, and system inception to retirement) [14].

Social and cultural issues and complexities exist in the implementation of digital twins. Frequently this is related to the types of data being collected, stored, and exposed by the digital twin [1]. Digital twins have been used by consumable product designers to understand how product use differs across cultures and locations [18]. Even a local sports team's home game schedule can be a factor in modeling and predicting factory production [1].

Cybersecurity is not the only concern when implementing digital twins. Current standards and architectures for IoT do not solve their interoperability problems [19]. Organizations contributing to IoT standards include the World Wide Web Consortium (W3C), the Internet Engineering Task Force (IETF), and the Internet Research Task Force (IRTF), among others [19]. Progress in IoT standardization includes these example services and protocols [19]:

- A Thing Description is a file containing semantic metadata about an IoT thing including its properties and behaviors
- Resource Directories are repositories of things and their network identification
- Constrained Application Protocol (CoAP) offers device communication over UDP and other transports. A CoAP datagram is illustrated in Figure 42.

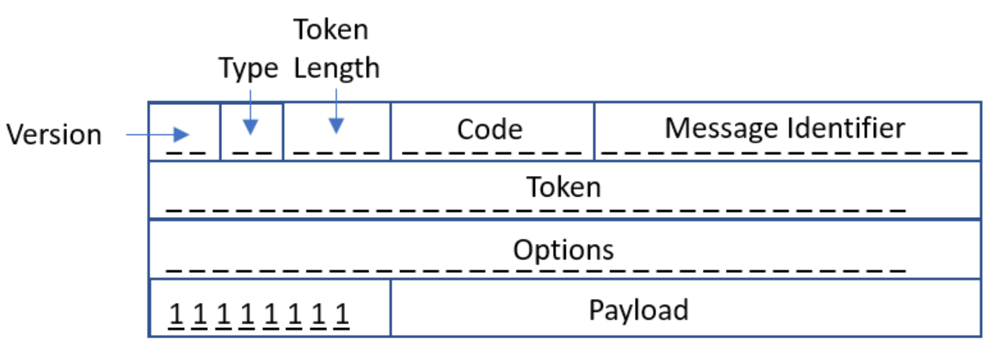

Figure 42: The message format of constrained application protocol

The first two bytes of a CoAP datagram indicate the version of the protocol. Version is followed by two bits indicating the type of the message. The type of message could include confirmable, which requires acknowledgment of receipt. Acknowledgment becomes another message type. The token length field indicates the length of the upcoming token field. The value of the token field is used to connect request messages to their responses. The message identification field can be used to identify duplicate messages as well as to match an acknowledgment to a confirmable message. After message options, a byte of all one's indicates the start of the message payload. CoAP messages use unreliable transports such as UDP but offer mitigating features such as retransmission of confirmable messages [20].

Standards and interoperability among devices are important because a digital twin may have a lifecycle many decades-long. During such a span of operations, many IoT devices that are informing the digital twin may be swapped in and out due to failure, enhancement, or upgrade. The lifecycle of a digital twin requires affordable and feasible interoperability of IoT devices. IoT devices should be reusable, discoverable, and adaptable. These attributes of IoT devices help a digital twin to become maintainable.

To evaluate performance and scalability, tools such as CoAPBench may be utilized to evaluate implementations. The CoAPBench employs virtual clients that simulate IoT device registrations. CoAPBench can scale many concurrent clients while measuring response times from the management layers of an IoT and digital twin system. For a digital twin to achieve the characteristic of fidelity, or sameness to its physical counterpart, many IoT devices will be integrated and informing the digital twin solution. Non-functional characteristics, such as the performance and maintainability of the system will be critical in the management of the digital twin over an extended lifespan. Characteristics such as reuse and discoverability of IoT endpoints will help accelerate the maintenance and enhancements of digital twins over their lifespan.

A development model and methodology for using APIs for digital twins have been put forward [21]. The development model begins with an objective tree and contextual diagram to cover the environment, relationships, and operations of the physical entity. The development of a digital twin must encompass the functionality of the physical counterpart, supporting and foundational data sources and integrations, as well as the context of the operating environment and culture. Using context diagrams and objectives trees are methods to explore and define the needs of a digital twin.

Test-driven development was suggested as a practice for implementing APIs in a test-first approach [21]. Utilizing OpenAPI specification aids design and test documentation and supports reuse. Traditional software development lifecycles place testing the system after its development. A better approach is to “shift left” and test during the design and development through practices such as Test-driven development and Behavior-driven development. These practices focus on the creation of unit tests and UI tests before any code being implemented. With these practices, testing comes before code development work, and thus “shifts left” in a traditional development cycle.

Performance engineering for digital twins must be done early, such as testing individual parts or components of the API operations. Testing for performance concerns early in the development helps avoid expensive redesign efforts. Figure 43 is from the 2019 work of Scheibmeir and Malaiya and illustrates the use of contextual diagrams, objective trees, TDD, and many more practices in the development of APIs for digital twins[21]. The model suggests API mediation but fails to extend into concerns for the user interface. Augmented reality has been suggested as an interface modality for digital twins [7].

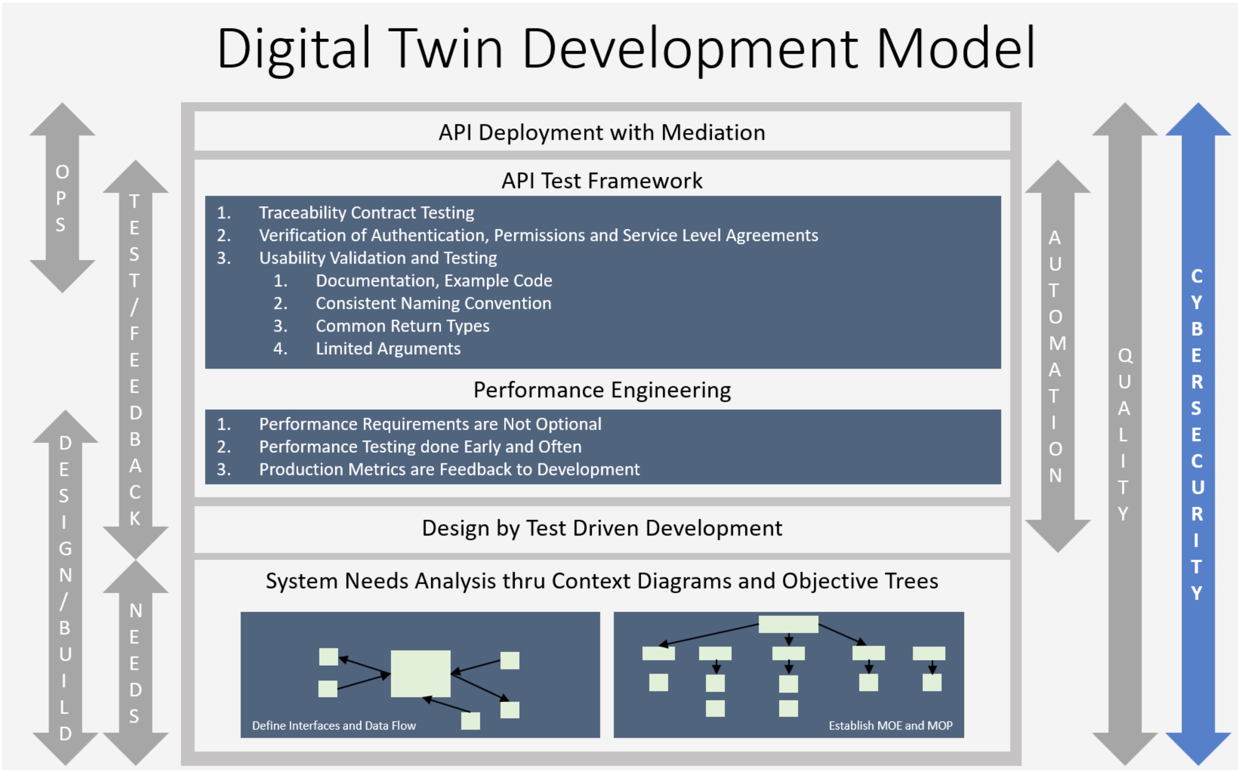

Figure 43: A framework for developing APIs for digital twins [21]

Use of Social Media Analytics in Research 6.2

Social media data is common to many research investigations. Social media data is publicly available, and offers velocity, variety, and volume of data. Researchers can extract valuable conclusions from social media due to its public nature and ease of access [22]. Twitter data has enhanced biased survey populations and assisted in research by supplying factors of longitude and latitude of participants [23]. A study by Bougie et al [24] found that 23% of tweets by the software engineering groups they followed were towards software engineering topics. Of that 23% of tweets, 62% were towards solving software engineering problems. Software engineering practitioners use social media platforms to learn about technology trends [25]. They do not cite scientific research in their blogs [26]. Beyond trend identification, social media platforms offer links to web resources, networking people, and directing our attention [27]. Searching for and accessing information are the leading factors among college students for accessing social media platforms [28].

Social Media Analytics Methodology 6.2.1

Utilizing R programs and the Twitter API, we have collected 24,275 tweets from August 2019 until July of 2021. This is not a comprehensive collection of all tweets referring to digital twins. Our collection of tweets is limited by unpaid access to Twitter's API and further constrained by daily limits and the R programs collecting tweets towards many different topics. While the analysis is limited, it informs on the public discourse about digital twins and our methodology will be discussed in enough detail to enable similar research for those who want to dig deeper in this area.

A content-based analysis is utilized within this research to determine themes among the tweets. Themes may include technology trends or industries where digital twin technology is frequently discussed. Time series analysis indicates ebbs and flows of the discussions and helps identify when peaks or lulls in the discussions are occurring. Sentiment analysis provides a numerical approach to how positive or negative the meaning of a tweet's language may be. Network graphs help identify relationships. This chapter will utilize network graphs to detect relationships between the industry discussions of digital twins and which technology trends are included and omitted from the discussion. When confronted with large amounts of free-form text, it may be useful to utilize clustering techniques to determine the distinct topics and conversations occurring. The cluster sizes are determined by the within-cluster sum of squares (WSS) and the average silhouette methods. A dendrogram is a data visualization object and a type of tree graphic. Dendrograms depict the closeness or sameness of objects after they have been clustered. These methods are useful when analyzing social media and other data sources and will be utilized throughout this chapter.

Time Series Analysis of Tweets about Digital Twins 6.2.2

Twitter supplies a created date field that identifies when Twitter users posted their communication. The earliest tweet within our collection was posted on August 29th, 2019. The last tweet in our collection is dated July 31st, 2021. Figure 44 is a time series chart identifying the date the tweets were posted to the Twitter platform and the number of tweets per day. A smoothed line is positioned along the time series to indicate the overall trend in the volume of tweets.

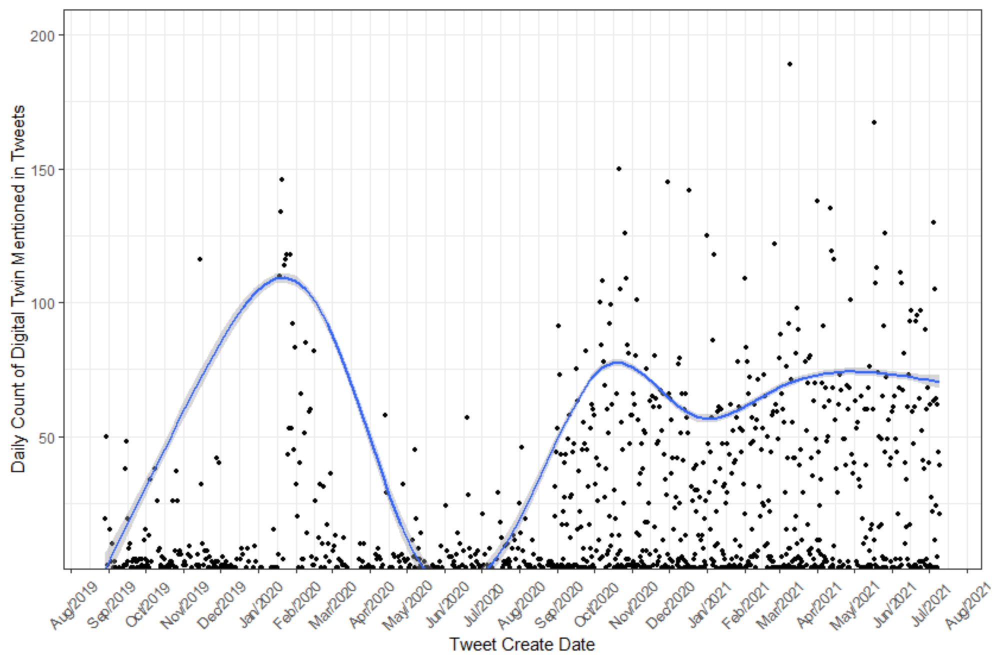

Figure 44: Time series chart of our collection of tweets referring to digital twins with a peak in January of 2020

The chart identifies a peak in the discussions of collected tweets during January 2020. To determine the trends driving up this peak we isolate by the posted date and identify tweets having the highest retweet counts. Retweets are a feature and behavior among Twitter users who can repost a tweet to propagate the message through their network. The tweet in Figure 45 was retweeted eighty times.

Figure 45 Within the collection of digital twin tweets, this tweet has the most retweets in January of 2020 [29]

The tweet's message is like many of the definitions reviewed earlier in this chapter and describes a digital twin as a virtual model that can bridge the physical and digital worlds [29]. The image represents a virtual replication of a city. Smart cities are a popular form of digital twins. The tweet utilizes many hashtags, such as #AR (augmented reality), #IoT (Internet of Things), and #AI (artificial intelligence), that draw Twitter users' attention and help gain more attention to the tweet based upon platform algorithms.

Unsupervised Clustering of the Digital Twin Tweets 6.2.3

We utilize a document term matrix as input into an unsupervised cluster analysis. The document term matrix is a large object that contains an identifier of each tweet, the words used within the tweet's text, and the frequency of the words. The clustering algorithm searches through the document term matrix and groups the tweets based upon patterns in the utilized words and their frequencies. To determine an appropriate number of groups, or clusters, to be created, we utilize the within-cluster sum of squares (WSS) and silhouette methods.

The WSS method will iterate through generations of clustering incrementing the number of individual groups with each generation. During each iteration, the squared distance between all the observations within the cluster and its center are summed. This is done for all clusters and the total WSS is then compared with the other generations each having an increasing number of clusters. The ideal number of clusters is frequently determined visually, known as the “elbow method”. The “elbow” is visually identified when the WSS decreases rapidly in initial generations of smaller n number of clusters and the decrease flattens as n increases. The WSS output is plotted in Figure 46 with a few potential “elbows” in the line occurring at two, four, and six clusters generated.

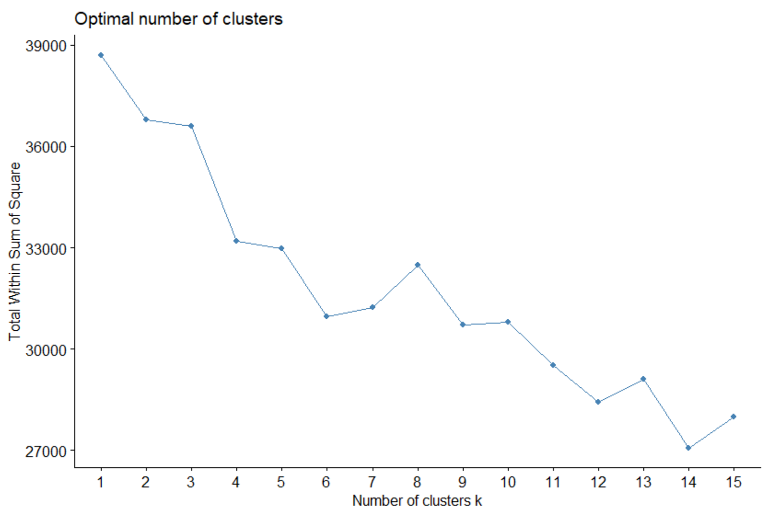

Figure 46: Within-cluster sum of squares indicates that the proper number of clusters, identified as “elbows” in the line, maybe two, four, or six groups

The silhouette method also strives to find the proper number of clusters in a collection of data. The method is like the WSS method in that it will iterate through generations of cluster creation and compare each generation. The comparison is performed across the distance between observations in a cluster and observations in the neighboring cluster. If many clusters exist within a small dimension, observations will be near neighboring observations, and this may indicate that too many clusters have been generated for the dataset. We utilize R libraries of nbclust and factoextra to quickly implement the WSS and silhouette methods. The output of the silhouette method is found in Figure 47 and identifies four clusters as the appropriate amount for our collection of digital twin tweets.

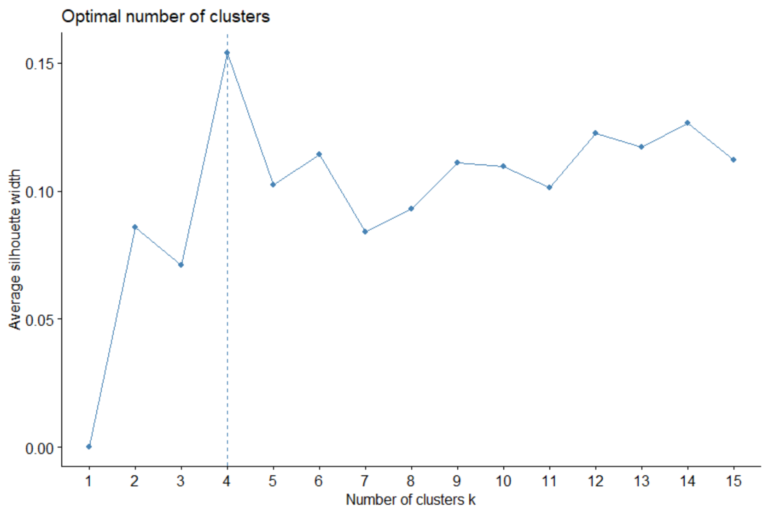

Figure 47: Average silhouette method indicates four clusters as the appropriate grouping size for our collection of tweets

Another helpful data visualization graphic when performing text analysis and hierarchical clustering is the dendrogram. Dendrograms are tree-based graphics that indicate relationships. Dendrograms are frequently created when observing the distance between observations in document term matrices and help visualize cluster distribution. The problem with dendrograms is that they do not scale well when the number of observations approaches many thousands. In these cases, the graphics become either quite large or very densely populated making discernment difficult. Because dendrogram diagrams do not scale well with large observations, we have cast only a sample of 1% of our 24,000 tweets. Dendrograms can be customized with specific visual formats such as the typical tree diagram, circular and in our case, we are utilizing the phylogenic shape. Phylogeny is the development of traits or taxonomic grouping. It can help discuss biology and the evolution of species. Here we utilize a phylogenic dendrogram to illustrate the evolution of the conversations within the digital twin tweets, illustrated in Figure 48.

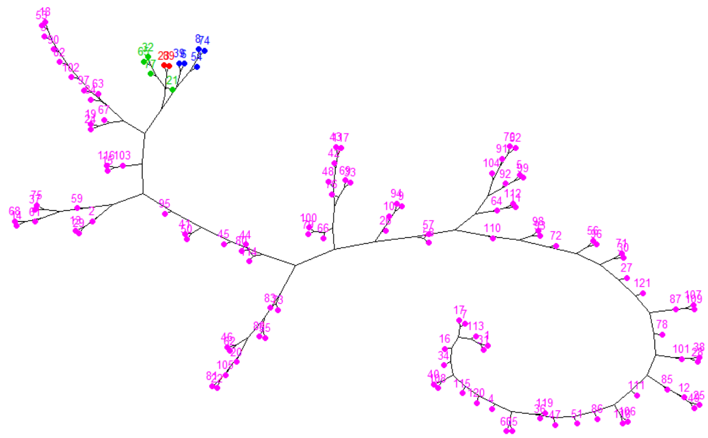

Figure 48: Phylogenic dendrograms can be created using distance calculations from document term matrices but dendrograms do not scale well with large numbers of observations

Trends were extracted from the four clusters by the frequency of mention. The largest cluster in the volume of tweets is the first group, magenta in the phylogenic dendrogram (only a sample of 1% of tweets were used to generate the graphic), and the top seven trends by mention come from this first cluster:

- the Internet of Things
- Artificial Intelligence
- industry use
- collaboration
- the virtual world
- novelty
- data

The remaining three clusters then provide three other trends to round out the top ten: machine learning, blockchain, and augmented reality. The largest cluster of tweets is displayed by word cloud in Figure 49, further illustrating many of the top trending concepts.

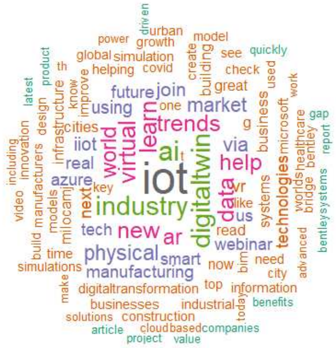

Figure 49: Word cloud graphic of the most frequent terms from cluster one

Tweets can be retweeted by users to further promote the message content. Figure 50 is the most retweeted post within the first cluster. The tweet in Figure 50 references predictions for the year 2020 created by a technology and business solutions provider, Global NTT [30].

Figure 50: The most retweeted post from the first cluster of tweets references technology predictions by Global NTT [30]

Digital twins are one of the emerging technologies that the predictions include. The tweet's embedded link is to an online article that summarizes the predictions and mentions that digital twins can collect data from instrumented assets, model behavior, identify patterns and create more accurate conclusions [31].

From the second cluster, the most retweeted post references the 44th episode of IoT Coffee Talk, an online webinar. This tweet merges business concerns such as the conversion of manual, human-driven, or paper processes into optimized and automated processes via digital transformation. These technologies may improve the efficiency of industry and consumer behaviors to also solve sustainability concerns. Some of the hashtags in Figure 51 are like those within Figure 50, #AI, #AR, and #DigitalTwins. New topics are introduced including 5G, edge, cloud, sustainability, and digital transformation.

Figure 51: A tweet by @RobTiffany links to a webinar episode hosting discussions on IoT technology [32]

The most retweeted post from the third cluster is again a reference to trends and predictions, this time the trends listed were identified by the research and advisory organization, Gartner. The tweet in Figure 52 links to an article that identifies eight trends in three categories with two additional cross-cutting trends. The three categories include Intelligent, Digital, and Mesh. Digital twins are identified as the fourth 2019 technology trend by Gartner [33]. The article by Panetta [33] further states that digital twins have:

- robustness in their modeling profile to support business outcomes
- link to physical assets to potentially model and control
- drive new business opportunities when big data analytics and AI are applied
- interaction to help evaluate future states such as modeling and simulation

Figure 52: The most retweeted post from the third cluster which references trends and predictions by an industry research and advisory organization[34]

The most retweeted post in the fourth cluster [35] offers some distinction from the previous three. This tweet references an open-source distributed ledger system that is like standard blockchain but utilizes a different algorithm requiring less energy [36]. Because IOTA can run on devices having less computational power and bandwidth, it enables the value and security of distributed ledger in the realm of IoT devices [36]. The tweet in Figure 53 mentions the tangle algorithm, which is used by the IOTA distributed ledger, is a technology that could secure digital twins by creating more trust in the IoT ecosystem.

Figure 53: The most retweeted post in the fourth cluster references a distributed ledger technology that can operate in the constrained realm of small IoT devices [35]

Twitter analysis by industry 6.2.4

Content-based analysis of the tweets has identified the mentions of specific industries. The International Labor Organization maintains a curated list of industries and descriptions [38]. This curated list can be utilized in a labeling algorithm to identify industry mentions within the tweets. The health industry is the most mentioned within this collection of tweets, followed by entertainment and utilities. The textile industry was not mentioned within our collection of tweets, illustrated in Figure 54.

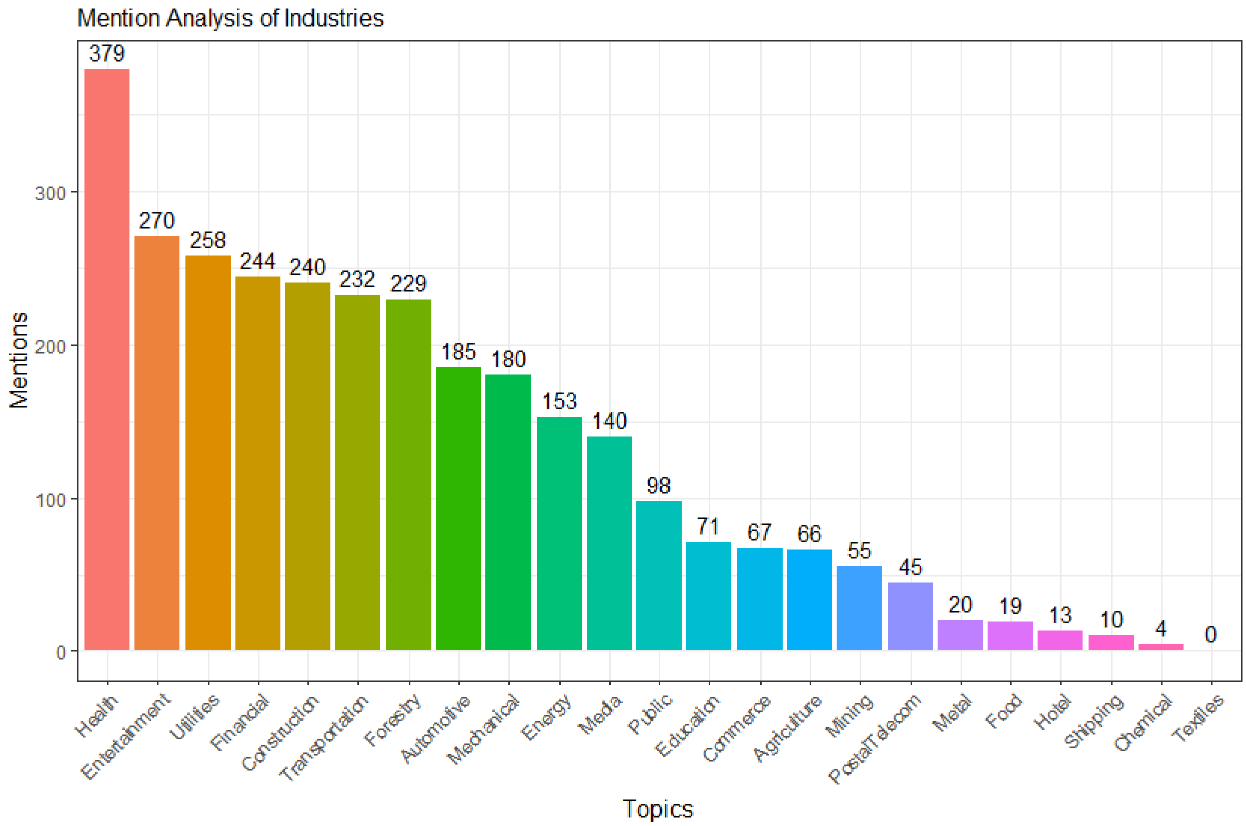

Figure 54: The health industry had the most mention within the collection of digital twin related tweets

Considering the sentiment and emotions that are prevalent in the tweets is an interesting research angle. Sentiment analysis typically reviews content on a continuum of negative to positive. Our sentiment analysis will review the tweets by industry and for specific emotions that may be felt or influenced by the message of the tweets, including sentiments such as anticipation or fear among others. This analysis will utilize the NRC lexicon to label the tweet's sentiment [38].

The most tweets were labeled to the *health* and *entertainment* industries (shown in Figure 54). It is more probable that a tweet using words that convey *anticipation* will reference the *health* industry (31.0%) compared to the *entertainment* industry (8.7%). The naïve Bayes algorithm was utilized to determine these probabilities. The formula for naïve Bayes is explained for our classification problem and data set in Equation 2.

Equation 2: The naive Bayes equation explained

Probability of *sentiment* given a specific *industry* in mentioned

Count of *industry* observations divided by total number of observations

Probability of *industry* mention given a specific *sentiment*

$$P(A|B) = \frac{P(B\,|\,A) \times P(A)}{P(B)}$$

Number of instances of a *sentiment*, divided by all observations

To calculate these probabilities, we may start with the numerator, which is the product of the probability of positive sentiment given a specific industry and the probability of a tweet having a relationship to the same specific industry. This example will focus on the sentiment of anticipation and the health industry as those factors have many observations. Table 14 identifies the sentiment factors across all industry classes and will inform our formulas for the health industry.

Table 14: Counts of sentiments by classes of industry

Industry	Anger	Anticipation	Disgust		Fear	Joy	Negative	Positive	Sadness	Surprise	Trust	Totals:
Agriculture		13			1	11		42			8	75
Automotive	1	1			1	1		8			8	20
Commerce		20			2	6	5	46		3	19	101
Construction	7	25	2		7	11	11	165	1	4	56	289
Education	3	18	1		2	11	2	62	1		16	116
Energy	2	9	1		2	21	6	75	2	1	29	148
Entertainment		44			50	38	45	137		43	3	360
Financial	6	22	1		7	17	4	80		11	33	181
Food	2	6				4	1	8			4	25
Forestry	10	33	4		10	30	7	120	3	9	64	290
Health	18	157	4		137	14	11	255	19	3	162	780

Industry	Anger	Anticipation	Disgust		Fear	Joy	Negative	Positive	Sadness	Surprise	Trust	Totals:
Hotel		2				3	1	5	1		3	15
Mechanical	1	5				7	3	22		2	10	50
Media		4						6			1	11
Metal	1	2			2	5	1	14		2	1	28
Mining	3	4			5	7	4	47		2	21	93
Postal/Telecom		4				8		30		7	8	57
Public	1	23			7	7	5	49		2	8	102

In this training data set of the classification model, there are 780 tweets that reference *health* and within those tweets, 157 have the sentiment of *anticipation*. The conditional probability of anticipation-sentiment (*B* in Equation 2) given an agricultural tweet (*A* in Equation 2) is presented in Equation 3:

Equation 3: Conditional probability of positive sentiment given a tweet references the agriculture industry

$$P(Anticipation|Health) = 157 \div 780$$

$$P(Anticipation|Health) = 0.201$$

The conditional probability of anticipation-sentiment given a tweet referencing health is 20.1%. This can be validated with the results found in Table 15. To complete the numerator, we need the product of the P(B|A) and the *a priori,* or the number of *health* related instances divided by the total number of size in the data set. Equation 4 determines the *a priori.*

Equation 4: The *a priori* is the probability of a tweet referencing *health* and is found by dividing the count of *health* instances by the total training data set count.

$$P(Health) = 780 \div 3411$$

$$P(Health) = 0.229$$

The numerator is divided by the probability of a tweet having the sentiment of anticipation. This is determined by dividing the number of anticipation instances for all classes

(507) by the total amount of training data instances (3411). Probability of a tweet having the sentiment of anticipation is determined in Equation 5.

Equation 5:

$$P(Anticipation) = 507 \div 3411$$

$$P(Anticipation) = 0.149$$

The *posterior,* or the probability of a tweet referencing the health industry if we know that the tweet contains the sentiment of anticipation can be determined using the completed equation in Equation 6.

$$P(Health|Anticipation) = 0.310 = \frac{0.201 * 0.229}{0.149}$$

This equation could be done a second time with the numerator updated for the *entertainment* industry to prove the statement, given a tweet has the sentiment of anticipation, there is a greater probability that the tweet references the *health* industry (31.0%) compared to the *entertainment* industry (8.7%). The conditional probabilities, such as Equation 3, can be confirmed using the output of the R functions NBclassifer and naiveBayes, listed in Table 15 below.

Table 15: Conditional probability of sentiment given an industry

Sentiment given Industry	**Anger**	**Anticipation**	**Disgust**	**Fear**	**Joy**	**Negative**	**Positive**	**Sadness**	**Surprise**	**Trust**
Agriculture	0.000	0.173	0.000	0.013	0.147	0.000	0.560	0.000	0.000	0.107
Automotive	0.050	0.050	0.000	0.050	0.050	0.000	0.400	0.000	0.000	0.400
Commerce	0.000	0.198	0.000	0.020	0.059	0.050	0.455	0.000	0.030	0.188
Construction	0.024	0.087	0.007	0.024	0.038	0.038	0.571	0.003	0.014	0.194
Education	0.026	0.155	0.009	0.017	0.095	0.017	0.534	0.009	0.000	0.138
Energy	0.014	0.061	0.007	0.014	0.142	0.041	0.507	0.014	0.007	0.196
Entertainment	0.000	0.122	0.000	0.139	0.106	0.125	0.381	0.000	0.119	0.008
Financial	0.033	0.122	0.006	0.039	0.094	0.022	0.442	0.000	0.061	0.182
Food	0.080	0.240	0.000	0.000	0.160	0.040	0.320	0.000	0.000	0.160
Forestry	0.034	0.114	0.014	0.034	0.103	0.024	0.414	0.010	0.031	0.221

Sentiment given Industry	Anger	Anticipation	Disgust	Fear	Joy	Negative	Positive	Sadness	Surprise	Trust
Health	0.023	0.201	0.005	0.176	0.018	0.014	0.327	0.024	0.004	0.208
Hotel	0.000	0.133	0.000	0.000	0.200	0.067	0.333	0.067	0.000	0.200
Mechanical	0.020	0.100	0.000	0.000	0.140	0.060	0.440	0.000	0.040	0.200
Media	0.000	0.364	0.000	0.000	0.000	0.000	0.545	0.000	0.000	0.091
Metal	0.036	0.071	0.000	0.071	0.179	0.036	0.500	0.000	0.071	0.036
Mining	0.032	0.043	0.000	0.054	0.075	0.043	0.505	0.000	0.022	0.226
Postal/Telecom	0.000	0.070	0.000	0.000	0.140	0.000	0.526	0.000	0.123	0.140
Public	0.010	0.225	0.000	0.069	0.069	0.049	0.480	0.000	0.020	0.078
Shipping	0.000	0.333	0.000	0.000	0.000	0.000	0.667	0.000	0.000	0.000
Transportation	0.022	0.181	0.007	0.014	0.018	0.040	0.567	0.018	0.043	0.090

Given a tweet references the food industry, there is a 8.0% probability that the sentiment of the tweet will be anger. The probability of the sentiment of trust occurring is highest for the industry of automotive; however, automotive referencing tweets only convey trust with a probability of 40.0%. The sentiment of disgust is rarely found in the tweet messages, the highest probability of disgust was found in messages labeled towards the industry of forestry (1.4%).

The R library e1071 offers a naïve Bayes function that eases the implementation of the algorithm. Unfortunately, given the quantity of data we have, and the factors supplied to the model, we only achieve an accuracy of 25.5%. To increase the accuracy of this model, first, increase the number of tweets in the collection and second, improve the factor selection beyond only utilizing the factor of sentiment.

Network graphs visually identify relationships. Within the analyzed conversations, not all industry-related tweets reference the top trends. Tweets that reference the food or hotel industries have very little relationship to trends. This is visible in Figure 55, a network graph where the industry nodes are yellow, the trend nodes are green, and the relationships between these labels are red lines. The construction industry tweets are the most inclusive to top trends.

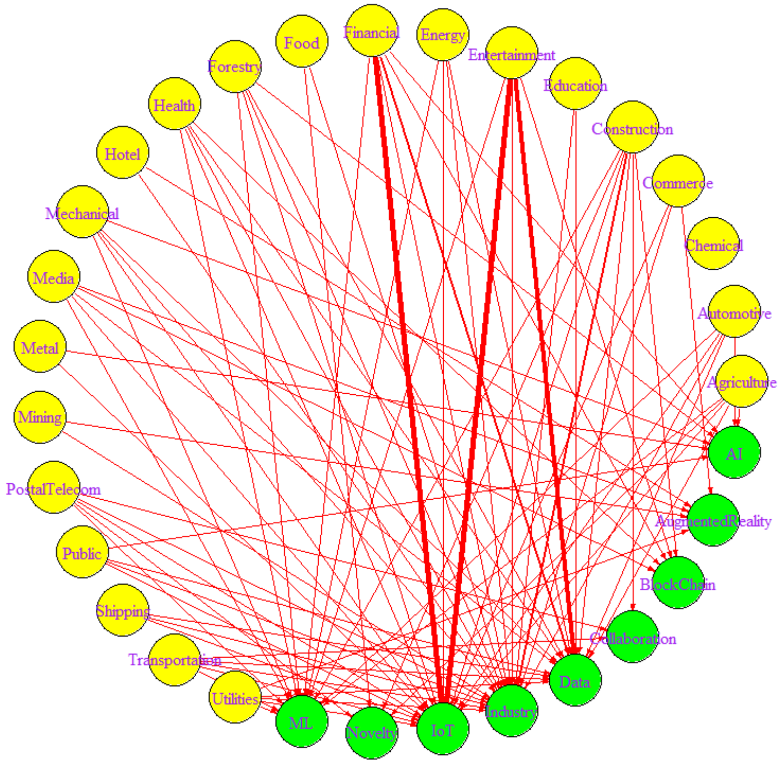

Figure 55: This network graph illustrates the relationships found in tweets between industries and trends

Background on Maturity Models 6.3

While we have noted the many definitions of the digital twin, determined popular industries in the public discussion, and uncovered the sentiment in the conversations, we have not uncovered what a good digital twin is. To help organizations determine the level of value, and to further improve and enhance their development process, we suggest a digital twin maturity model.

Maturity models help organizations achieve higher levels of capability within a discipline or process [39]. To increase the capability or capacity, an organization first places itself along a trajectory that is determined by current performance [40]. Achieving a state of greater capability along the same course becomes the goal. A maturity model establishes the milestones of capability and the distance between current and goal states.

Assessments of maturity inform organizations and their leadership teams about their current capability and readiness. Organizations frequently utilize a questionnaire to place their competencies or system capabilities along the maturity model. These can be self-assessments or utilize consultants. The questions and the maturity model assessment effort evaluate Key Performance Indicators (KPIs) to position the organization and system capabilities.

Organizations have two options when requiring a maturity model. The first option is to apply a generic model and the second option is to build a specific model for a given problem domain [40]. To build a model that is specific to a problem domain, five factors must be considered: domain context, capability characteristics, interaction with domain experts, quantitative research, and qualitative research [40].

The Digital Twin maturity model 6.4

The creation of a maturity model for digital twins requires defining the benefit that would come from using the model. Kluth et al. describe a maturity model as a representation to evaluate business processes [41]. Kohlegger et al. describe a maturity model as a representation of increasing capabilities over distinct levels [42]. A maturity model for digital twins is a tool and associated process to measure increasing and distinct milestones of value derived from a digital twin by its capabilities.

After defining the benefit, the next step in creating a maturity model is to determine the characteristics and parameters of digital twins for distinction along a path of increasing system capability or capacity. Some foundational parameters already established include governance, supportive technology, connectivity, value generation, and the competencies of the organization [43]. Maturity models frequently include people and culture factors such as the skills, organizational structures, and processes, as well as supporting technology [44].

A digital twin maturity model can be informed by existing models for Industry 4.0. Industry 4.0 describes the integration of people, data, and equipment to allow flexibility and autonomy in manufacturing [45]. Industry 4.0 represents the transition away from mostly mechanical to mostly digital manufacturing [46]. Given a digital twin is of a factory, it would be integral to Industry 4.0. Digital manufacturing is aided by these technologies and principles [47]:

- Changeability
- Decentralized Decisions
- Interoperability
- Real-time Reaction
- Simulation

Other technology trends and principles that are key to Industry 4.0 include big data, cloud, additive manufacturing, AR, robotics, and machine-machine-human integration [48]. The IoT technology is foundational for these mechanisms [40]. IoT provides big data, may include additive manufacturing and robotic instrumentation, and can inform both machines and humans in the loop. A digital twin of an Industry 4.0 plant is the composition of these mechanisms for modeling, monitoring, simulating, and securing the physical plant relative to its environment. Because organizations differ from each other due to resources, composition, culture and other factors, copying another organization's digital twin implementation will not necessarily lead to success [40]. However, a maturity model can help guide the capabilities and improvements of a digital twin along a path to implement these mechanisms.

There are many different maturity models. Some of the models are well known and not specific to digital twins, such as the CMMi model. Other models have a higher correlation to digital twins based upon their focus on digitalization, such as the SMSRA and M2DDM. Further models have been created by technology and consulting companies that offer solutions or expertise. Those models include examples such as Rockwell Automation, Price Waterhouse Coopers, and Siemens (these models and others are listed in Table 16).

Table 16: Example maturity models and descriptions

Example Models	**Short Description**
Maturity Model for Data-Driven Manufacturing (M2DDM)	The Maturity Model for Data-Driven Manufacturing (M2DDM) contains six levels of maturity(begins at level 0). The 4th level is *Digital Twin* and characterized by smart systems, decentralized decisions, and centralized intelligence to keep humans in the loop. The 5th, and highest level, is the self-optimizing-factory. [49]

Example Models	Short Description
Smart Manufacturing Systems Readiness Assessment (SMSRA)	The Smart Manufacturing Systems Readiness Assessment (SMSRA) provides manufacturing organizations with an indication of their current factory state compared against a reference model of capabilities. The last stage is *transformed* implying the business has executed a change to its business model [50].
Complexity Management Maturity	The first level of the Complexity Management Maturity is initial and represents that an organization has not yet quantified the amount of complexity at hand [39].
C3M	The C3M model presents five levels of maturity for IT-based case management systems (CSM) across three phases of CSM adoption; pre-CSM, CSM, post-CSM. C3M is novel as it presents levels of capabilities and the risks that may be associated with the levels of benefits [51].
Capability Maturity Model Integration (CMMI)	Capability Maturity Model (CMM) was constructed in 1986 and updated in 2006 to include tech and process as the CMMI model. CMMI includes the phases of Initial, Repeatable, Defined, Managed, and Optimizing.
Test Maturity Model Integration (TMMi)	TMMi utilizes the same structure as CMMi and helps organizations gauge and improve their software testing practices [52]
Industry 4.0/Digital Operations Self-Assessment	PWC's self-assessment places an organization's Industry 4.0 capability concerning their target state and offers them a benchmark to the positions of industry competition [51]. Cognet et al. compared the PwC and IMPULS models and found that the IMPULS model has 84% coverage of the PwC digital maturity model's KPIs [44].
The Connected Enterprise Maturity Model	Created by Rockwell Automation, this five-stage maturity model offers best practices for modernizing culture and technologies when networking operational technology (OT) and information technology [54]
Digitization Roadmap	The digitization roadmap by Siemens is constructed to help organizations transform their business. Six areas, such as process, security, and collaboration, are reviewed, benchmarked and an associated ROI study is completed to evaluate financial consequences of improvement activities [55].

A good maturity model removes confusion by isolating the factors and priorities that will help an organization achieve the next level of capability. Parente and Federo suggest removing conjunction, equifinality, and asymmetry for causality in models to be clear to organizations [40][56]. Asymmetry is a characteristic of causality that may explain one result and then fails to explain another result [56]. Asymmetry can create doubt in the accuracy of a model. Equifinality implies that similar benefits and capabilities may be the outcome of more than one level of maturity. When equifinality exists in models, organizations will cease to increase the risk or cost in implementation as the value may not increase. A conjunction is a relationship between technologies, processes, or culture that holds back value creation until all related tenets increase in maturity together. If such related conditions are spread across maturity levels, intermediary benefits offered at lower levels won't become actual value until much higher levels of maturity are achieved. Maturity models should not suffer conjunction, equifinality, or asymmetry.

The ERP 4.0 maturity model by Basl et al [57] has six levels across dimensions of business model, technology, data, and processes. To construct the model, Basl et al analyzed trends from survey data and layered the trends into the maturity model levels based upon their frequency found from the survey [57]. The survey was completed by 26 ERP system suppliers [58]. Trends having the most frequency of being acknowledged by the system suppliers were positioned higher into the levels of the maturity model. The trends were identified through the survey included cloud, IoT, blockchain, digital twins, edge computing, AI, big data, social networks, and AR/VR [58]. These trends are very similar to those identified through social media analytics and are illustrated as green network nodes in Figure 55. The most frequent trends include cloud, IoT, and AR. Trends with lesser frequency include extending asset life, optimizing performance, and implementing blockchain. Other trends included big data, mobile ERP apps, and in-memory computing (IMC) [57]. A segment of Basl et al ERP 4.0 maturity model is illustrated in Table 17.

Table 17: A subset of Basl's ERP 4.0 maturity model [57]

Level	Description and Inclusion
0	Traditional RDBMS system, with basic ERP process automation, and no cloud adoption
1	Mobility, additional automation, and digitization of processes
2	The complexity, digital capabilities, and analysis all increasing
3	Initial migration to cloud services, business intelligence efforts underway, continued increase in process automation
4	As-a-Service implementations, IoT integration, Digital Twin capabilities
5	AI, RPA, all Cloud deployment, all business processes automated

The digital twin maturity model has been informed according to the guidance by de Jesus and Lima of using context, characteristics, expertise, survey (social media analysis), and qualitative research [40]. Academic research, commercial solution and providers' models, and social media analytics were input factors for the creation of the digital twin maturity model. Basl's method utilized in the ERP 4.0 maturity model creation uses trend popularity to determine the maturity levels. While not the final version of our model, the approach by Basl does offer insight into public opinion and the volume of driving trends (as illustrated in Figure 56).

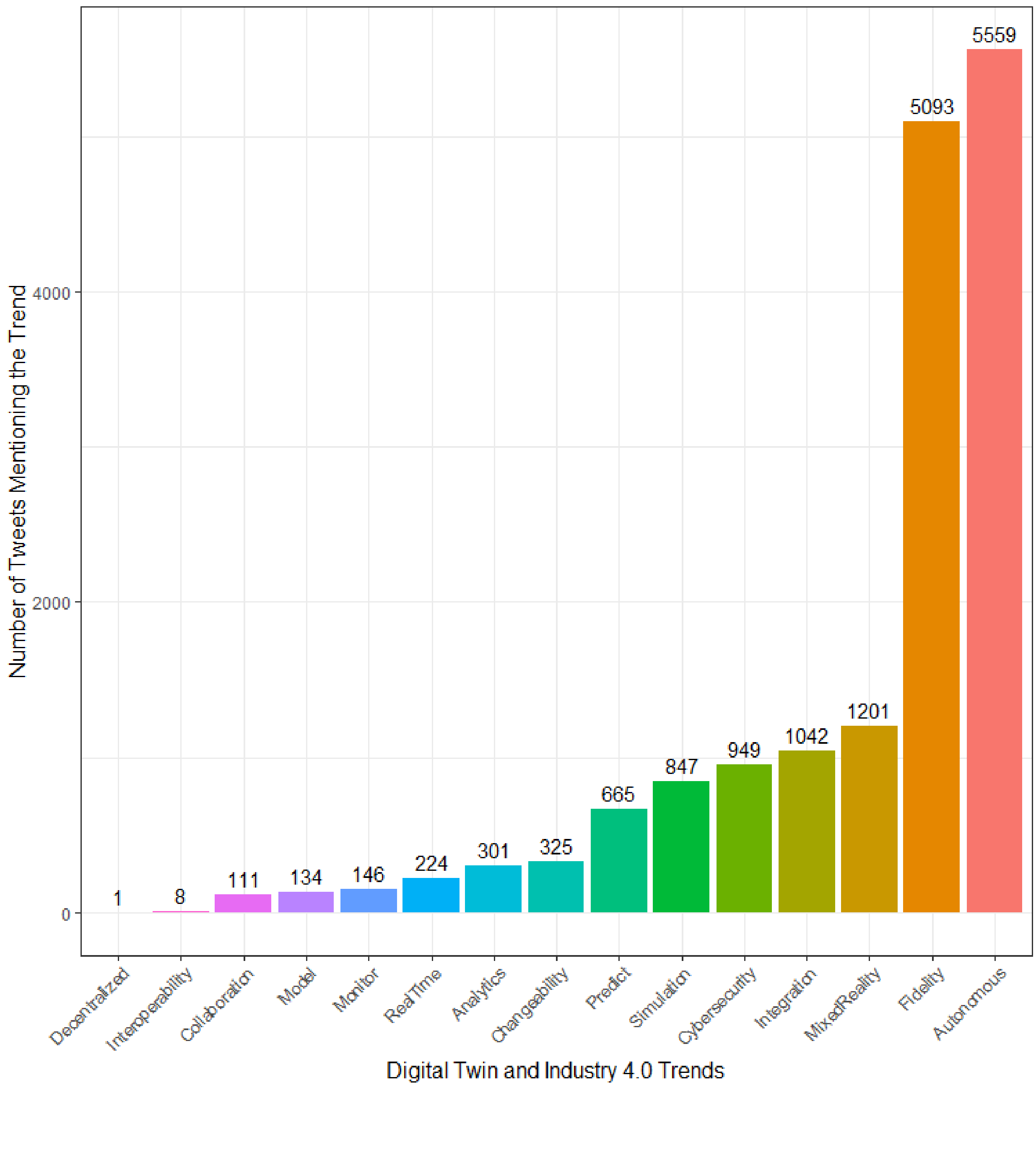

Figure 56: Amount of maturity level/trend mention in our social media analysis

Trends identified from the Industry 4.0 models and academic research, such as *decentralized* and *interoperability*, have little mention within our collection of tweets. *Collaboration* has many more references compared to the two least mentioned topics. Another jump exists between the trends of *changeability* and *predictability;* however, the topics of *fidelity* and *autonomy* retain the most conversation found in the collection of tweets.

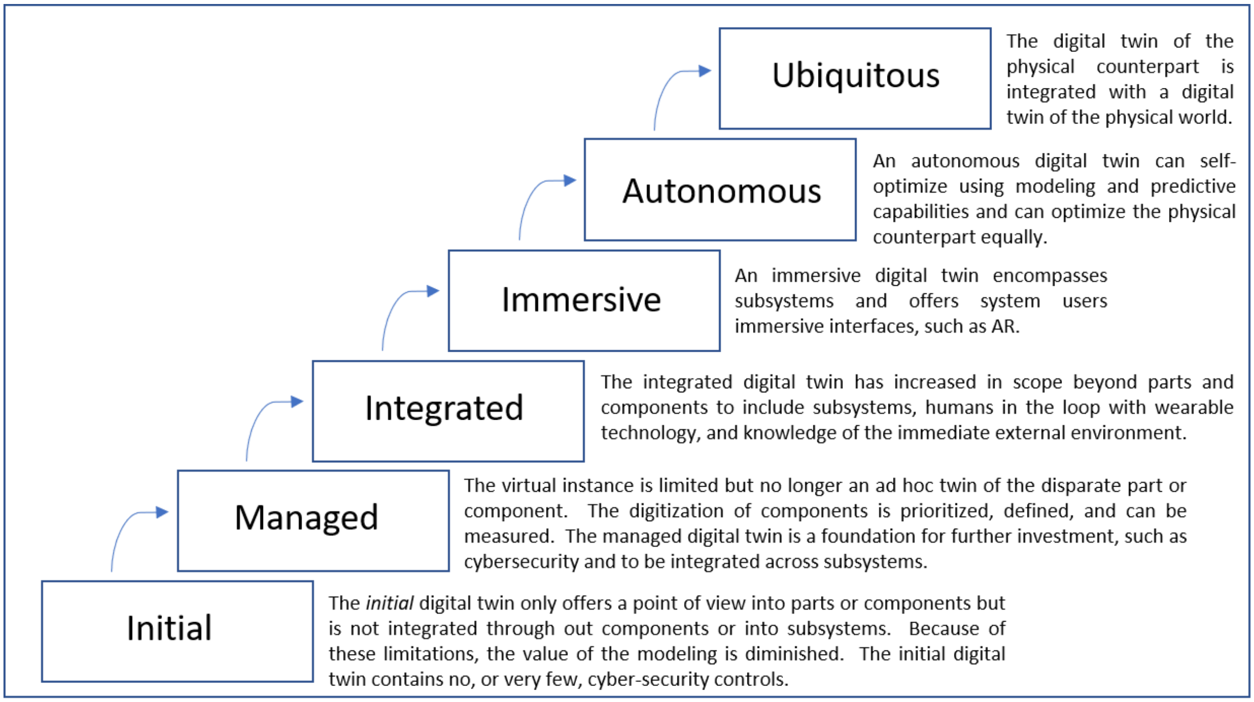

Figure 57: The six levels of the digital twin maturity model

The digital twin maturity model is composed of six levels. The lowest maturity level is the initial digital twin. The initial digital twin is limited in scope, such as only instrumenting a few parts and components. The initial digital twin offers limited insight and is far from being complete, integrated, or even secured. The second level of maturity is the *managed* level. At the managed level, the twin has a prioritized roadmap and coverage beyond ad hoc parts and includes parity with system components. The third level of maturity increases the digital twin's capability via integration and interoperability. At this third level of *integration,* the digital twin can model, monitor, and predict many of the physical subsystems. The fourth level of maturity, *immersive,* is mostly defined by its human interface. At the fourth level of maturity, the digital twin is assessable using immersive interfaces, such as augmented or virtual reality. The *autonomous* level of the digital twin has the cybersecurity, integration, and authority, among other characteristics, to self-optimize its physical counterpart. Finally, the sixth level of maturity of a digital twin is when it consumes the context of its environment. This is the *ubiquitous* level. This level requires investment and technology that will be beyond the scope of most organizations. Achieving the ubiquitous level requires instrumenting the physical world, beyond the immediate assets, to understand global weather patterns, political, social, and economic phenomena, as well as other growing concerns. Table 18 provides the capabilities and their descriptions.

Table 18: The six levels of capability and a short description of their enablement

Capability	**Description of enablement**
Initial	At this level of maturity, the digital twin can model a selection of parts or a few components of the system. The digital twin can inform human operators and offers a viewpoint towards collaboration. It is far from a smart or autonomous capability.
Managed	Digital twins increase the cybersecurity risk footprint by increasing integration touchpoints and consuming data in transit, storage, and processing. A managed digital twin is measured for its ability to secure itself and the physical asset. The managed digital twin has moved beyond ad hoc instrumentation of parts into a prioritized roadmap that incorporates cybersecurity concerns.
Integrated	A complex system is composed of many systems and subsystems. At this level, the digital twin incorporates all targeted data sources into a unified virtual instance of the physical counterpart.
Immersive	A digital twin at this level offers a modern and immersive interface with AR or VR capabilities. Beyond monitoring the components, the VR interface may offer simulated experiences of the components and the system.
Autonomous	Once the digital twin is integrated, informed, and secured, it may become *smart* or optimized without decisions from a human control interface.
Ubiquitous	Complex systems operate within the context of their environment. A ubiquitous digital twin of a physical asset must integrate with a digital twin of the physical world, such as climate models. This level of maturity requires investment and integrations that organizations will scope out of their implementation for more years to come.

Referring to Parente and Federo's guidance for models to be effective for organizations [40][56], we test our digital twin maturity model for conjunction, equifinality, and asymmetry. Conjunction in a maturity model exists when the benefit promised by achieving a lower level, such as the *integrated* digital twin, cannot be reaped until the digital twin reaches an advanced level, such as *immersive*. In our model, for example, value is delivered to an organization at the integrated digital twin level, as that level of maturity allows the twin to grow from modeling individual parts or components into modeling entire subsystems. Furthermore, value is achieved at the integrated level through engaging with users with wearable technology, such as understanding the physical location of system operators for safety reasons. It is clear then that value arrives at the *integrated* level without requiring the *immersive* level to have been met.

Similarly, asymmetry would damage the trust in the digital twin model when a characteristic fails to explain its importance in each of its succeeded levels. For example, the integrated digital provides an API that can be consumed by a headset offering an immersive experience. The integrated digital twin also provides interfaces to the sensors and actuators that will be utilized to digitally annotate the physical world through augmented reality. Furthermore, an integrated digital twin is required for the autonomous digital twin to exist. The autonomous digital twin requires integrations to the many parts, components, and subsystems to control and optimize the physical asset. Even the *autonomous* digital twin requires integration to the digital twin of the physical world. If we moved a step down in maturity, down from the integrated digital twin to the managed, all the previous features and benefits would exist in a product roadmap but not in the implementation. The managed digital twin is more than a simple roadmap and vision, it offers an implementation whose limited existence is now counted (managed and measured) so that vulnerabilities, risk, and remediation are a part of the planning and implementation. Without applying cyber-security early into the maturity model, future benefits would have a greater risk. Any future maturity state beyond the *initial* digital twin will always offer the original benefit of the digital point of view into a limited part or component.

The last area to defend the digital twin maturity model includes the characteristic of equifinality. Equifinality implies that similar benefits and capabilities may be the outcome of more than one level of maturity. If a digital twin were at the maturity level of *initial*, we would not want to allow the twin to become *autonomous* nor would the benefits of an autonomous system be reached at the *initial level*. The small scoped system could likely ruin many integrated parts and components, as it is not yet informed of the entire system’s states, such as whether dependencies are operating, within appropriate thresholds, failed, or shutdown. The *initial* twin would need to reach the integrated phase to have this knowledge and should not have widespread integrations with other systems without first safely being counted, measured, and secured in the *managed* level. The benefit at each phase of our model can be reached at that level, without delaying the benefit until future phases. It is important to note that while cybersecurity is a component of the *managed* level, cybersecurity must be addressed throughout later phases.

Conclusion and Future Work 6.5

This chapter introduced findings from social media analytics on digital twins as well as a new maturity model. From the social media analytics, the top three trends identified included the IoT, AI, and industrial uses. An analysis into the industrial uses found the health industry as the most mentioned, followed by entertainment and utilities. The textile industry was not mentioned within the collection of tweets used in this research.

Sentiment analysis was performed on the messages within the tweets and a comparative analysis was offered across industries. Given a tweet references the food industry, there is an 8.0% probability that the sentiment of the tweet will be anger. The probability of the sentiment of trust occurring is highest for the industry of automotive; however, automotive referencing tweets only convey trust with a probability of 40.0%. The sentiment of disgust is rarely found in the tweet messages, the highest probability of disgust was found in messages labeled towards the industry of forestry (1.4%). Given a tweet has the sentiment of anticipation, there is a greater probability that the tweet references the *health* industry (31.0%) compared to the *entertainment* industry (8.7%).

Network graphs were utilized to visually identify relationships. Within the analyzed conversations, not all industry-related tweets referenced the top trends. Tweets that reference the food or hotel industries had very little relationship to top trends.

The collection of tweets identifies a peak in the discussions during January 2020. The tweet having the most retweets was retweeted eighty times [29]. That popular tweet's message was like many of the academic definitions reviewed in this chapter, as a virtual model that can bridge the physical and digital worlds.

To help organizations determine the level of value, to further improve, and to enhance their development process, we suggest a digital twin maturity model. The digital twin maturity model is composed of six levels: initial, managed, integrated, immersive, autonomous, and ubiquitous. The maturity model was discussed in terms of conjunction, equifinality, and asymmetry. These three characteristics should not exist in maturity models as they reduce trust in the accuracy and the value that maturity models offer. Future research should focus on case studies, implementing the maturity model, and further evaluating it for accurate causality of benefits achieved in the various phases.

REFERENCES:

[1] Jones, D., Snider, C., Nassehi, A., Yon, J., & Hicks, B. (2020). Characterising the digital twin: a systematic literature review. *CIRP J. Manuf. Sci. Technol.* 29, 36–52. https://doi.org/10.1016/j.cirpj.2020.02.002

[2] Rosen, R., Von Wichert, G., Lo, G. & Bettenhausen, K. D. (2015), About the importance of autonomy and digital twins for the future of manufacturing, *In 15th IFAC Symposium on Information Control Problems in Manufacturing INCOM*, 48(3), 567–572.

[3] Shafto, M., Conroy, M., Doyle, R., Glaessgen, E., Kemp, C., LeMoigne, J. & Wang, L. (2010), Draft modeling, simulation, information technology & processing roadmap, *Technology Area* 11.

[4] Grieves, M. (2014). Digital twin: manufacturing excellence through virtual factory Replication.

[5] Eckhart, M., & Ekelhart, A. (2019). Digital twins for cyber-physical systems security: state of the art and outlook. *In Security and Quality in Cyber-Physical Systems Engineering. Springer International Publishing*, 383–412. https://doi.org/10.1007/978-3-030-25312-7_14

[6] Negria, E., Fumagallia, L., & Macchia, M. (2017), A review of the roles of digital twin in cps-based production systems. *International Conference on Flexible Automation and Intelligent Manufacturing,* 939–948.

[7] Schroeder, G., Steinmetz, C., Pereira, C., Muller, I., Garcia, N., Espindola, D., & Rodrigues, R. (2016). Visualising the digital twin using web services and augmented reality. *In 2016 IEEE 14th International Conference on Industrial Informatics (INDIN)*, 522–527.

[8] Scheibmeir, J., & Malaiya, Y.K. (2020). Multi-model security and social media analytics of the digital twin. *ASTEJ.,* 5, 323-330.

[9] Stark, R., & Damerau, T. (2019). Digital twin. *In: The International Academy for Production Engineering, CIRP Encyclopedia of Production Engineering. Springer*. https://doi.org/10.1007/978-3-642-35950-7_16870-1

[10] Tao, F., Zhang, M., Liu, Y., & Nee A.Y.C. (2018). Digital twin driven prognostics and health management for complex equipment. *CIRP Ann. Manuf. Technol*, 66, 169-172.

[11] Haag, S., & Anderl, R. (2018). Digital twin – proof of concept. *Manuf. Lett.* 15, 64–66. https://doi.org/10.1016/j.mfglet.2018.02.006

[12] Glaessgen, E., & Stargel, D. (2012). The digital twin paradigm for future NASA and U.S. air force vehicles. *53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference.*

[13] Uhlemann, T., Lehmann, C., & Steinhilper, R. (2017). The digital twin: realizing the cyber-physical production system for industry 4.0. *Procedia CIRP.* 61, 335-340. https://doi:10.1016/j.procir.2016.11.152

[14] Hearn, M., & Rix, S. (2019). Cybersecurity considerations for digital twin implementations report. *IIC Journal of Innovation*.

[15] Girma, A. (2018). Analysis of security vulnerability and analytics of internet of things (IoT) platform. *In Information Technology - New Generations. Advances in Intelligent Systems and Computing*, 738, https://doi:10.1007/978-3-319-77028-4_16

[16] Atalay, M., & Angin, P. (2020). A digital twins approach to smart grid security testing and standardization. *In: IEEE International Workshop on Metrology for Industry 4.0 & IoT,* 435-440, https://doi:10.1109/MetroInd4.0IoT48571.2020.9138264

[17] Pal, D., Arpnikanondt, C., Funilkul, S., & Chutimaskul, W. (2020). The adoption analysis of voice-based smart IoT products. *IEEE Internet Things J.* 7(11), 10852–10867. https://doi.org/10.1109/JIOT.2020.2991791

[18] Burnett, D., Thorp, J., Richards, D., Gorkovenko, K., & Murray-Rust, D. (2019). Digital twins as a resource for design research. *In: Proc. Pervasive Displays 2019 - 8th ACM Int. Symp. Pervasive Displays, PerDis,* 6–7. https://doi.org/10.1145/3321335.3329685

[19] Novo, O., & Di Francesco, M. (2020). Semantic interoperability in the IoT. *ACM Trans. Internet Things* 1(1), 1–25. https://doi.org/10.1145/3375838

[20] Shelby, Z., Hartke, K., & Bormann, C. The constrained application protocol (CoAP). *RFC Editor. RFC 7252.* https://doi.org/10.17487/RFC7252

[21] Scheibmeir, J., & Malaiya, Y. (2019). An API development model for digital twins. *In Proceedings - Companion of the 19th IEEE International Conference on Software Quality, Reliability and Security, Institute of Electrical and Electronics Engineers Inc.,* 518–519. https://doi.org/10.1109/QRS-C.2019.00103

[22] Cruickshank, I.J., & Carley, K.M. (2020). Characterizing communities of hashtag usage on twitter during the 2020 COVID-19 pandemic by multi-view clustering. *Appl Netw Sci.* https://doi.org/10.1007/s41109-020-00317-8

[23] Martin, Y., Cutter, S., & Li, Z. (2020). Bridging twitter and survey data for evacuation assessment of hurricane matthew and hurricane irma. *Natural Hazards Review.* https://doi.org/10.1061/(ASCE)NH.1527-6996.0000354

[24] Bougie, G., Starke, J., Storey, M., & German, D.M. (2011). Towards understanding twitter use in software engineering: preliminary findings, ongoing challenges and future questions. *In: Web2SE '11. ACM,* 31–36. https://doi:10.1145/1984701.1984707

[25] Storey, M. et al. (2010). The impact of social media on software engineering practices and tools. *In: Proceedings of the FSE/SDP workshop on Future of software engineering research. ACM.* 359–364.

[26] Williams, A. (2018). Do software engineering practitioners cite research on software testing in their online articles?: a preliminary survey. *ACM.* 151–156. https://doi.org/10.1145/3210459.3210475

[27] Büchi, M. (2017). Microblogging as an extension of science reporting. *Public Understanding of Science,* 953–968. https://doi:10.1177/0963662516657794

[28] Gómez-García, M., Matosas-López, L., & Ruiz-Palmero, J. (2020). Social networks use patterns among university youth: the validity and reliability of an updated measurement instrument. *Sustainability*. https://doi:10.3390/su12093503

[29] Nappo, S. (@stephanenappo) "Virtual models, aka #digitaltwins, further bridging #physical & #digital worlds". 21-Jan-2020. Tweet.

[30] Rai, O., (@Omkar_Raii), "Predictions by @GlobalNTT reflect ho #emergingtech such as #AI, #ML, #IoT, #RPA #Cybersecurity, #DigitalTwins, #AR, #VR, #Blockchain & #BigData will significantly improve productivity, growth & innovation across entire work, live and play environments." 13-Nov-2019. Tweet.

[31] BW Online Bureau, "'Future Disrupted' predictions for 2020: Data, automation and IoT will enable virtual societies," *Smart Cities*, Nov-2019. [Online]. Available: http://bwsmartcities.businessworld.in/article/-Future-Disrupted-predictions-for-2020-Data-automation-and-IoT-will-enable-virtual-societies-/13-11-2019-178854/. [Accessed: 18-Jul-2021].

[32] Tiffany, R (@RobTiffany), "Welcome to #IoT Coffee Talk #44 where we chat about #Digital #IIoT #Automation #DigitalTwins #Edge #Cloud #DigitalTransformation #5G #AI #Data & #Sustainability over a cup of coffee." 24-Mar-2021. Tweet.

[33] Panetta K, "Gartner Top 10 Strategic Technology Trends for 2019," *Smarter with Gartner*, 15-Oct-2018. [Online]. Available: https://www.gartner.com/smarterwithgartner/gartner-top-10-strategic-technology-trends-for-2019/. [Accessed: 18-Jul-2021].

[34] Quindazzi, M (@MikeQuindazzi). "10 Strategic #Tech Trends for 2019 >>> #Gartner via @MikeQuindazzi >>> #AI #MachineLearning #DigitalTwins #AR #Blockchain #EdgeComputing #Autonomous #4IR >>> https://gtnr.it/2DymMLQ" 14-Sep-2019. Tweet.

[35] RolSOuLi (@M6sp2004), "#IOTA #iotatoken #coloredcoins #DigitalTwins $IOTA $TANGLE $MIOTA #MachineLearning #iotcommunity #iottrends #iottechnology #MIOTA #TANGLE #iotdevices #iotworld #iotsolutions #BigData #CloudComputing Ya!!! Fecha Oficial para la migración a Chrysalis" 18-Mar-2021. Tweet.

[36] Ullah, I., de Roode, G., Meratnia, N., & Havinga, P. (2021). Threat Modeling-How to Visualize Attacks on IOTA?. Sensors (Basel, Switzerland), 21(5), 1834. https://doi.org/10.3390/s21051834

[37] International Labor Organization. Industries and sectors. Accessed: May 16, 2021. Available: https://www.ilo.org/global/industries-and-sectors/lang--en/index.htm

[38] Mohammad, S., & Turney, P. (2010). Emotions evoked by common words and phrases: using mechanical turk to create an emotion lexicon. *In: Proceedings of the NAACL-HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*.

[39] Mittal, S., Khan, M.A., Romero, D., & Wuest, T. A critical review of smart manufacturing & Industry 4.0 maturity models: Implications for small and medium-sized enterprises (SMEs). *J. anuf. Syst.*, 49, 194–214.

[40] De Jesus, C., & Lima, R. M. (2020). Literature search of key factors for the development of generic and specific maturity models for industry 4.0. *Applied sciences,* 10(17), 5825–. Web.

[41] Kluth, A., Jäger, J., Schatz, A., Bauernhansl, T. (2014). Method for a systematic evaluation of advanced complexity management maturity. *Procedia CIRP.*, 19, 69-74. https://doi.org/10.1016/j.procir.2014.05.041

[42] Kohlegger, M., Maier, R., & Thalmann, S. (2009). Understanding maturity models - results of a structured content analysis. *In Proceedings of the 9th International Conference on Knowledge Management (I-KNOW'09)*, 51–61.

[43] Colli, M., Madsen, O., Berger, U., Møller, C., Wæhrens, B.V., & Bockholt, M. (2018). Contextualizing the outcome of a maturity assessment for Industry 4.0. *IFAC-Pap.*, 51, 1347–1352.

[44] Cognet, B et al. (2020). Towards a novel comparison framework of digital maturity assessment models. *Product Lifecycle Management in the Digital Twin Era. Springer International Publishing*, 58–71. Web.

[45] Agostini, L., & Filippini, R. (2019). Organizational and managerial challenges in the path toward Industry 4.0. *Eur. Innov. anag.,* 22, 406–421.

[46] Oztemel, E., & Gursev, S. (2018). Literature review of Industry 4.0 and related technologies. *J. Intell. anuf.*, 31, 127–182.

[47] Bakkari, M., & Khatory, A. (2017). Industry 4.0: strategy for more sustainable industrial development in SMEs. *In Proceedings of the IEOM 7th International Conference on Industrial Engineering and Operations Management*.

[48] Crnjac, M., Veža, I., & Banduka, N. (2017). From concept to the introduction of industry 4.0. *Int. Industrial Eng. Manag.*, 8, 21–30.

[49] Weber, C., Königsberger, J., Kassner, L., & Mitschang, B. (2017). M2DDM – A maturity model for data-driven manufacturing. *Procedia CIRP*., 63, 173-178. https://doi.org/10.1016/j.procir.2017.03.309

[50] Jung, K., Kulvatunyou, B., Choi, S., & Brundage, M.P. (2016). An overview of a smart manufacturing system readiness assessment. *In: Nääs I. et al. (eds) Advances in Production Management Systems. Initiatives for a Sustainable World. APMS 2016. IFIP Advances in Information and Communication Technology,* 488, https://doi.org/10.1007/978-3-319-51133-7_83

[51] Koehler, J., Hofstetter, J., & Woodtly, R. (2012). Capabilities and levels of maturity in IT-based case management. *In: Barros A., Gal A., Kindler E. (eds) Business Process Management. BPM 2012. Lecture Notes in Computer Science,* 7481, https://doi.org/10.1007/978-3-642-32885-5_4

[52] TMMi Foundation. (2020). Model Aims and Objectives. Retrieved July 24, 2021, from https://www.tmmi.org/model-aims-and-objectives/

[53] PWC. (n.d.). Industry 4.0 Self-Assessment. Retrieved July 24, 2021, from https://i40-self-assessment.pwc.de/i40/landing/

[54] Parkinson, B. (2015). The Connected Enterprise® Maturity Model: Metrics that Matter. Retrieved July 24, 2021, from https://www.rockwellautomation.com/en-us/company/news/blogs/the-connected-enterprise-maturity-model--metrics-that-matter.html

[55] Siemens. (2020). Https://www.plm.automation.siemens.com/media/global/pt/Siemens SW Digitalization roadmap Fact Sheet_tcm70-71287.pdf. Retrieved July 24, 2021, from https://www.plm.automation.siemens.com/media/global/pt/Siemens SW Digitalization roadmap Fact Sheet_tcm70-71287.pdf

[56] Parente, T.C., & Federo, R. (2019). Qualitative comparative analysis: justifying a neo-configurational approach in management research. *RAUSP anag.*, 54, 399–412.

[57] Basl, J., & Novakova, M. (2019). Analysis of selected ERP 4.0 features and proposal of an ERP 4.0 maturity model. *Research and Practical Issues of Enterprise Information Systems. Cham: Springer International Publishing*, 3–11. Web

[58] Novakova, M. (2019). *Trends of enterprise information systems in 4.0 conditions*. [Diploma Thesis, University of Economics]

CHAPTER 7: QUALITY CHARACTERISTICS FOR DIGITAL TWINS

Awareness of digital twin technology is growing, and the concepts and practices continue to evolve from early definitions. However, research towards the quality of a digital twin is limited. We utilized a collection of research articles as input into a machine-learning algorithm to derive a novel quality model for digital twins. We performed a comparative analysis from our new quality model to the established ISO 25010 instance across a collection of tweets gathered from the Fall of 2019 to early 2021. The social media analytics found that quality characteristics assembled into our newly derived quality model account for 7.5 times more social media mention than those from the ISO 25010 model. Sentiment analysis of our derived quality characteristics found that sentiment was lowest in the digital twin tweets categorized towards *cybersecurity*. We also evaluated the potential use cases of digital twin technology across the collected social media conversations. Of the selected use cases, *autonomous vehicles* were the most popular in mentions (5,517), followed by *smart cities* (2,593) and then *smart homes* (2,338). We performed a content-based analysis of open-source and commercial test automation tools from the social media discussions. While the commercial test automation technology Tricentis had the most mention, it was SoapUI, an open-source API testing tool, that had the most positive sentiment. We built a naïve Bayes model to predict whether tweets about digital twins would be favorites or be retweeted. We found that the factor contributing most to the popularity of a digital twin tweet was topical. Tweets towards topics of *augmented reality* and *DevOps* concepts had the most retweet or favorited actions. By modeling popular tweets that reference the Industrial Internet of Things or Intelligent Transportation Systems, we identified the leading conditional probabilities as the topics of *smart homes* (32.%) followed by *autonomous vehicles* (30.1%).

Introduction 7.1

Digital Twins (DT) growth is on the rise. The digital twin concept was first applied in aviation, and then towards the manufacturing sector [1]. Digital twins allow system users to monitor, control, predict and secure the physical twin [2]. Scheibmeir and Malaiya have found that the concept of the digital twin is popular in the education, health and medical, and government sectors [2]. The same study found that the sentiment of digital twin conversations on social media platforms was most positive in the industries of engineering and manufacturing, then commerce, and energy. Many benefits of digital twins have been noted.

- Digital twins are utilized for cost savings by monitoring equipment for failure [3].
- Digital twins can help designers understand how a consumer product is being utilized [4].
- Digital twins are useful for providing information at both an individual and aggregate [4]
- Over time, digital twins provide a source of big data that can be utilized in machine learning algorithms [5].

The Novel Interface of Augmented Reality 7.1.2

The history of AR could be made brief by highlighting milestones such as the invention of the first Head Mounted Display (HMD) by Ivan Sutherland during the 1960s [6]. The term ‘augmented reality’ was coined in the 1990s by Tom Caudell and David Mizell [7]. Also, in the 1990s a 2D marker for camera tracking and object detection was introduced, and after that ARToolKit was developed by Hirokazu Kato to ease the development of building AR applications [8]. In 2004, the first see-through AR application was built using a mobile phone [9]. Mobile and web technologies such as Android and iOS operating systems, devices and modern frameworks, and SDKs such as WebXR have all offered a revolution for this technology.

Modern mobile devices offer the most common platform for AR applications. However, there are limitations to AR experiences over mobile devices including language, learnability, and timeliness [10]. A benefit of the mobile platform is that the technology platform easily moves with participants as they select different physical paths [10]. Using the BYOD model allows for accelerated time to value as AR users are already associated with their mobile devices. Interactive technology can also influence emotions such as authenticity and cognitive engagement [11]. Table 19 offers definitions of experience relating to application user interfaces and end-user interactions [10].

Table 19: Definitions of "experience" for system users [10]

Definitions of "experience" in context to the user of an application	**Author(s), date**
User experience involves emotion, affect and experience beyond the instrument	M. Hassenzahl and N. Tractinsky, 2006
Experience consists of meaning and history	M. Hassenzahl, 2008
Experience is the result of persons engaged in the process of creating meaning	J.H. Falk and L.D. Dierking, 2010
User experience is the experience created by a product for the person who uses it	J.J. Garrett, 2011

Augmented reality applications can be utilized by individual users, or in immersive settings where experiences are shared among multiple users. While augmented reality is available on many devices today, immersive, or shared AR experiences are still novel [12]. To reduce risk in the novel interface, a quality model has been proposed [13]. The model from that work is depicted in Figure 58. It should be noted the absence of security, as a characteristic, is a limitation of the model.

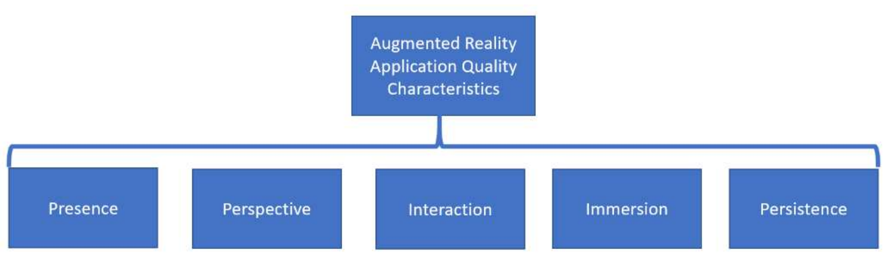

Figure 58: Quality Model for Augmented Reality Applications

The AR quality model contains five characteristics that should be discussed among product teams and system stakeholders. For the characteristics found to be most relevant, a test plan must be established to assure for the required characteristics:

- The presence characteristic helps to measure occlusion and collision between virtual and physical events. Presence, done well allows users to be present in both physical and virtual worlds simultaneously.
- Perspective is the realistic rendering of virtual objects concerning their physical counterparts so that height, width, and depth are accurately maintained.
- Interaction and the amount of it affect performance. Performance may degrade when a consumer device has limited battery life and the physical environment is hosting many virtual objects.
- Immersion is our ability to be in both physical and virtual presence at the same time. Immersion, and especially a shared immersive experience, requires portability. For many users of an augmented reality application to experience immersion, the application must be compatible across many device types and versions.
- Persistence utilizes cloud technology to enable distributed users to share anchor points and other object and scene metadata so that they can become immersed in each other's virtual and physical worlds.

A Review of Software Quality 7.1.3

In 2016, a tertiary study of 101 secondary studies was completed by Garouis and Mantyla [14]. A large disappointment in this research was that Behavior Driven Development (BDD) was not mentioned once, although it was coined by Dan North in 2006 [15], a decade before Garouis and Mantyla's research.

BDD was used in a case study published by Lubke et al. [16]. The application under test (AUT) was a web application and API that are utilized by banks in Switzerland for land registration across regions. These regions have different and complex legislation. While manual testing was utilized during the first releases of the product, as the environment and use cases became more complex, test automation was required. They incorporated modeling using Business Process Model and Notation (BPMN). They utilized BDD to give them an increased collaboration with non-technical teammates, as well as living documentation of the system. BDD uses domain-specific language and is typically written in a context-action-outcome format. Lubke et al. built a generator that reads the text behind their BPMN model and generates test automation. Five rules were utilized in their modeling and test case generation [16]:

- Test cases must describe business processes served by web services.
- Test cases must be modeled in non-technical language.
- Test cases should be automated for execution but also used to describe manual tests.
- Creating automation should be faster than the creation of manual test cases.
- Test cases must be easy to adapt as the underlying web application changes.

The generator ingests the model along with technical bindings of objects and processes as well as a skeletal test suite to create the test automation. Test models can be reused, with different bindings, to present new test cases that may be specific to versions of the web applications. Value from this approach has been in the communication of testing, the maintenance of test cases, and the accuracy of test artifacts [16]. A key benefit of using scenarios in development and test is that they support learnability, reduce fear, and enable more collaboration [16]. Another benefit of using the model is that test cases can be easily designed with non-technical team members [16].

GUI testing is system testing of a software's user interface by executing application and end-user processes [17]. Regression testing focuses on retesting the GUI after a series of code changes have been implemented to assure that previous features have not been impacted by recent change. Banerjee et al. studied 136 articles on the topic of GUI testing [17]. Themes and trends the authors found include.

- The most common metric was the number of faults detected.
- Most of the articles in the study were about new or improved testing processes.
- More than half of all articles were generating automated tests from models.
- 23 of the articles utilized record/replay for test article generation.
- Less popular methods included formal methods, AI planning, and statistical analysis.
- Test oracles are utilized to determine if a test has passed or failed. Popular test oracles found by Banerjee et al. included state reference and crash detection methods.
- 112 tools were identified in the review.

Santiago, King, and Clarke created a framework for applying AI to test automation in three phases (illustrated in Figure 59) [18]. The first phase perceives an application's objects, states, and actions. The next phase, Act, includes using a new defined language to create a model of the objects and actions. Finally, automated agents utilize computer vision to detect error messages or error states, or utilize NLP to read application logs searching for indicated failures. This detection of failures represents the observe phase. The observations form new knowledge. AI for test automation will change software testing roles, processes, and the culture in the next 5 – 10 years [18].

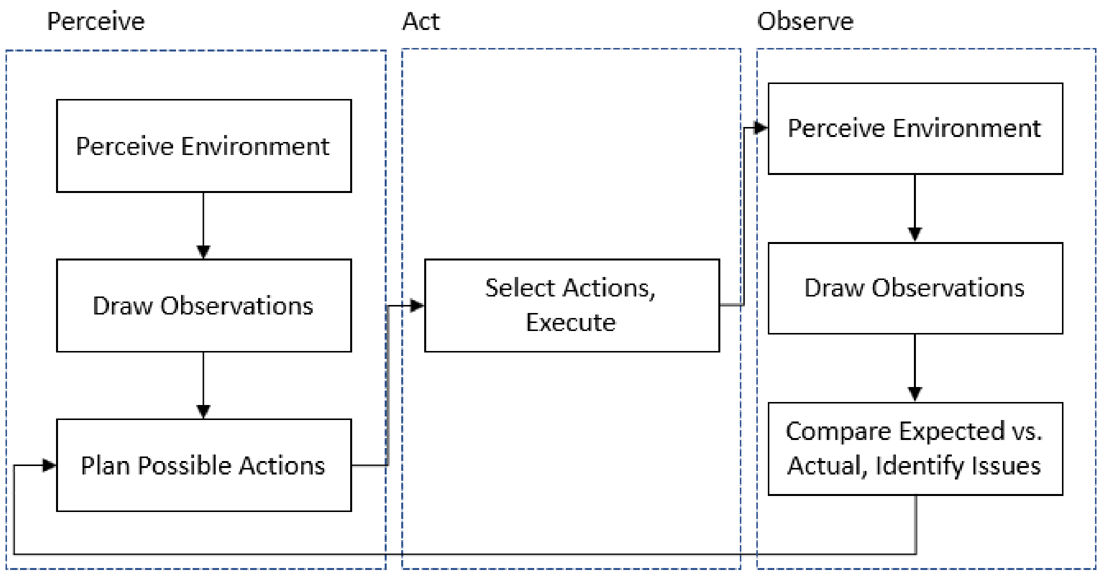

Figure 59: An Illustration adapted from Santiago, King & Clarke [18]

ISO 25010 identifies eight characteristics of software quality. Security is one of the eight characteristics. Information security and cybersecurity are global concerns with over 50 nations having published official documentation [19]. While these two terms are often used synonymously, they have differences [20]. Information and Communication Technology (ICT) security is another concept that suffers from being used in place of information security or cybersecurity but also has differences. Von Solms & Van Niekerk noted the following definitions in their work to differentiate these concepts [20].

- Securing information ensures business continuity and minimizes business damage by limiting the impacts of security incidents [21].
- Information security protects information and critical elements, including the systems and hardware that process, store, and transmit information [22]
- Merriam-Webster dictionary defines cybersecurity as “measures taken to protect a computer or computer system (as on the Internet) against unauthorized access or attack” [23].

- The International Telecommunications Union (ITU) defines cybersecurity as "Cybersecurity is the collection of tools, policies, security concepts, security safeguards, guidelines, risk management approaches, actions, training, best practices, assurance and technologies that can be used to protect the cyber environment and organization and user's assets" [24].

While there are many similarities between information security and cybersecurity, Von Solms and Van Niekerk argue that a major difference is that cybersecurity must protect people themselves [20]. Cyberbullying has become a major concern in modern society [25]. Cybersecurity should include the protection of people from cyberbullying [20]. Will autonomous cars reduce the ability to utilize a vehicle to bully or hurt others? The argument is that being the victim of a cyberbullying attack does not require that information is breached or suffers a loss of integrity. In this way, cybersecurity has ethical and cultural concerns beyond those of information and ICT security. Figure 60, an adaptation from the work of Von Solms & Van Niekerk [20], illustrates areas of overlap and differences.

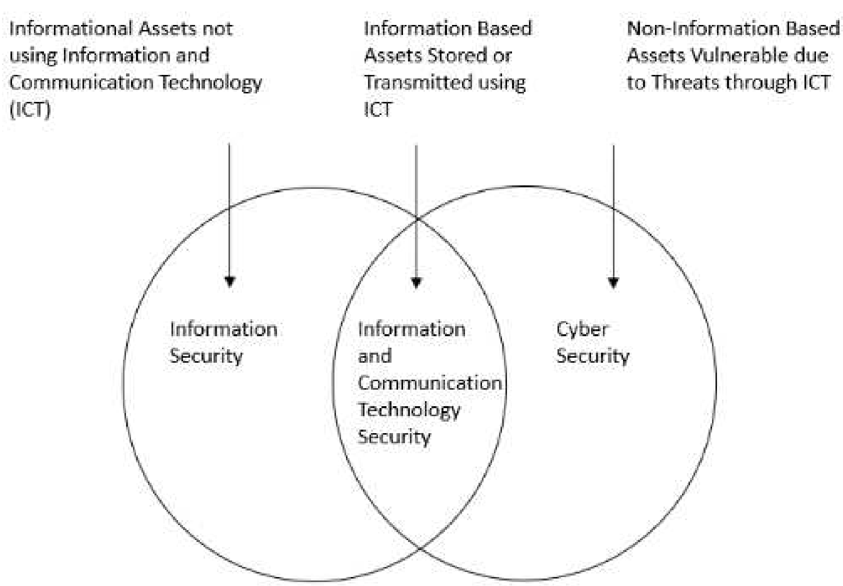

Figure 60: Venn diagram depicting asset relationships to security categories adapted from Von Solms & Van Niekerk [20]

Finding vulnerabilities in software is often the responsibility of ethical hackers and software testers. The process is frequently done in an ad-hoc procedure and is not optimized [26]. While many tools offer automation in these areas, humans are still required to find defects and vulnerabilities [27]. Aranda and Venolia found that most bugs are identified through manual testing [28]. Votipka et al. utilized a semi-structured interview study of 25 participants to identify similarities in testing and hacking methods [26]. They identified how these practitioners learn such skills and what factors limit successfully finding vulnerabilities. Practitioners report experience is the most significant factor in vulnerability finding [26]. Both hackers and testers reported that learning from formal education was not a significant source of knowledge. The four areas identified towards learning vulnerability discovery:

- Learn by doing, real-world work
- Learn by doing, exercises such as hack days
- Learn through the community
- Learn from reading published defect reports

One finding from the work of Votipka et al. is that hackers and testers utilize a similar process to find vulnerabilities [26]. The illustration below is an adaption from Votipka et al. and outlines the six process steps that were common across both testers and developers.

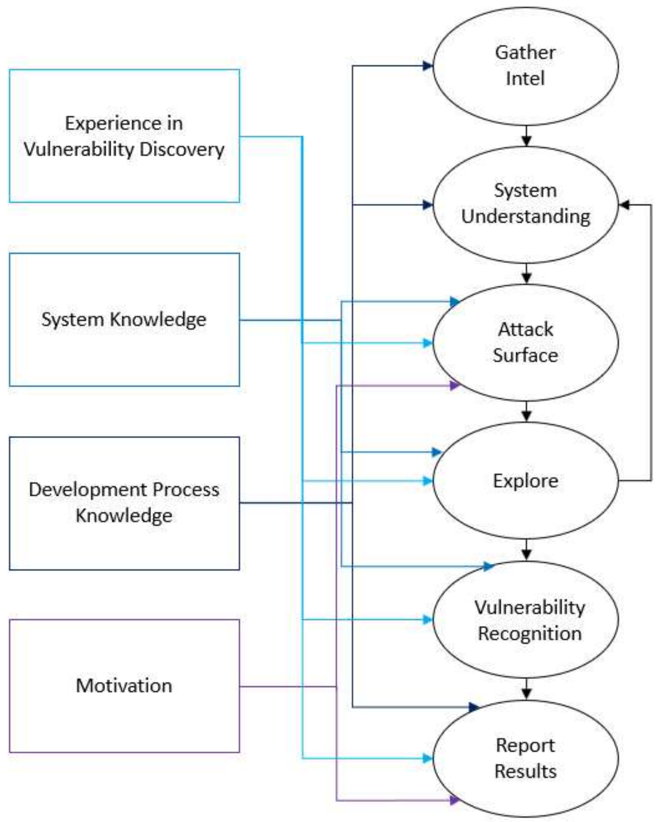

Figure 61: Six Steps for Defect and Vulnerability Discovery adapted from works of Votipka et al. [26]

To their surprise, the researchers did not find a correlation between having access to the development environment and the success of finding vulnerabilities. This may be due to hackers preferring black box techniques such as fuzz testing over static analysis.

Methodology 7.2

Data Collection 7.2.1

This research utilizes eight R scripts that poll the Twitter API once per day beginning in the Fall of 2019. Each script stores the returned tweet's text and metadata in separate MySQL databases hosted in the AWS cloud. The table below identifies the 119 search terms used in the API calls sorted by alphabetical order. The search terms were partitioned into the eight R scripts by classification, such as being a technology, commercial or open-source tool, or practice, and are labeled as such during early data quality and analysis efforts. The tweet search labels are utilized in the supervised machine learning models. Within the search terms are hashtags for topical searches, simple text, and account identifiers of users and organizations. Accounts are denoted in Twitter as starting with the '@' symbol. The search terms represent various technologies, practices, and technology products that may aid in the lifecycle and quality of digital twin systems and IoT.

Table 20: Search terms used for Twitter API fetch

1 Search Terms Utilized for Acquiring Tweets				
APITesting	#failfast	#soapui	@DevOpsOnline	@SeleniumHQ
#appium	#failforward	#softwaretesting	@eggplantio	@smartbear
#AR	#featureflags	#sparkar	@Experitest	@soapui
#arkit	#fitforpurpose	#specflow	@fbplatform	@testautomation
#arkitnews	#fitforuse	#systemsengineering	@froglogic	@testIO
#AugmentedReality	#gherkin	#tdd	@Functionize	@tricentis
#bdd	#humanfactors	#testautomation	@github	@Vuforia
#behaviordrivendevelopment	#innovationculture	#testdriven	@gitlab	@wikitude
#changeleadership	#iot	#thefutureofwork	@GoogleARCore	@yourzephyr
#ChangingYourTeam	#leadershipculture	#virtualreality	@hiptest	a/b testing
#chaosengineering	#learnfromfailures	#VR	@IBM	Bdd
#continuousdelivery	#manualtesting	#Webdriver	@Jira	behavior driven development
#continuousintegration	#MBSE	#WinAppDriver	@KarateDSL	behavior-driven development
#ContinuousQuality	#mindset	@ApacheJMeter	@mablhq	blue/green release
#continuoustesting	#MobileTesting	@AppiumDevs	@microfocusdev	blue-green release
#culturechange	#multiexperience	@AR_Maxst	@parasoft	canary release
#culturehack	#oktofail	@ArcoreGoogle	@pCloudy_	culture of learning
#culturehacking	#omnichannel	@ArrestedDevOps	@perfectomobile	data driven culture
#cultureleadership	#PsychologicalSafety	@Bitbucket	@postmanclient	people dynamics
#DevOps	#qualityculture	@broadcom	@QASymphony	shaping culture
#DigitalTwin	#radicalcandor	@cainc	@ranorex	TDD
#ergonomics	#RestAssured	@cucumberbdd	@ready_api	test driven development
#escapemodel	#Robotium	@devops_research	@robotium	test-driven development
#exploratorytesting	#Selenium	@DEVOPSINST	@saucelabs	

The API library utilized in the R scripts is the open-sourced twitteR (capital 'r') library. The image below omits authentication secrets but illustrates basic usage of the library to find and store tweet data using the search term '#digitaltwins'.

```r
library(twitteR)
library(RCurl)
library(ggplot2)
library(reshape)
library(sentimentr)
library(RMySQL)

consumer_key <- ''
consumer_secret<- ''
access_token <- ''
access_secret <- ''

register_mysql_backend("", "", "", "")
setup_twitter_oauth(consumer_key, consumer_secret, access_token, access_secret)

practicestweets <- search_twitter_and_store('#digitaltwins',
                                            retryOnRateLimit = 20, lang= "en")
```

Figure 62: Example usage of R library twitteR (capital 'R')

On the Validity of Twitter as a Data Source 7.2.2

Twitter is a social media platform that enables users to share non-peer-reviewed communication widely. To alleviate some of the concerns of utilizing tweets as a research data source, we briefly introduce other work which has found validity in using Twitter data. For example, Martín, Cutter & Li utilized Twitter data in their research of Floridians' evacuation behaviors during hurricanes Irma and Matthew [29]. They found that Twitter offered a countering age bias to surveys, where surveys tended to have an elderly bias and that Twitter is biased toward younger populations. Furthermore, the research found that Twitter was a new and valuable source of research data for certain aspects of behavior [29]. D'heer and Verdegem utilized Twitter as a data source to better understand television audiences [30]. They found that Twitter is a social media platform where users criticize without expecting to drive change [30]. D'heer & Verdegem believe the debate of utilizing social media analytics in research should focus on cultural aspects and meaning [30]. Gómez-García, Matosas-López, & Ruiz-Palmero surveyed 640 university students on their usage of social media [31]. They found that the top factor for using social media platforms was for searching for information [31]. The surveyed university students also placed importance on social media platforms for exchanging and discussing information [31]. By following groups of software engineers on the Twitter platform, Bougie et al. found that 23% of the groups' tweets were related to software engineering [32]. Among the tweets referencing the topic of software engineering, 62% were about solving software engineering problems [32]. Research by Williams found that practitioners rarely cite academic research in their online content [33]. The work of Storey et al. found that software engineers utilize social media to become informed of new technology [34]. Büchi studied communication on Twitter and found that traditional information sources, such as The New York Times, have a presence [35].

To further improve the results of our usage of Twitter as a source of data, we start our analysis with a foundation of peer-reviewed academic papers. We guide our use of Twitter data with a machine learning framework through unsupervised hierarchical clustering and content-based analysis of a literature review. Then, from the clusters of academic papers, trends, topics, and terms, we construct the foundation of the digital twin quality model. Only then do we analyze the popular and public viewpoints from Twitter data to help solve our research question regarding the characteristics of quality for digital twins.

Measurements 7.2.3

Measures such as document clustering and term frequency-inverse document frequency were taken from the literature review using natural language processing techniques and then contextualized via social media analytics. Digital twin characteristics were extracted from the literature review. To begin extracting the digital twin characteristics, the utilized research notes were clustered using a hierarchical method and visualized as a dendrogram. Then the clusters were evaluated with both term frequency and term frequency-inverse document frequency (TDIDF) measures. The dendrogram clustering of the background research notes is illustrated later.

Tweets were labeled based upon the search term categories that identified and collected the tweets. Within the Twitter platform, a Tweet may be 'liked' and 'retweeted' by Twitter users. These tweet factors will be a part of the analysis and used as factors within ML models to determine which labels influence Twitter user behavior. Digital twin quality characteristics, technologies, and development practices were evaluated for the number of mentions. Use cases and industries, such as autonomous vehicles and smart cities, were also tracked and measured. Sentiment analysis is also utilized. The sentiment analysis informs the research of how specified characteristics, technology, or practices are held in terms of negative and positive perspectives.

Empirical Analysis and Results 7.3

Clustering and TF-IDF Analysis of the Literature Review Research Papers 7.3.1

Many of the journal and conference papers utilized and referenced in the background section of chapter six and this chapter have been clustered using a hierarchical algorithm based upon the term frequency-inverse document frequency (TFIDF) of the words found within the research papers. The literature review articles were segmented into five clusters. The output of the cluster is illustrated in the dendrogram in Figure 63.

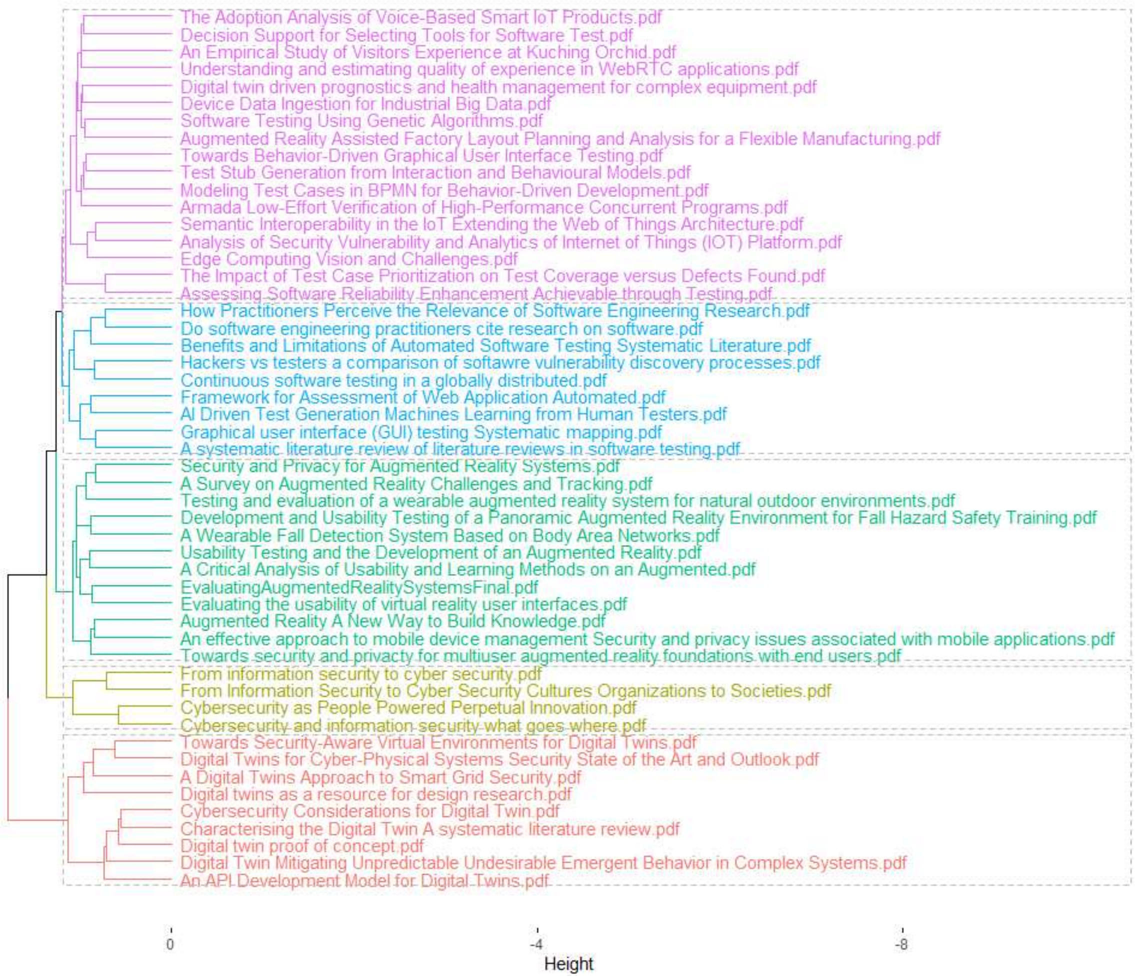

Figure 63: Literature review papers hierarchically clustered into five groups using term frequency-inverse document frequency

After the clustering, the groups' terms were sorted by the highest TFIDF scores or offering the most distinction. While attempting distinction, it did occur that some of the top terms were found within multiple documents within the same group. There may be less than 10 rows in any given facet as some terms are charted more than once, e.g., Group 2 and the term "twin", which was identified in four documents but also few enough documents to still have a compelling TFIDF score. This is illustrated in Figure 64.

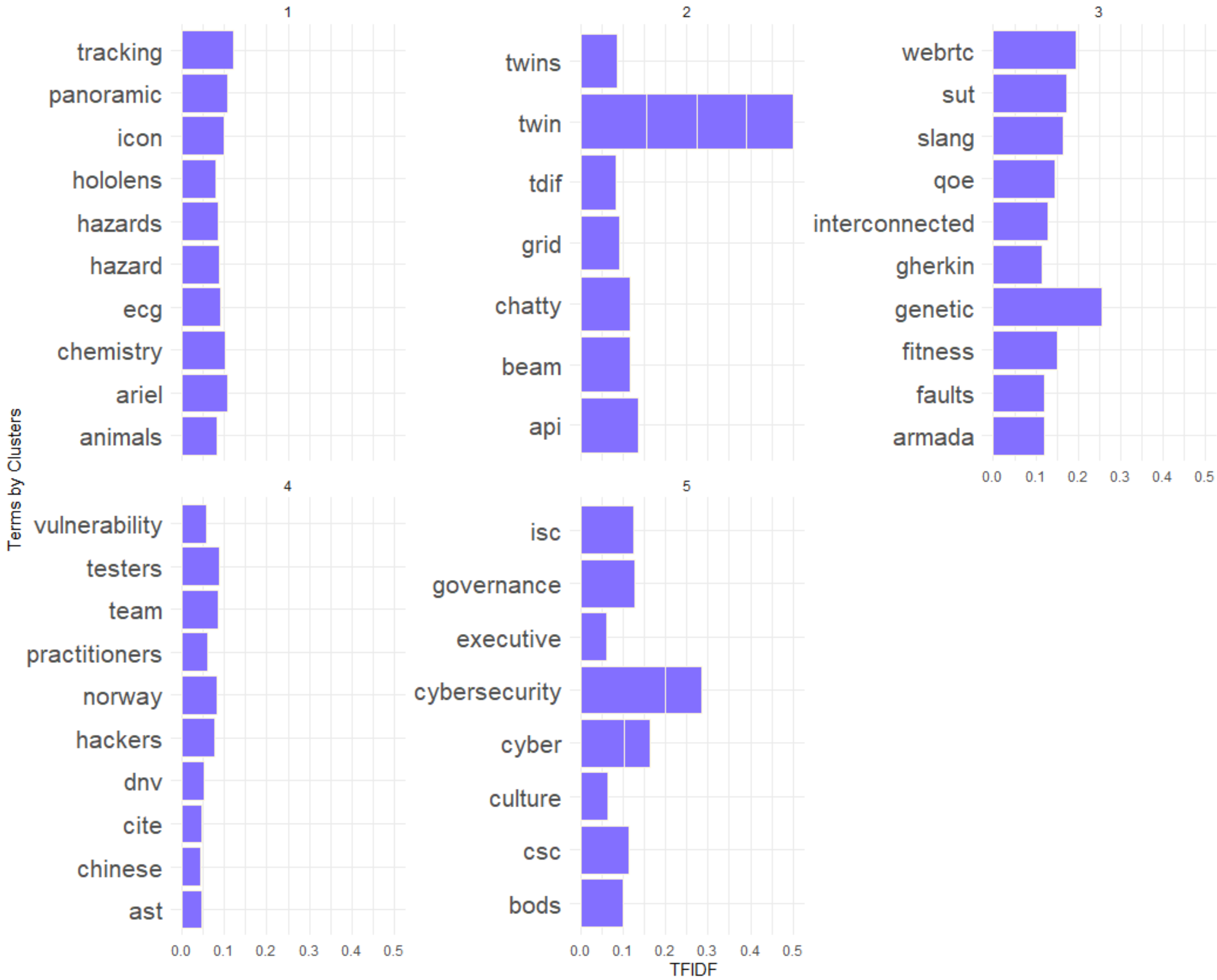

Figure 64: Facets of terms from clustered literature review papers and then weighted by term frequency-inverse document frequency

Group one terms, from Figure 64, represent use cases of IoT and digital twins for human safety. The terms “hazard” and “ECG” were related to work in protecting individuals who risk injury due to falling, such as construction crewmembers. The wearable devices and associated digital twins offer a level of monitoring and alerting in the case that a fall or vehicular accident is detected. Quality concerns, in this case, may include timeliness of communication, survivability of the IoT, and resilience. A digital twin exhibits resilience when the system maintains an appropriate state even when its dependencies become damaged. The first cluster of popular terms also uncovers the visual aspects of digital twins including the “HoloLens” technology and a “panoramic” viewpoint.

Group two includes the term "chatty". This was in the description of a dashboard that would assist designers in understanding how their prototypes or finished products were being used by consumers. Designers should utilize digital twins to understand how product use differs across cultures, locations, and user needs [4]. A digital twin dashboard would illustrate to designers how their innovative smart cities or autonomous vehicles were being used or potentially misused by consumers. This same cluster identifies the term "API", which has immense importance for digital twins. Digital twins utilize data and integrations to become models of their physical counterparts [4]. The fidelity of the digital twin to its physical counterpart is related to the integrations and the number of physical states and attributes that the digital twin can model, monitor, and predict.

The third group contains the acronym QoE for quality of experience. In the background section, we provided many definitions of "experience". Boni et al. produced five key performance indicators to determine the QoE: call establishment time, both audio and video quality, audio and video being in sync, and end-to-end delay [36]. While immersive visual quality has been discussed in the AR quality model [13], the audio was not. Having audio and video in sync could be a tenet within the characteristic towards timeliness, as could the end-to-end delay KPI. "Interconnected" appears in group three. The characteristic of interconnectedness can reinforce those of integration and fidelity. Interconnected and may also denote that a connection requires two-way communication, such as in the definition of a digital twin by Grieves [37]. "Gherkin" is a domain-specific language used within the practice of Behavior Driven Development. The term Gherkin and BDD relate to software testing. BDD was missing in mention from the literature by Garouis and Mantyla [14]. Discovering "Gherkin" as distinctive within this ML-driven review of articles may also inform new research questions in software testing practices for digital twins.

The fourth cluster of papers from the literature review shed concerns for "teams" and "practitioners". The term "hackers" that appears in this group is in the context of white hat hackers or individuals who identify vulnerabilities so that they may be remedied, and not for exploitation. While geographic terms also appear most of the terms and concepts from the fourth cluster focus on the individuals and teams involved in digital twin development. We may add to our quality characteristic the ability of a system to be testable, maintainable, and supported, as well as secured. These systems characteristics help teams support a system throughout its lifecycle.

From group five in the figure above, “executive” and “bods” (board of directors), indicate high-level concerns such as costs or risks. “Governance” is found as a distinct term within this group. The acronym “CSC” was used by Reid and Niekerk in their research to discuss cybersecurity culture [38]. “Culture” is also a term and concept that appears within this fifth cluster. From group 5, we may inform future research of digital twins towards the importance of governance and organizational culture as well as how digital twins work within society. These higher-level concerns are grouped into a characteristic of oversight. Oversight is the ability to audit a digital twin’s effect on the owning organization as well as the external impact on the environment and society it operates within.

Using machine learning methods, we have constructed a quality model for digital twins. Table 21 offers high-level characteristics and the initial creation of a bag of words to begin analyzing the public’s voice of these characteristics from our corpus of tweets.

Table 21: Research Question 1. What are the characteristics of quality for digital twins?

Characteristic	**Bag-of-words**
Twin Fidelity	"fidelity|exact|complete|integration|api|truth|true|accurate|accuracy|representation|interoperability"
Twinning Frequency	"rate|frequency|timeliness|twinning|time|late"
Resilience	"resilience|reliability|reliable|survivability|survive|mttr|mtbf|failure|repair|testability|maintainability"
Quality of Experience (QoE)	"audio|video|clarity|augmented|experience"
Oversight	“business|invest|cost|executive|expense|talent|culture|board of directors
Cybersecurity	"cybersecurity|security|information security|confidential|integrity|breach|hack|governance"

Natural Language Processing of the Tweet Collection towards the Quality of Digital Twins 7.3.2

When we analyze the 2,288,535 collected tweets using the assembled bag-of-words from Table 20, we find that 418,761 tweets contain mention of the quality characteristics we derived from our background clustering and TF-IDF analysis. Table 22 highlights the characteristics and the counts of mention.

Table 22: Digital Twin Quality Characteristics Mentions in social media

Characteristic	Mention Count
QoE	274,008
Frequency	73,669
Oversight	30,596
Fidelity	16,553
Cybersecurity	9,572
Resilience	4,196

The pie chart in Figure 65 illustrates the drastic difference in characteristic mentions between QoE and Resilience. Search terms associated with QoE encompass more counts of mention than all other characteristics combined.

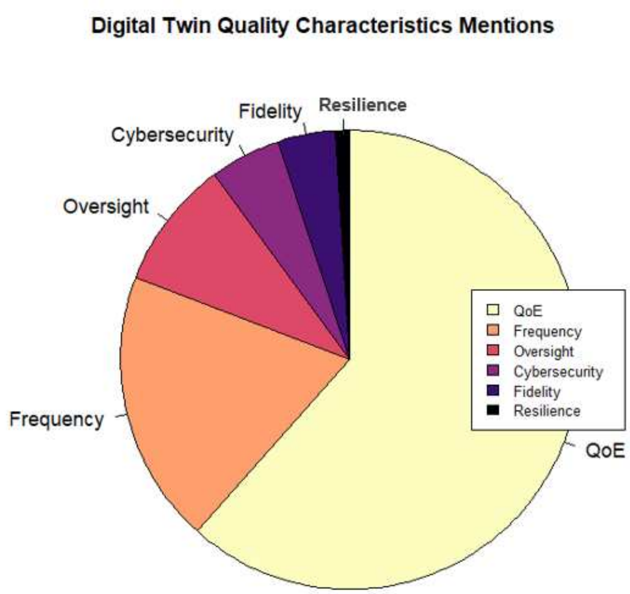

Figure 65: Pie Chart illustrating QoE as a Leading Mentioned Characteristic among the Corpus of Tweets

To judge the relevance of our derived digital twin quality model, we compare its characteristics to those of the ISO 25010 model. The comparison of characteristics between the two models is carried out by a content-based analysis of our corpus of tweets regarding digital twins.

The quality characteristics we derived from our literature review were found in 418,761 tweets. By creating a corpus of search words related to the ISO 25010 model, we found only 59,141 tweets. This is a stark difference. The following bar charts illustrate the content-based analysis between characteristics from the ISO 25010 model (in Figure 66) and mention the characteristics found within our derived quality characteristics.

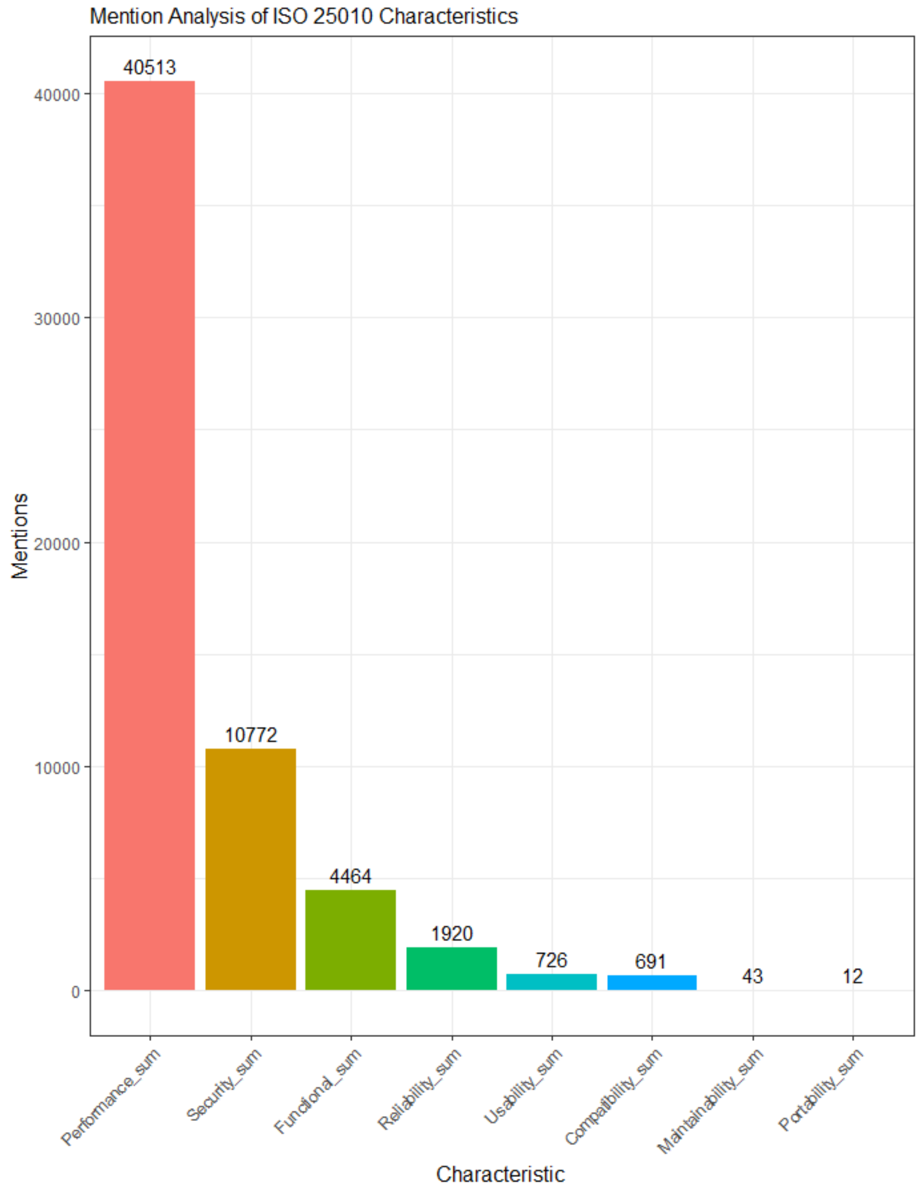

Figure 66: Bar chart of mentions of ISO 25010 characteristics within the corpus of tweets

The characteristics from our derived model have many more mentions in the corpus of tweets. The leading characteristic for mention from the ISO 25010 model was Performance which had 40,513 mentions within the corpus of tweets collected in this research effort. The second highest characteristic of our newly derived model, the twinning frequency, had nearly double the mentions with 79,079.

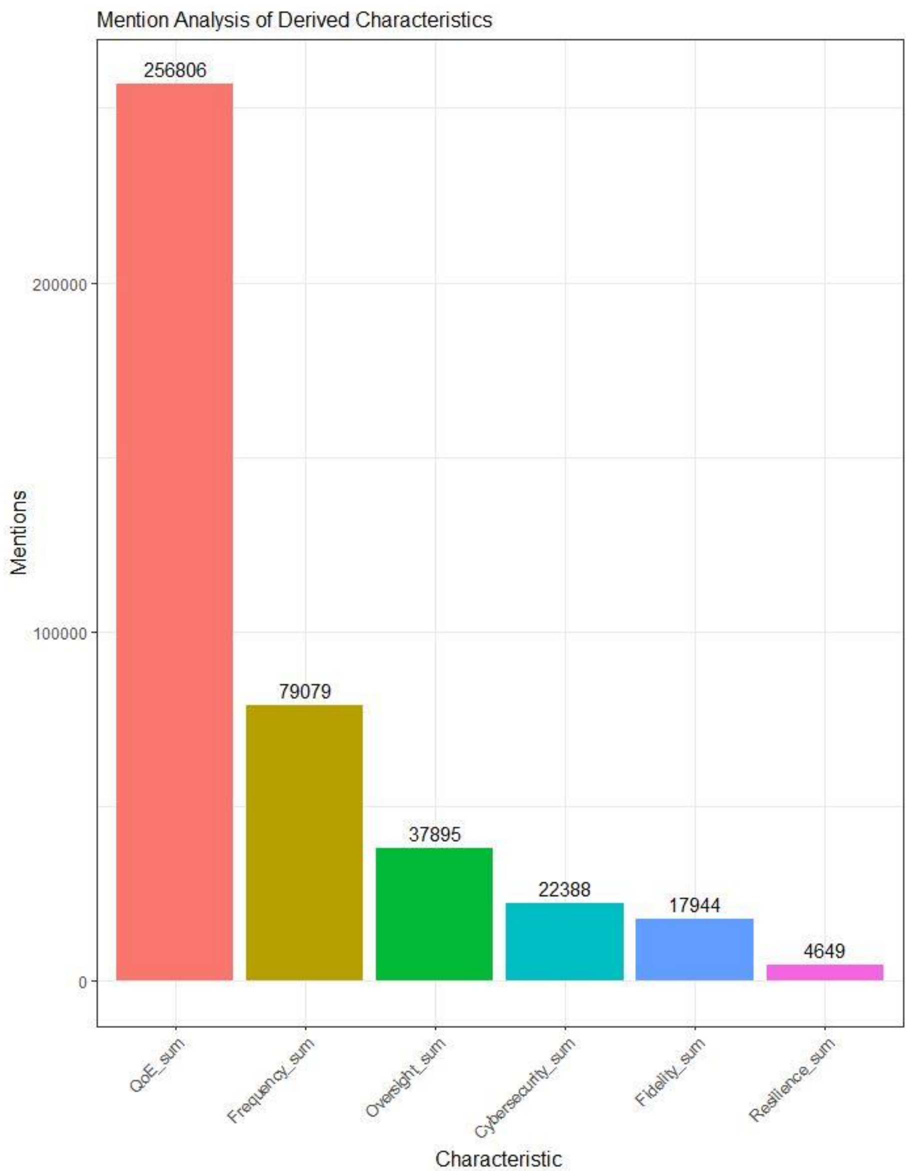

Figure 67 Mention Analysis from Research Tweet Collection referring to Derived Quality Model Characteristics

QoE had the most mentions, favorites, and retweets within the corpus of tweets. The cybersecurity category has the least positive sentiment of 0.0965 (the sentimentR package has values ranging from negative as -2 to positive as +2). The category of resilience had the most positive sentiment score, 0.2378. The least sentiment of a single tweet categorized towards the characteristic of resilience was -0.8471. Whereas the least sentiment of a single tweet within the cybersecurity category was found to be -1.0. Table 23 identifies how many tweets have been favorited, retweeted, and the average sentiment by characteristic.

Table 23: Tweet analysis of derived quality characteristics for digital twins

Category	**Sum of Favorited Tweets**	**Sum of Retweets**	**Average Sentiment**
Cybersecurity	9,132	908,963	0.0965
Oversight	24,070	1,288,657	0.1283
Frequency	74,159	4,282,937	0.1091
QoE	207,866	19,090,359	0.1074
Resilience	2,501	418,104	0.2378
Fidelity	15,133	629,468	0.1420

We performed a content-based analysis on Intelligent Transportation Systems (ITS) and other IoT use cases. A bag-of-words mention analysis was performed across the corpus of tweets for the use cases. The chart below illustrates IIoT (30,210) as the most popular discussion of use case examples within this collection of tweets, followed by autonomous vehicles (5,881), smart city (2,833), and then smart home (2,605).

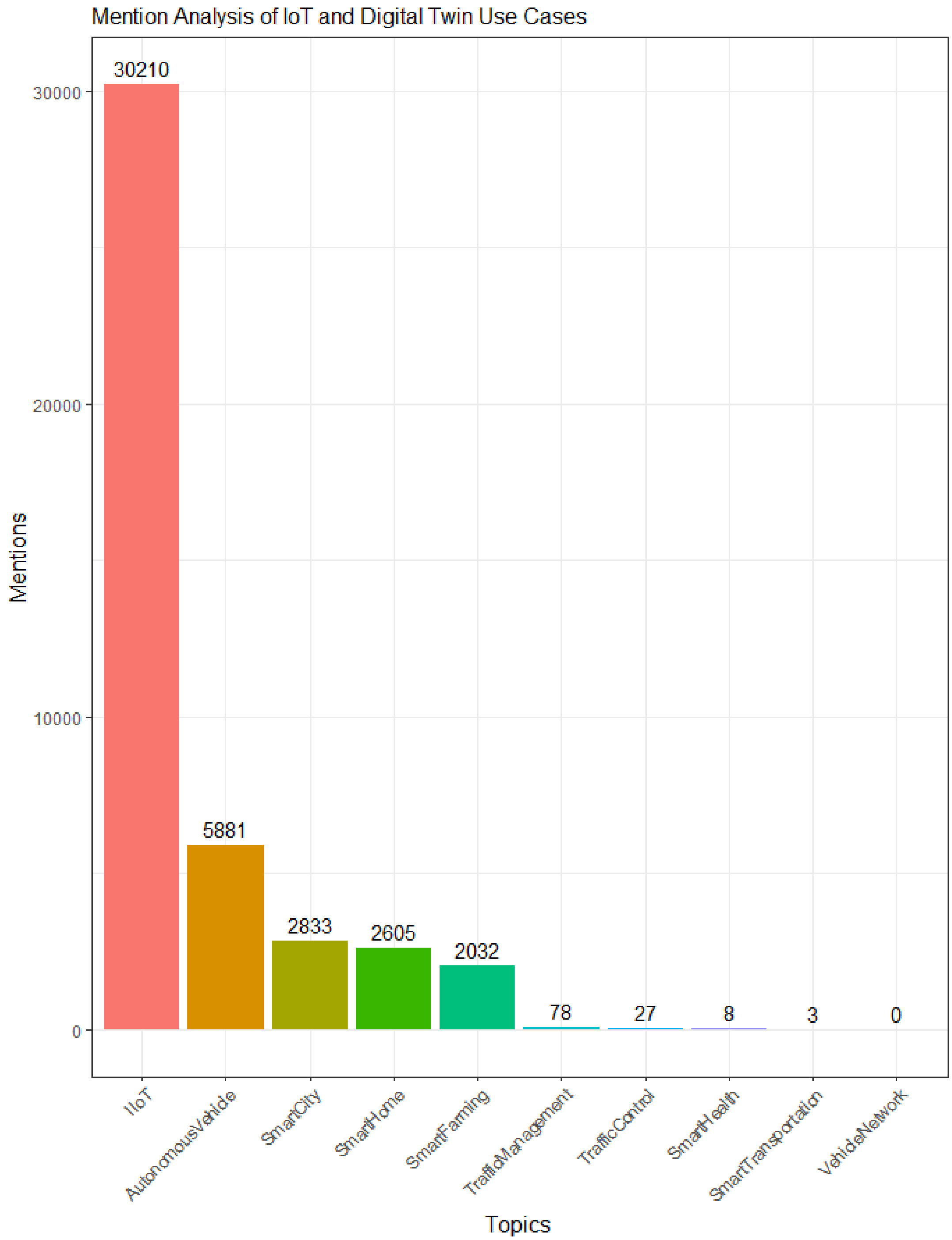

Figure 68: Bar chart of select use case mention within our corpus of tweets

Figure 68 indicates the volume of conversations about various IoT and digital twin use cases, but the chart fails to indicate trends over time. In Figure 68, the same topics' daily mentions are evaluated over a time series. For the corpus of tweets collected in this research, topics that are trending up include IIoT, Smart Health, Smart Home, and Traffic Control. The topics that with decreasing daily mention over the collection of tweets include Autonomous Vehicle, Smart City, and Smart Farming. These trends may indicate popularity of digital twin use cases in the future.

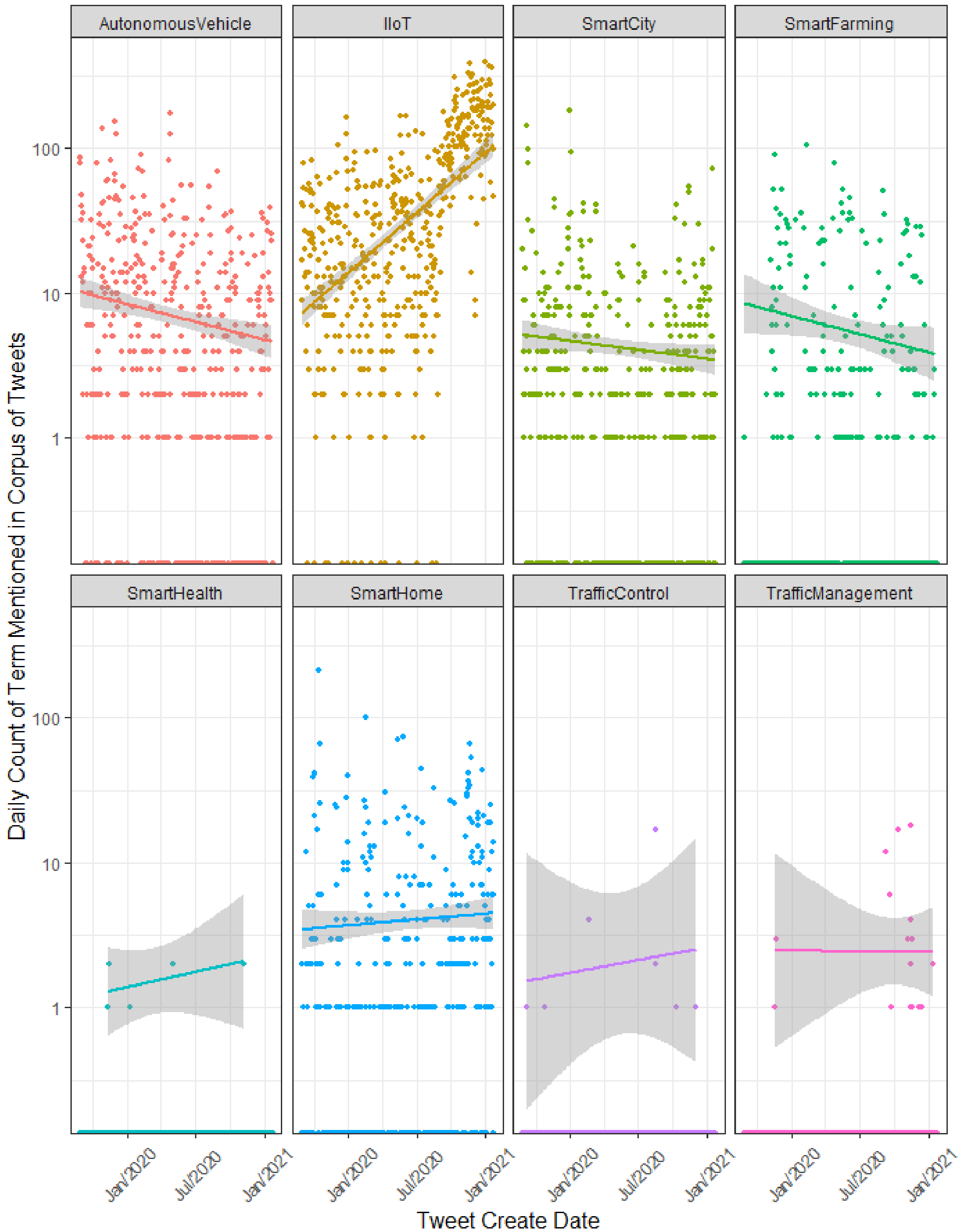

Figure 69: Facets of daily topic mention trends (grey ribbon is the 95% confidence interval)

To achieve a level of quality within a digital twin, testing must evaluate the system against our digital twin quality model. There are both testing tools and practices that may help to facilitate this testing. The following table lists testing and deployment practices and the number of times the practices are mentioned within our corpus of tweets. A large disappointment in the research by Garousi & Mantyla [14] was that Behavior Driven Development (BDD) was not mentioned once. This is a critical finding for future research questions in software testing practices. As we see in Table 24, BDD is the most popular of practices among those that were analyzed.

Table 24: Software quality, deployment and testing practices Twitter mentions

Practice	**Mentions in Corpus of Tweets**
Behavior Driven Development	624
Test Drive Development	570
Chaos Engineering	441
Feature Flags/Toggles	161
Blue-Green Deployment	75
Canary Release	20

Rafi et al. utilized both a systematic literature review and a survey of industry testers to evaluate the benefits and challenges of test automation [39]. 86% of respondents agreed that the reusability of test cases is a benefit of test automation. 75% agree that test automation results in better test coverage and improved product quality [39].

We analyzed the corpus of tweets for mention and sentiment analysis of many open source and commercial test automation products. A limitation of this analysis is that more test automation tools exist than we evaluated within this research. Banerjee et al. studied 136 articles on the topic of GUI testing and found 112 different test automation tools mentioned in their research [17]. We calculated the sentiment z-score for each test automation technology. We plotted the commercial and open-source technologies into a quadrant. The number of mentions for each test automation technology is plotted on the y-axis. The y-axis uses a logarithmic scale to help compress the height of the plot. The sentiment is plotted on the x-axis by z-score. The upper right quadrant in Figure 70 indicates leading tools by scoring high in mention and above-average sentiment.

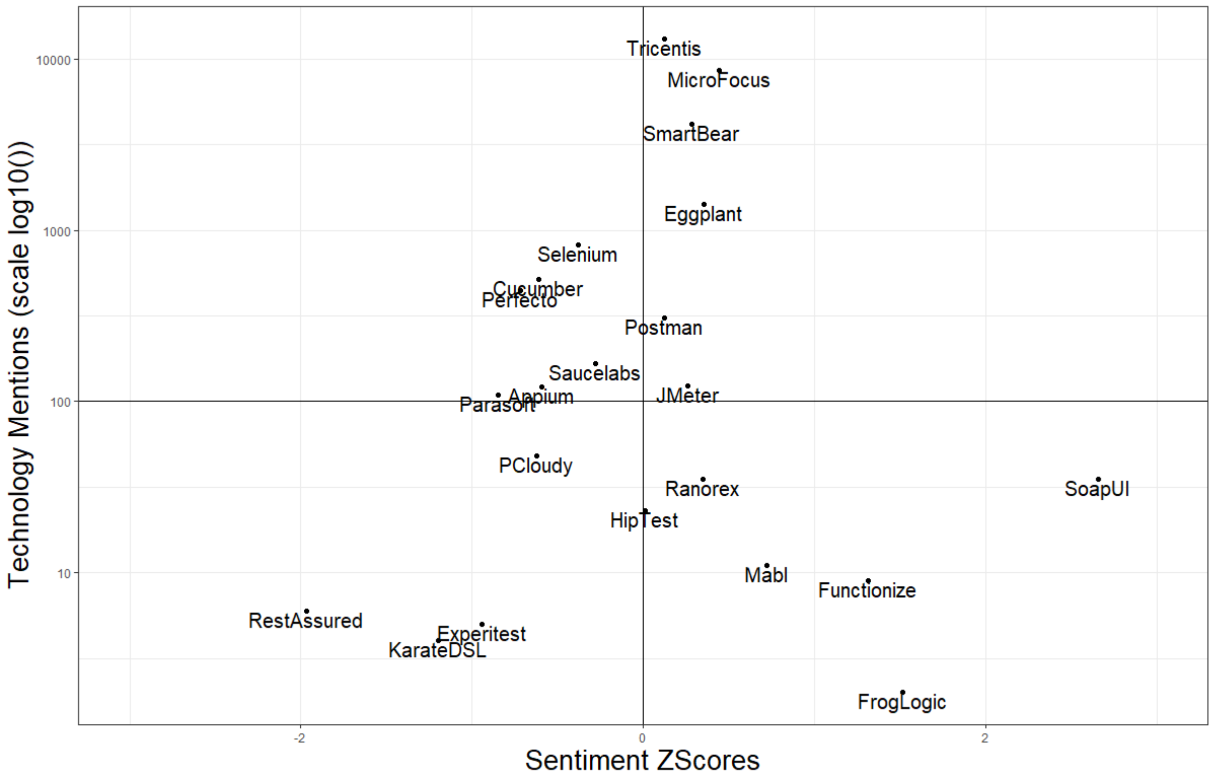

Figure 70: Open source and commercial software testing tools plotted in quadrant by Twitter mention and sentiment

Tweets can inform and can be acted upon by Twitter users, such as by liking a tweet or by retweeting the tweet. We utilized a naïve Bayes model to predict whether tweets would be favorited or retweeted. Given a tweet was either liked or retweeted, there was a 51.1% probability that the tweet was labeled as referring to the topic of augmented reality. Tweets that referenced the concept of DevOps had the next highest probability of being liked or retweeted at 41.1%. The model achieved 89% accuracy. By modeling the tweets having IIoT or ITS mentions and given the tweet had either been favorited or retweeted, the most likely tweet topic was smart homes (32.%) followed by autonomous vehicles (30.1%), illustrated in Table 25.

Table 25: Conditional probabilities of IIoT and ITS tweet topics and being a favored or retweeted tweet

	Tweet Topic								
Favorited / Retweeted	Autonomous Vehicle	IIoT	Smart City	Smart Farming	Smart Health	Smart Home	Smart Transportation	Traffic Control	Traffic Management
No	0.2464	0.2319	0.1812	0.0290	0.0072	0.2609	0.0072	0.0217	0.01449
Yes	0.3007	0.1384	0.1542	0.0587	0	0.3245	0.0011	0.0081	0.01432

Conclusions, Limitations, and Future Research 7.4

There are many driving factors for the growth in research and utilization of digital twin technology. The recent pandemic and the need for further digital touchpoints will not slow this trend. While there is extensive research covering the definitions and benefits of digital twins, research covering the quality and evaluation of them is limited.

Our research utilized natural language modeling of both academic research papers and a collection of tweets. Quality characteristics of digital twins were first distilled from the academic work and then sought for relevance across the social media discussions. We found 7.5 times more conversations of characteristics in our newly derived digital twin quality model than mentions of characteristics found in the ISO 25010 model. The QoE characteristic was the most popular in social media discussions. The characteristic of resilience had the highest sentiment. On average, no characteristic had average sentiment scores approaching the high side of positive (a score of 2) nor the low end of the negative (-2). Autonomous vehicles (5,517) were discussed in the collection of tweets more than double compared to discussions of smart cities (2,593) or smart homes (2,338). Topics trending in mention over time-series analysis included IIoT, Smart Health, Smart Home, and Traffic Control. By modeling the tweets having IIoT or ITS mentions using naïve Bayes, and given the tweet has either been favorited or retweeted, it is most likely that the given tweet would be about smart homes (32.%) followed by autonomous vehicles (30.1%).

This research is limited in its approach as it analyzes research and social media without direct evidence from a case study. Only one microblogging platform, Twitter, was utilized in the data collection. The research does not answer questions surrounding measurements that may be taken to evaluate the quality of a digital twin. Future research should case study these characteristics as guide rails in the test and evaluation plans of digital twins.

REFERENCES

[1] Negri, E., Fumagallia, L., & Macchia, M. (2017). A review of the roles of digital twin in cps-based production systems. *In Proceedings of the 27th International conference on Flexible Automation and Intelligent Manufacturing* 939–948.

[2] Scheibmeir, J., & Malaiya, Y.K. (2020). Multi-model security and social media analytics of the digital twin. *ASTEJ.,* 5, 323-330.

[3] Grieves, M. (2016). *Origins of the Digital Twin Concept*. https://doi.org/10.13140/RG.2.2.26367.61609

[4] Burnett, D., Thorp, J., Richards, D., Gorkovenko, K., & Murray-Rust, D. (2019). Digital twins as a resource for design research. *In Proceedings of the 8th ACM Int. Symp. Pervasive Displays. PerDis*., 6–7. https://doi.org/10.1145/3321335.3329685

[5] Daily, J. & Peterson, J. (2017). *Predictive Maintenance: How Big Data Analysis Can Improve Maintenance.* Springer International Publishing. https://doi.org/10.1007/978-3-319-46155-7_18

[6] Sutherland, I.E. (1968). A head-mounted three-dimensional display. *In Proceedings of the Fall Joint Computer Conference, American Federation of Information Processing Societies*.

[7] Caudell, T., & Mizell, D. (1992). Augmented reality: an application of heads-up display technology to manual manufacturing processes. *In Proceedings of the Twenty-Fifth Hawaii International Conference on System Sciences.*, 2:659-669. https://doi:10.1109/HICSS.1992.183317

[8] Kato, H., & Billinghurst, M. (1999). Marker tracking and HMD calibration for a video based augmented reality conferencing system. *In Proceedings of the 2nd IEEE and ACM International Workshop on Augmented Reality (IWAR 99)*, 85-94.

[9] Mohring, M., Lessig, C., & Bimber, O. (2004). Video see-through AR on consumer cell phones. *In Proceedings of the 3rd IEEE/ACM international Symposium on Mixed and Augmented Reality*, 252-253.

[10] Othman, M.K., Idris, K.I., Aman, S., & Talwar, P. (2018). An empirical study of visitors' experience at kuching orchid garden with mobile guide application. *Adv. Human-Computer Interact.* https://doi.org/10.1155/2018/5740520

[11] Pallud, J. (2017). Impact of interactive technologies on stimulating learning experiences in a museum. *Information & Management*, 54(4), 465–478. https://doi.org/10.1016/j.im.2016.10.004

[12] Lebeck, K., Ruth, K., Kohno, T., & Roesner, F. (2018). Towards security and privacy for multi-user augmented reality: foundations with end users. *In Proceedings - IEEE Symposium on Security and Privacy, Institute of Electrical and Electronics Engineers Inc.,* 392–408. https://doi.org/10.1109/SP.2018.00051

[13] Scheibmeir, J., & Malaiya, Y.K. (2019). Quality model for testing augmented reality applications. *In Proceedings of the 10th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference*, *IEEE*, 219-226. https://doi/10.1109/UEMCON47517.2019.8992974

[14] Garousi, V., & Mantyla. M.V. (2016). A systematic literature review of literature reviews in software testing. *Information and Software Technology*, 80:195-216. https://doi.org/10.1016/j.infsof.2016.09.002

[15] North, D. (2006). "Behavior Modification". *Better Software Magazine* 2006;03.

[16] Lubke, D. & Van Lessen, T. (2016). Modeling test cases in BPMN for behavior-driven development. *IEEE Softw.*, 33, 5, 15–21. https://doi.org/10.1109/MS.2016.117

[17] Banerjee, I., Nguyen, B., Garousi, V., & Memon, A. (2013). Graphical user interface (GUI) testing: systematic mapping and repository. *Information and Software Technology*, 55:1679–1694. https://doi.org/10.1016/j.infsof.2013.03.004

[18] Santiago, D., King, T.M., & Clarke,P.J. (2018). AI-Driven test generation: machines learning from tuman testers. *In Proceedings from Pacific NW Software Quality Conference*

[19] Klimburg, A. (2012). *National Cyber Security Framework Manual*, NATO CCD

[20] Von Solms, R., & Van Niekerk, J. (2013). From information security to cyber security. *Comput. Secur.,* 38, 97–102. https://doi.org/10.1016/j.cose.2013.04.004

[21] Von Solms, R. (1998). Information security management (3): the code of practice for information security management. *Information Management & Computer Security* 1998;6(5):224-225.

[22] Whitman, M.E., & Mattord, H.J. (2009). *Principles of information security*. 3rd ed. Thompson Course Technology.

[23] Merriam-Webster. *Cybersecurity*. Retrieved June 24, 2021, from https://www.merriam-webster.com/dictionary/cybersecurity

[24] International Telecommunications Union. *Cybersecurity*. Retrieved June 24, 2021, from https://www.itu.int/en/ITU-T/studygroups/com17/Pages/cybersecurity.aspx

[25] Martin, N.J., & Rice, J.L. (2011). Cybercrime: understanding and addressing the concerns of stakeholders. *Computers & Security*, 30, 803-814.

[26] Votipka, D., Stevens, R., Redmiles, E., Hu, J., & Mazurek, M. (2018). Hackers vs. testers: a comparison of software vulnerability discovery processes. *In Proceedings - IEEE Symposium on Security and Privacy, IEEE Inc.*, 374–391. https://doi.org/10.1109/SP.2018.00003

[27] Shoshitaishvili, Y., Weissbacher, M., Dresel, L., Salls, C., Wang, R., Kruegel, C., & Vigna, G. (2017). Rise of the HaCRS: augmenting autonomous cyber reasoning systems with human assistance. *In Proceedings of the 24th ACM SIGSAC Conference on Computer and Communications Security*, *ACM*.

[28] Aranda, J., & Venolia, G. (2009). The secret life of bugs: going past the errors and omissions in software repositories. *In Proceedings of the 31st International Conference on Software Engineering, IEEE Computer Society,* 298–308.

[29] Martin, Y., Cutter, S.L., & Li, Z. (2020). Bridging twitter and survey data for evacuation assessment of hurricane matthew and hurricane irma. *Natural Hazards Review*, 21(2), https://doi.org/10.1061/(ASCE)NH.1527-6996.0000354

[30] D'heer, E., & Verdegem, P. (2015). What social media data mean for audience studies: a multidimensional investigation of Twitter use during a current affairs TV programme. *Information, Communication & Society*, 18:2, 221-234. http://doi:10.1080/1369118X.2014.952318

[31] Gomez-Garcia, M., Matosas-Lopez, L., & Ruiz-Palmero, J. (2020). Social networks use patterns among university youth: the validity and reliability of an updated measurement instrument. *Sustainability*. 12(9), 3503–. https://doi.org/10.3390/su12093503

[32] Bougie, G., Starke, J., Storey, M., & German, D.M. (2011). Towards understanding twitter use in software engineering: preliminary findings, ongoing challenges and future questions. *In Proceedings of the 2nd International Workshop on Web 2.0 for Software Engineering. ACM,* 31–36. https://doi.org/10.1145/1984701.1984707

[33] Williams, A. (2018). Do software engineering practitioners cite research on software testing in their online articles?: a preliminary survey. *ACM*, 151–156.

[34] Storey, M., Treude, C., Van Deursen, A., & Cheng, L. (2010). The impact of social media on software engineering practices and tools. *In Proceedings of the FSE/SDP workshop on Future of software engineering research. ACM.*, 359–364.

[35] Büchi, M. (2017). Microblogging as an extension of science reporting. *Public Understanding of Science.* 26(8), 953–968. https://doi.org/10.1177/0963662516657794

[36] Garcia, B., Gallego, M., Gortazar, F., & Bertolino, A. (2019). Understanding and estimating quality of experience in WebRTC applications. *Computing 101*, 11, 1585–1607. https://doi.org/10.1007/s00607-018-0669-7

[37] Grieves, M. (2015). Digital Twin: Manufacturing Excellence through Virtual Factory Replication.

[38] Reid, R., & Van Niekerk, J. (2014). From information security to cyber security cultures. *In Proceedings of the 2014 Information Security for South Africa Conference, IEEE*., https://doi.org/10.1109/ISSA.2014.6950492

[39] Rafi, D.M., Reddy, K., Moses, K., Petersen, K., & Mäntylä, M.V. (2012). Benefits and limitations of automated software testing: Systematic literature review and practitioner survey. *In 2012 7th International Workshop on Automation of Software Test, AST 2012 - Proceedings, IEEE*., 36–42. https://doi.org/10.1109/IWAST.2012.6228988

CHAPTER 8: CONCLUSION AND FUTURE WORK

Digital twins are a valuable technology that enable organizations to optimize processes and resources. Digital twins and emerging technologies including augmented reality and the IoT still have aspects of novelty, which implies risk, driving the need for quality practices. This research has identified many aspects of quality and maturity, including helpful commercial and open-source technologies. Emerging technologies also drive conversations on Twitter, which has been a useful data source to gauge the voice of stakeholders. Contributions of this research and areas where further investigation is needed follow.

Research Contributions

This research has focused on the quality attributes of digital twins. Digital twins help organizations monitor, model, and predict their physical counterparts. The contributions of this research are as follows:

- Produced the API Development Model for Digital Twins
- Assembled the quality characteristics for augmented reality applications
- Developed a framework that utilizes computer vision to detect failed presence and perspective characteristics in AR apps
- Constructed the IoT Cybersecurity Communication Scorecard
- Defined six layers of maturity for Digital Twins
- Distilled quality characteristics for Digital Twins using machine learning methods.

Emerging technologies, such as augmented reality and the internet of things, are utilized by digital twin systems. Integrations to IoT help to inform the digital twin instance and to provide volumes of data that assist in modeling. Integrations with immersive technologies such as augmented reality provide humans in the loop with visibility into self-optimizing systems. This research presented an API development model for digital twins. The API development model for digital twins provides a framework for creating integrations with a focus on their maintainability, performance, security, and functional correctness. These are accomplished with practices such as test-driven development, performance engineering, and API contract testing, among others.

The API development model provides a framework for digital twin integration up to deployment and operations using a mediation layer. In this way, it was limited in that it did not approach the user interface. Augmented reality technology has been suggested as a user interface for digital twins. Chapter two presents a software quality model for augmented reality applications. The quality model is constructed with five characteristics that should be present and tested for in high quality AR applications: presence, perspective, interaction, immersion, and persistence. Another novel contribution in this dissertation is the application of computer vision as an oracle to detect when the presence and perspective characteristics are defective in AR applications.

Digital twins provide a blueprint of cyber physical systems and their integrations. In the wrong hands, this information would empower malicious activities. For these reasons, digital twins should be secured systems and should be utilized to help increase the security posture of their physical counterpart. To help organizations understand a small aspect of their security posture, we introduced the IoT cybersecurity communication scorecard. This scorecard uses two factors, communication density and sentiment, to provide an IoT cybersecurity communication benchmark per industry segment.

The final two areas of contribution involve the quality and maturity of digital twins. First, we provided a maturity model for digital twins consisting of six layers. Each layer helps organizations plot their as-is and to-be roadmaps for digital twin investment and value acquisition. The layers include the initial, managed, integrated, immersive, autonomous, and ubiquitous digital twins. Second, we utilized a novel machine learning approach of distilling quality characteristics from literature review and cross examined against social media analytics. Organizations who are investing in digital twins should evaluate their systems for QoE, frequency of twinning, fidelity, oversight, and cybersecurity posture.

Beyond the novel approaches, models, and frameworks presented in this research, many analyses were performed across social media regarding augmented reality, IoT, digital twins, and related vendors and technologies.

This work analyzed over two hundred thousand tweets, collected during 2019 and 2020, relating to augmented reality technology. The education industry had the highest mentions among the augmented reality tweets. The tweets labeled towards the transportation industry had the highest sentiment.

Social media conversations on the topic of digital twins were also reviewed. Only 10% of the first collection of digital twin tweets (n=3,102) had any mention of cybersecurity concerns. This was concerning as the industry most mentioned within that collection of tweets was the healthcare industry. Information confidentiality and integrity are certainly concerns for the healthcare industry, despite it not being a frequent discussion in the collection of tweets. The top three industries having the highest tweet sentiment were engineering, commerce, and energy. A larger corpus of tweets was assembled to use in a machine learning model. In total, 422,963 industrial internet of things (IIoT) tweets and 497,174 cybersecurity tweets were collected. A naïve Bayes model reached a 70.3% accuracy at differentiating a tweet that was about cybersecurity versus a tweet about digital twins.

We evaluated 684,503 contemporary tweets on the topic of the Internet of Things to shed light on public opinion, technology trends, popular industry usage and the popularity and sentiment of technology providers in this space. The number of tweets mentioning cybersecurity concepts rose slightly to 12%. Again, it was found that the most discussed industry within the tweets was healthcare. In reviewing trends within the tweets, no cybersecurity concepts were identified in the top ten trends. The top three trends identified within the IoT tweets were data science, machine learning and big data. We performed a network analysis to identify relationships between trends and industries, such as what industries have the greatest or least inclusion of trending concepts and technology. The shipping and postal/telecom industry tweets had the fewest relationships to identified trends.

More charts and analyses are available in the previous chapters. Results covering content based, sentiment, and network analysis are available. The R code and raw data utilized in chapter five is available to the public in Gitlab repository and listed in that chapter's references.

Future Research Areas

Not all aspects of the quality of digital twins were pursued or provided answers for within this research. Future areas of research should include the security of AR applications. AR applications will utilize a device camera to help blend the digital models into the physical world. An AR application that is vulnerable would offer malicious actors access to the device camera and the user's surroundings. While AR was discussed within this research as a possible interface for digital twins, there are other immersive technologies. Further research should be done on the application of haptic experiences and wearable technologies as interfaces to digital twins. While digital twins are frequently considered for industries such as manufacturing or use cases such as industry 4.0, smart devices, or autonomous vehicles, research should be pursued on the application of digital twins for natural entities such as plants and forests. Another aspect of requiring further research is the inclusion of distributed ledgers and blockchain technologies to help inform and secure the integrity of digital twin technologies.

Other areas of future research include reviewing the social media discussions and how the voice of stakeholders regarding augmented reality, IoT, and digital twins change over time. Cybersecurity conversation was found in very few of the IoT (12%) and digital twin (10%) related tweets. It should be determined how time, regulation, technology improvements, and breach events alter the percentage of these conversations over time. Future research should also review which industries are having the most mention, such as whether the education industry continues to be mentioned the most within tweets regarding augmented reality. Trends should also be identified in future research and compared to the trends that were identified in the conversations we reviewed in this research. For example, within the IoT related tweets, the top three trends were data science, machine learning, and big data. In chapter five, we proposed the IoT Cybersecurity Communication Scorecard. Future research should analyze the social media landscape and regenerate the scorecard, identifying new benchmarks.