JOURNAL OF INFORMATION SYSTEMS APPLIED RESEARCH

Volume 14, Issue. 4
December 2021
ISSN: 1946-1836

In this issue:

iscap
Information Systems & Computing Academic Professionals

The **Journal of Information Systems Applied Research** (JISAR) is a double-blind peer reviewed academic journal published by ISCAP, Information Systems and Computing Academic Professionals. Publishing frequency is three to four issues a year. The first date of publication was December 1, 2008.

JISAR is published online (https://jisar.org) in connection with CONISAR, the Conference on Information Systems Applied Research, which is also double-blind peer reviewed. Our sister publication, the Proceedings of CONISAR, features all papers, panels, workshops, and presentations from the conference. (https://conisar.org)

The journal acceptance review process involves a minimum of three double-blind peer reviews, where both the reviewer is not aware of the identities of the authors and the authors are not aware of the identities of the reviewers. The initial reviews happen before the conference. At that point papers are divided into award papers (top 15%), other journal papers (top 30%), unsettled papers, and non-journal papers. The unsettled papers are subjected to a second round of blind peer review to establish whether they will be accepted to the journal or not. Those papers that are deemed of sufficient quality are accepted for publication in the JISAR journal. Currently the target acceptance rate for the journal is about 40%.

Questions should be addressed to the editor at editor@jisar.org or the publisher at publisher@jisar.org. Special thanks to members of ISCAP/EDSIG who perform the editorial and review processes for JISAR.

JOURNAL OF INFORMATION SYSTEMS APPLIED RESEARCH

Integration of Information Systems: Robotic Process Automation

William H. Money
wmoney@Citadel.edu
Management and Entrepreneurship Department
The Citadel
Charleston, South Carolina 29409 USA

Abstract

This paper offers an overview of Robotic Process Automation (RPA). It approaches this technology from the perspective of systems integration, and the presently suggested value of Robotic Process Automation (RPA) in addressing information systems integration issues. The background requirements for organization process integration are discussed, and the methods organizations employ to achieve system integration are reviewed. RPA is described and its application and suggested benefits are summarized. The current literature indicates that RPA technology has been focuses primarily on the transactional processes that occurs between routinized and repetitive business processes and back office work that are performed by different information systems or manual follow-on processes. The future of RPA has been hypothesized to include bots that learn and implement analytical processes, and complex work steps requiring more reasoning. Significant research is needed to understand the benefits of RPA, and its growing popularity in organizations.

Keywords: Integration, Automation, Robotic, Processes, Value Chain Integration.

1. INTRODUCTION

This paper discusses a relatively new technology (Robotic Process Automation) that appears to be an innovations solution for addressing some of the critical organizational and business problems that require systems integration or manual support after systems processes are completed. This technology may be applied both within and among organizations. The RPA technology appears (inductively) to address a significant problem fort many organizations – integrating information systems work processes.

Integration has long been a critical concern for organizations. March and Simon (1958:195) defined the level of integration as "…lowest level at which all activates relating to a particular goal can be coordinated…" Thompson (1967:40 - 41) described integration (coordination of successive stages of production occurring in variety of fields) as rational behaviors designed to reduce crucial contingencies. The evolution of the integration concept is expanded below.

Integration has also been a significant factor in the development and implementation of information systems. Brancheau and Wetherbe (1987) described integration as one of the top ten issues in information systems management in both 1984 and 1987. Hasselbring (2000) discussed the importance and difficulties with integrating heterogeneous information systems to legacy systems, and inter-organizational processes that utilize information systems are highly autonomous. It is noted that this makes the integration process an even more challenging task.

Hasselbring (2000) explains that there are both external and internal drivers for integration. The external difficulties associated with system can be traced to the value chains that extend beyond organizational borders. This occurs because external supplier and customer

information systems become linked as critical parts of each other's data and information architectures. From an internal perspective, organization units contain data in many application areas supported by a wide variety of workstation level tools, database, and collection tools that seek to share data among desperate, and organizationally autonomous information systems.

Exchanging data among these differentiated systems and the organizational units they support and passing essential data to other units to complete work processes is often complex, time consuming, and very costly. Information systems that are separated by departmental or other formal organizational decision making lines with distinct process and different but equally important business purposes are difficult to combine. The integration situation is alarming as organizations grow, combine, develop new products and services and the amount of work increases. Possible exacerbating circumstances include budgets decreases, and the unit processes morphing to involve more steps and procedures to ensure quality and accuracy. Greater labor costs, longer unit processes, increased governance, and pressures for high accuracy create a recipe for a disastrous planning, budgeting, and management brew.

There are other proposed method for addressing this concern. For example, Simon, Karapetrovic, and Casadesús (2012) developed and proposed a model to address the difficulties and recognize the benefits of standardized management systems. They assessed the level of integration of different MS elements such as the resources, documentation, goals and procedures. They conclude that managers and practitioners become aware of the challenges and obstacles of systems integration, and address them early in the process so they do not delay the completion of the integration process.

Integration demanded by organizations requires extensive coordination of shared resources and dependencies among activities in and across systems. A variety of information systems technical and functional solutions to this problem have been developed over the past 30 years. Since about 2010 a relatively new technology (Robotic process Automation (RPA) has become a major factor in improving systems integration.

This paper describes the integration problem, various enacted solutions and their limitations, and the status of the RPA technology available today. Form a research perspective, the work done on RPA is somewhat limited and focuses on case studies. Key question such as – when is the RPA technology appropriate, what criteria should be considered before adoption and implementation of the technology, and what are the limitations have not been explored well. This paper seeks to provide a discussion of the issues regarding the technology, and suggest why researching and developing a comprehensive understanding of RPA technology is important.

The popular literature and case studies indicate selection of this technology (over other approaches) seems to be based on assessment of the criteria driving the integration effort. These criteria are summarized and offered as a set of heuristics for the use of this new technology at the end of this paper.

Integration Background

Difficulties with the integration of business process has plagued organizations for many years. Information systems have been "inserted" into this essential mission of attempting to manage processes and their data, and improve the coordination and integration of the work within and between organizations.

The linkages between information systems and the integration mission can be traced back to Porter and Millar's (1985). They discussed the (then coming) role that information technology plays in the value chain and competitive advantages sought by organizations. They state that information technology must be viewed as encompassing information that businesses create as well as the many linked technologies that process the information. Their seminal work identified many examples of value creating activity that manages and uses information in work processes. They discussed the advantage that is derived from the information-processing component that execute steps required to capture, manipulate, and channel the data necessary to perform the value chain activity. The data handling improvements they identify (that could lead to competitive advantages from improved data handling) were extensive. They attribute potential improvements to: bar codes for error handling reductions, databases for knowledge and experience storage, management of services with data, improved weather satellite data uses, financial analysis through data, transfer of data between suppliers and manufacturers, data for improved designs for manufacturing coordination, uses of office support data, and communication data. It is important to note that their examples relate to

the integration effects of these many data sharing and transmission opportunities.

Many other researchers have used the value chain framework to address the role of information systems in achieving helpful integration. For example, White and Person (2001) utilized Porter's value system concepts as the framework for integrating a firm's activities within the supply chain. They emphasized the importance of integrating customer service activities into the decision making process of manufacturing organizations. They further argue that just-in-time systems and new technologies (product, process and information) provide the mechanisms for integration of the various activities across the supply chain. However, they do not specifically show how this "integration" can be achieved using information technology (with low costs) across the value chain.

And there's the rub. Integration is costly and difficult to maintain. This paper will provide an overview of the methods used to achieve integration. But first, it will set the stage for the RPA technology by providing an overview of what it is, and why it is (apparently) being introduced into organizational processes as an integration tool.

Robotic Process Automation (RPA) and Processes

There are two underlying concepts important in Robotic Process Automation (RPA). One must be able to identify and understanding work "processes", and be able model the work processes. Conceptually, simple processes are groups of actions, steps, operations, and decisions or other related activities taken to achieve some goal or purpose. Process thinking is "viewing actions as groups of activities" with a purpose, not just seeing each step in isolation. This process concept is important for effective management, as Porter and Millar noted. It has been an important part of organizational quality management, performance improvement, and productivity enhancement steps for almost three decades. Process thinking requires systematic assessments of work actions and steps, concentrating on work that meets a customer's needs, targeting an objective or goal, focusing on value-adding steps and activities, utilizing user feedback in developing (improving) the process and always keeping in mind the end result – increased productivity.

RPA is software, in its simplest form - a bot. It is a specialize type of software (or code) that maps data from one layers of code to another layer. It may be viewed from a higher-level systems perspective as a connection subsystems from a layer architecture. The translation of data from one layer of code in a computer to another layer enables everyone to understand what the data are. RPA is a software "presentation" layer that is programed to find, access (read), and then re-enter correct data into a different specific location in a file or record. It also provides a way to view or display the data. Robotic process automation operates within this "presentation" layer of software. RPA is not an invasive technology that requires changes in a system or application. What this means is that applying this technology does not require that one change the existing process steps, calculations, comparisons, or actions. Thus, an organization can maintain its currently operating applications without massive modifications while improving performance.

Finally, RPA may also add capability to a current process. Thus, a new information system or a complex technical solution is not necessary when the RPA layer is used to access and perhaps enhance or modify the use of the data so it meets a new requirement (new report, further validation, incorporation into a different process, or a new test or comparison).

The bottom line is that the functionality in the presentation-layer automation software (RPA) is specific and matches the rules and steps in the operational work processes and flows. This rules-based action approach is not subjective. It automates (though a bot) a wide variety of back-office tasks from data entry, comparisons, and validation to automated ordering and payments.

2. PREVIOUS APPROACHES TO THE INTEGRATION PROBLEM

A number of major approaches to integration have been offered for organizations. The approaches have great value, but also bring limitations and costs before they can be used.

Same Data Solution

Information systems have called for organizations to normalize and simply "use the same data" for many years. This paper will not attempt to address the many reasons that this has proven to be an elusive objective. Information systems meet different functionalities, large organization develop specific terms and uses for the same data, and the enormous coordination and communication

efforts required to achieve this elusive target are simply not "free."

Application Level Solutions

It has been recommended that organizations attempt to design and build information systems and applications in parallel. This is a form of managing the coexistence and coordination of multiple concurrent activities. The communication among the system components and their synchronization are common problems that occur when this is attempted. Coordination language is used to synchronize the activity of those computations through component cooperation. (Hasselbring, 2000).

Example systems using this integration approach include inter-organizational systems designed to enhance supply-chain visibility. The systems improve coordination between buyers and suppliers through electronic integration. (Grover, & Saeed, 2007). This tightens linkages in the supply chain, but data do not indicate how the conditions under which transaction exchanges are conducted impact the use of integrated systems. These transactional characteristics are important antecedents to integration under conditions where demand uncertainty, complexity, market fragmentation, and market volatility capture key characteristics that make integration important and very valuable. Data collected from the electronics industry show that firms tend to deploy integrated systems when complexity of the component is high, market fragmentation is low, and an open information-sharing environment exists. Thus, from a managerial perspective, integration is the appropriate configuration under conditions of high product complexity and open information-sharing environment, but it precludes the firm from participating in the open market and gaining brokerage benefits. (Grover, & Saeed, 2007).

Enterprise Solutions

At the enterprise level, a diversity of information systems is often employed for integration such as custom applications, e-business solutions and Enterprise Resource Planning (ERP). The solutions all support the organizational and financial business processes, but the diverse and incompatible systems restrict the automation of business processes and create a proliferation of integration difficulties. Organizations have used integration software called Application Integration (AI) to deliver flexible and more manageable Information Systems (IS) and infrastructures. The Application integration is achieved by linking functionality from disparate systems with adapters and message brokers. The case study of a multinational petroleum company that adopted this solution required up to 60% of overall project time to integrate the systems due to the necessary re-engineering of business processes by phasing out systems and reducing redundancy in functionality. (Themistocleous & Irani, 2002). Further work by Irani, Themistocleous, and Love. (2003) and Themistocleous, Irani, and Kuljis (2004) concludes that the capabilities of Enterprise Application Integration (EAI) technologies can support a direct shift away from disparate systems operating in parallel toward a more common shared architecture. They viewed this opportunity as a possible emerging paradigm shift since integration of IS is in-line with the needs of a business altering its IS life cycle. This makes evaluating the full impact of the system difficult, as it has no definitive start and/or end. This case study of IS applications within an e-Government framework can be viewed as a portfolio of technologies to improves infrastructure integration.

The organization level approaches which developed by 2000 sought to identify and integrate independent functions and productive resources across an organization through resource planning and optimization. Approaches involved sending emails and message with data at the lowest levels, and the establishment of common datasets and sources. Examples include SAP which implements organizational integration with a single database, and utilized messaging services for integrating autonomous ERP systems (Hasselbring (2000). Systems using this approach included TSI Software's Mercator which offered specialized functionality in pre-built application adapters to move data for data conversions and messaging services between the SAP R/3 and PeopleSoft ERP systems.

This solution often required reengineering the organization functional business processes to align with the ERP system. However, organizational componentization continued to support the business processes implemented through existing legacy systems.

Architectural Solutions

In order to solve the poor information sharing capability and business adaptability, by integrating logistics information system based on Service-Oriented Architecture (SOA), a fast and flexible integration method for enterprise information system was presented. The analysis and experiments show that it effectively reduced

the cost of system adjustments, shortened the adjustment time, and improved the efficiency of execution and the quality of adjustment, so that the market competitiveness of enterprise improved. (Wang & Wang, 2010)

General Solutions

Ball, Ma, Raschid, & Zhao (2002) discuss the need for supply chain integration (SCI) methodologies as being driven by increases in the globalization of production and sales. They offer an integration architecture, describe the software components of a prototype implementation, and discuss a variety of information sharing methodologies. Their framework of a multi-echelon supply chain process model spans multiple organizations, and promotes intra-organizational knowledge sharing.

The Ball, et. al. (2002) integration is required because the supply chain prototype consists of six main components including ERP, SCM, a simulation, middleware, collaboration software and visualization and decision tools. The ERP component contains multiple ERP instantiations for individual supply chain members. The SCM component integrates with the ERP instantiations to support planning and execution across the total supply chain. They note that the integration of the SCM component and the ERP components forms the integrated Supply Chain Infrastructure (SCI) architecture. The middleware component uses an integration manager, a message broker, data adaptors and a variety of APIs for communication.

3. THE FUNCTIONAL BREADTH OF THE INTEGRATION PROBLEM

A variety of modeling tools have been used to examine the problem and number of functional area examples are provided. The modelling approaches are illustrated by three widely used modelling methods: IDEF0 which is used to establish functional models, IDEF3 which captures process descriptions, and DFDs that describe data-flows among the functional activities. (Shen, Wall, Zaremba, Chen, & Browne, 2004). These tools illustrate the approaches different modelling methods follow at varying levels of granularity, and they types of information required to complete the models. Shen, et.al, 2004) propose that a set of business process models be combined to capture the advantages of each modelling method and maximize the effect of the distinct modelling efforts. They illustrate the effectiveness of this modelling framework in designing an order using a combination of the target enterprise's legacy systems and a catalogue the tools to facilitate the exchange of information (e.g. order request, estimated ship dates, credit checking, etc.) between the customer and the target enterpriser using fax or email or through another communications tool.

Control, Accounting and Reporting

There is a complex relationship between information system integration approaches, such as Enterprise Resource Planning, and management control. Chapman and Kihn (2009) analyze information system integration data architecture. They posit that the single database concept and the variety of ways in which information might be utilized in practice means that a centralized database will link to (positive) business unit performance. They contend that system integration fosters the four design characteristics that provide an environment where management control will be effective and positively related to perceived system success and business unit performance (based upon PLS analysis of survey data). Their conceptualization is that flexibility, innovation, discovery and testing of assumptions foster an ability to drill down into detailed data below the summarized data provided at higher levels.

Healthcare

Nyella and Mndeme (2010) describe the goal and process of restructuring the Health Information Systems (HIS) in dev eloping countries by standardizing and integrating various vertical reporting systems. The pressure resulting from the vertical nature and support for the systems rendered the integration goal challenging and unachievable.

Zapletal, Rodon, Grabar, & Degoulet (2010) examined how clinical data warehouses (CDW - subject oriented, integrated, time-variant, non-volatile collections of data used in support of management decisions) integration with clinical information systems (CIS – containing data for biomedical research) to provide functionality that is not easy to implement with traditional operational database systems. They examined the technology, data, restitution, and administrative functions of this rare integration effort. UML use cases and the mapping rules from the shared integrated electronic health records were matched. Clinicians and investigators were able to conduct clinical research, quality evaluations and outcome studies because of the integration. These indirect benefits support the continuous use of an integrated system. Value is readily

demonstrated with new clinical data from tissue bank systems and biomedical research data that are integrated with legacy data sources. (Zapletal, Rodon, Grabar, & Degoulet, 2010)

Building Information Modeling
Building Information Modeling (BIM) seeks to utilize a digital representation of a facility's physical and functional characters with information in three dimensions (3D) to meet the needs of Architecture, Engineering, Construction and Facilities Management (AEC/FM) functions required to construct a highly usable facility.

Building Information Modeling (BIM) and Geographic Information System (GIS) is a promising and highly challenging topic to transform information towards the generation of knowledge and intelligence that can be utilized in the civil, building and infrastructure sectors. However, the original different purposes of the technologies have introduced significant challenges for the integration. The development and dissimilarities of various GIS and BIM applications show that integration approaches are developed for targeted reasons and focus on solving different specific problems. The parameters influencing the choice of approaches are "EEEF" criteria: effectiveness, extensibility, effort, and flexibility. Semantic web technologies provide a promising and generalized integration solution that comes with large efforts required at an early stage and the isolated development of ontologies within one particular domain. Openness is suggested as a key of the success of BIM and GIS integration. (Liu, X., Wang, Wright, Cheng, Li, & Liu, R., 2017)

Vehicles
Sch6ner, and Dose (1992) in early research on autonomous vehicles examined an approach to task-level system integration used to plan and control autonomous vehicle motion. They demonstrate the system capabilities and its ability to integrate redundant as well as complementary information with software simulations. They concluded that autonomous systems fulfilling tasks such as moving towards a goal and avoiding sensed obstacles face many problems because they must coordinate sensory and effector modules. They note that system integration is possible in the sense that all information provided by the various sensory modules and all information required by the various effector modules becomes part of the planning dynamics. Dynamic processes must then separate convergent information, and integrate actions by selecting a representative, from non-redundant information, which is kept invariant.

4. ROBOTIC PROCESS AUTOMATION

RPA is semi-automation and automation, effectively applied to rule-based, routine, and predictable tasks in combination with structured, understood, and stable data. (Primer, 2015).

Productivity and Performance
Employee productivity improvements are a byproduct of processes that are assigned to software robots. The software robot can process (without errors) the repetitive and more tedious jobs. It does not lose its place, slow down, skip an action, or forget. Fewer errors need to be identified, diagnosed, and assigned for to "rework." Fewer correcting adjustments are made to the outputs of the competed and standardized work. This provides time for the employees to concentrate on the exceptions or more difficult projects requiring a higher level of skill and training. These are the higher value-added activities that require more in-depth analysis, recognition of the exception, and personal interaction or problem solving decisions. This work adds value to a department or office and may be worth more, thereby increasing productivity. [It is also possible that this work will improve morale, and enhance employee retention.] (Asatiani, & Penttinen, 2016; Fung, 2014; Geyer-Klingeberg, Nakladal, Baldauf, & Veit, 2018, September).

The expanded capability and time available to the employee may also support the organization because these factors contributes to the overall customer experience by solving the customer's problem sooner. Employees will have more time to devote to the customer-facing roles. They are then able to analyze situations requiring personal involvement and direct communication thereby improving customer satisfaction. (Jovanović, Đurić, & Šibalija, 2018; Lacity, & Willcocks, 2015)

It must be emphasized that RPA reduces or possibly eliminates errors, especially those that are made by humans who process data, transpose numbers, or perform actions out of a required sequence. The most typical errors not made by bots are those common mistakes of transcription (error in copying) and transposition when information is input in the wrong order because people simply make errors when they type numbers rather than words. This accuracy improvement is real and measurable. RPA improves accuracy. Employees might make

mistakes, but software robots will not make those mistakes. RPA will also minimize or eliminate complications with employee errors attributed to training omissions, someone filling in (when the regular employee is out), cultural and language barriers, and errors attributed to processing that must occasionally occur in different locations or across time zones.
(Peláez & Kyriakou, 2008; Primer, (2015).

Application of RPA
The scope of the RPA technology is very impressive, and robust. It can be applied across many functions and seems practical for many different process focused tasks (definable, repeatable and rules-based). It can be executed at the direction of employees and assist them in their work by helping diagnose when decisions are not always clear (the data don't legitimately fit) and the rules base is not complete for all situations. RPA has multiple operating modes. It can operate in attended mode where an employee "triggers" the bot for day-to-day operations. The bot can function in an "unattended mode" on a server based on user-determined triggers such as a date and time like 12:00 AM on Friday. Thus, the RPA bot can serve as an independent automated process that does not demand human intervention in order to execute a work process and make or execute a decision if all the rules are clear and the decisions are pre-determined. RPA is very adaptive and fits many situations because of its internal capabilities.

RPA has several essential features that provide it competencies beyond those found in scripting, screen scraping, and sequential process management. 1) RPA utilizes dropping and dragging via icons that represent steps in a process. It is straightforward. Process code is then produced automatically without extensive programming, computer training or expertise. 2) The RPA bot accesses data produced by other computer systems or programs. It emulates exactly how an employee accesses this information (because the bot is created to do just this task). RPA can assume that logon ID and password are required to access what is normally seen or obtained by the worker from the other system's presentation layer. Therefore, the RPA bot is never interfering or invasive. 3) Finally, RPA is a secure and scalable technology that executes on the enterprise-protected platform. It can be configured, audited, and managed at the enterprise or organizational level that utilizes this technology.

The output of this bot appears to work "like a macro," but with more capabilities and functionality that is not restricted to an application like Word or Excel. Think of a very smart, tireless, and sophisticated desktop assistant. The bot is a powerful "aid" that performs scripting and screen scraping (record and replay), acts quickly, and is able to record (without error) what it is doing. Then the bot aid replicates the assigned task repeatedly – like a true robot. It is trained by watching worker selections, recording mouse clicks, matching inputs from the key board and completing the process as the user does. However, the bot is not intelligent – and does not know why it is doing this work since it only performs the assigned set of actions when called upon. (Madakam, Holmukhe, & Jaiswal, 2019; Peláez, & Kyriakou, 2008; Schmitz, Dietze, & Czarnecki, 2019).

5. CONCLUSION

The limited research and descriptive case studies suggest that RPA may have substantial potential for information systems integration in routine and standardized tasks that involve legacy systems.

RPA focuses primarily on the transactional processes that occur with more routine processes and back office work. However, the future for RPA might well include bots that learn and implement analytical processes, and complex work steps requiring more reasoning. They could act as a human might respond to data or situations that are more involved. RPA may evolve into more sophisticated processes that can modify the response required, and evaluate the data in light of the context or concern as well as interact and iterate steps and responses.

To accomplish this in future applications, RPA will need to comprehend and understand contextual situation. Artificial intelligence (AI) and machine learning algorithms, and cognitive computing systems can respond in this manner. However, these tools are trained to recognize and respond appropriately, and not programmed to be "intelligent." The future of RPA may be to work beside and integrate with employees to aid humans and support decision making with more and more detailed analysis. (Schmitz, Dietze, & Czarnecki, 2019; Slaby, 2012; Willcocks, Lacity, & Craig, 2015).

6. RECOMMENDATION: WHEN AND HOW SHOULD RPA BE USED?

IT processes have been targeted as work processes where RPA can deliver significant benefits for about 10 years. The more general characteristics of processes which may benefit most are predicted to include those with: high volumes and value of transactions, frequent access to multiple systems, environment stability, limited need for human intervention and exception handling, manual, prone to errors or re-work, and readily decomposed into steps with , and with a clear understanding of current manual costs. Eight use cases that are targets for this technology include server support, storage, networks, application, security automations, account identity and password management, automated job scheduling, and ITPA integration. The potential benefits include IT service: repeatability, predictability, integration, productivity, satisfaction, risk reduction, cost effectiveness, and improved business performance. (Fung, 2014)

7. REFERENCES

Asatiani, A., & Penttinen, E. (2016). Turning robotic process automation into commercial success–Case OpusCapita. Journal of Information Technology Teaching Cases, 6(2), 67-74.

Ball, M. O., Ma, M., Raschid, L., & Zhao, Z. (2002). Supply chain infrastructures: system integration and information sharing. ACM Sigmod Record, 31(1), 61-66.

Brancheau, J. C., & Wetherbe, J. C. (1987). Key issues in information systems management. *MIS quarterly*, 23-45.

Chapman, C. S., & Kihn, L. A. (2009). Information system integration, enabling control and performance. Accounting, organizations and society, 34(2), 151-169.

Doan, A., & Halevy, A. Y. (2005). Semantic integration research in the database community: A brief survey. AI magazine, 26(1), 83-83.

Fung, H. P. (2014). Criteria, use cases and effects of information technology process automation (ITPA). *Advances in Robotics & Automation*, *3*.

Geyer-Klingeberg, J., Nakladal, J., Baldauf, F., & Veit, F. (2018, September). Process Mining and Robotic Process Automation: A Perfect Match. In *BPM (Dissertation/Demos/Industry)* (pp. 124-131).

Grover, V., & Saeed, K. A. (2007). The impact of product, market, and relationship characteristics on interorganizational system integration in manufacturer-supplier dyads. Journal of Management Information Systems, 23(4), 185-216.

Hasselbring, W. (2000). Information system integration. Communications of the ACM, 43(6), 32-38.

Irani, Z., Themistocleous, M., & Love, P. E. (2003). The impact of enterprise application integration on information system lifecycles. Information & Management, 41(2), 177-187.

Jovanović, S. Z., Đurić, J. S., & Šibalija, T. V. (2018). Robotic process automation: overview and opportunities. *Int. J. Adv. Qual*, *46*.

Lacity, M., & Willcocks, L. (2015). What knowledge workers stand to gain from automation? *Harvard Business Review*, *19*(6).

Lacity, M., Willcocks, L. P., & Craig, A. (2015). Robotic process automation at Telefonica O2.

Ilo, N., (2018). Robotic Process Automation Implementation in Record-to-Report Process.

Lacity, M., Willcocks, L., & Whitley, E. (2016). Service automation: robots and the future of work.

Liu, X., Wang, X., Wright, G., Cheng, J. C., Li, X., & Liu, R. (2017). A state-of-the-art review on the integration of Building Information Modeling (BIM) and Geographic Information System (GIS). ISPRS International Journal of Geo-Information, 6(2), 53.

Madakam, S., Holmukhe, R. M., & Jaiswal, D. K. (2019). The future digital work force: robotic process automation (RPA). *JISTEM-Journal of Information Systems and Technology Management*, *16*.

Nyella, E., & Mndeme, M. (2010). Power tensions in health information system integration in developing countries: the need for distributed control. *The Electronic Journal of Information Systems in Developing Countries*, 43(1), 1-19.

Peláez, A. L., & Kyriakou, D. (2008). Robots, genes and bytes: technology development and social changes towards the year 2020. *Technological Forecasting and Social Change*, *75*(8), 1176-1201.

Porter, M. E., & Millar, V. E. (1985). How information gives you competitive advantage.
https://hbr.org/1985/07/how-information-gives-you-competitive-advantage[28/11/2016 11:34:33]

Primer, A. (2015). Introduction to Robotic Process Automation.

Schmitz, M., Dietze, C., & Czarnecki, C. (2019). Enabling digital transformation through robotic process automation at Deutsche Telekom. In Digitalization Cases (pp. 15-33). Springer, Cham.

Schöner, G., & Dose, M. (1992). A dynamical systems approach to task-level system integration used to plan and control autonomous vehicle motion. Robotics and Autonomous systems, 10(4), 253-267.

Shen, H., Wall, B., Zaremba, M., Chen, Y., & Browne, J. (2004). Integration of business modelling methods for enterprise information system analysis and user requirements gathering. Computers in Industry, 54(3), 307-323.

Schmitz, M., Dietze, C., & Czarnecki, C. (2019). Enabling digital transformation through robotic process automation at Deutsche Telekom. In Digitalization Cases (pp. 15-33). Springer, Cham.

Simon, A., Karapetrovic, S., & Casadesús, M. (2012). Difficulties and benefits of integrated management systems. Industrial Management & Data Systems.

Simon, J. G. M. H. A. (1958). Organizations.

Slaby, J. R. (2012). Robotic automation emerges as a threat to traditional low-cost outsourcing. HfS Research Ltd, 1(1), 3-3.

Themistocleous, M., Irani, Z., Kuljis, J., & Love, P. E. (2004, January). Extending the information system lifecycle through enterprise application integration: a case study experience. In 37th Annual Hawaii International Conference on System Sciences, 2004. Proceedings of the (pp. 8-pp). IEEE.

Themistocleous, M. & Irani, Z. (2002). Evaluating and adopting application integration: the case of a multinational petroleum company. In 35th Annual Hawaii International Conference on System Sciences, Big Island, HI, 2002. Proceedings of the (pp. 3715-3723, doi: 10.1109/HICSS.2002.994502).

Thompson, J. (1967). D. (1967) Organizations in action. New York.

Wang, W. X., & Wang, C. G. (2010). Integration framework of enterprise information system based on service-oriented architecture. Computer Engineering, 36(18), 29-31. (No paper)

White, R. E., & Pearson, J. N. (2001). JIT, system integration and customer service. International Journal of Physical Distribution & Logistics Management.

Willcocks, L. P., Lacity, M., & Craig, A. (2015). Robotic process automation at Xchanging 05.

Zapletal, E., Rodon, N., Grabar, N., & Degoulet, P. (2010). Methodology of integration of a clinical data warehouse with a clinical information system: the HEGP case. In MedInfo (pp. 193-197).

The Sting of Adoption: The Technology Acceptance Model (TAM) with Actual Usage in a Hazardous Environment

Edgar Hassler
hassleree@appstate.edu

Preston MacDonald
preston.macdonald@gmail.com

Joseph Cazier
cazierja@appstate.edu

Center for Analytics Research and Education
Appalachian State University
Boone, NC 28608-2049, U.S.A.

James Wilkes
wilkesjt@appstate.edu
Computer Science Department
Appalachian State University
Boone, NC 28608-2049, U.S.A.

Abstract

The Technology Acceptance Model (TAM) has been used for decades to explain adoption of business technology in a traditional office environment. It was later expanded to consumer side software. However, prior work still has some limitations. 1. These studies generally rely on self-reported intent to use measures rather than actual usage and 2. Prior research seldom looks at adoption in hazardous usage environments. This study extends prior research by looking at actual usage of new software in a hazardous environment, that of a bee yard. Results show that user perceptions of ease of use and usefulness are predictive of a user's intent to use the software and that the user's intent translates into actual usage of the software. Additionally, evidence is presented suggesting the need for an extension of the model to better reflect hazardous physical tasks and environmental conditions.

Keywords: TAM, Technology Acceptance Model, Beekeeping, Apiary Management

1. INTRODUCTION

The Technology Acceptance Model (TAM) has been used for decades to help guide and explain the adoption of information systems by various groups. However, some groups have been slower to adopt information technologies than others.

In this study we look at a profession crucial to both agriculture and the environment that has been particularly slow to adopt these

technologies, the world's beekeepers. Given the importance of pollination to agriculture, being necessary for roughly ¾ of the world's food crops, the pollinators that beekeepers manage are critically important to our food supply.

With annual honeybee losses now approaching 50% in much of the world due to pests, pathogens and environmental factors, there is much cause for concern about the future of these pollinators.
Information technologies can be used to help beekeepers be more successful. Using the lens of TAM, we examine how we can understand and perhaps encourage the adoption of these technologies by beekeepers.

To validate TAM in the beekeeping domain, we partnered with a leading apiary management software provider known as Hive Tracks. Through this partnership, we were able to collect the data required, including an online survey and actual usage data from their software database. This provides the advantage of seeing how responding user's intentions to use software translates to their actual usage.

Hive Tracks has a focus on research and citizen science (Hive Tracks, 2018) providing a unique opportunity to utilize a large volume of quality data. With over 19,000 users, Hive Tracks is growing in popularity with the beekeeping community. Beekeepers use this system for many beekeeping activities, such as managing and monitoring hives, recording inspections, inventory management, calendar scheduling, data recording, and community collaboration.

The primary objective of this study is to validate TAM as a means of examining beekeeping software and, if validated, to utilize TAM to understand what changes are necessary to facilitate wider adoption of such software.

2. BACKGROUND

TAM has an extensive body of research that both supports and criticizes the theory it embodies (Chuttur, 2009; King & He, 2006; Lee, Kozar, & Larsen, 2003; Legris, Ingham, & Collerette, 2003; Li, Qi, & Shu, 2008; Qingxiong & Liping, 2004; Schepers & Wetzels, 2007; Sharp, 2007; Turner, Kitchenham, Brereton, Charters, & Budgen, 2010). In this section, we provide a very brief background of the model and its use in the associated field of agriculture. Further information regarding the TAM constructs can be found in the model development section that follows.

TAM Origins

TAM is a derivation of the Theory of Reasoned Action (Fishbein & Ajzen, 1975), outfitted for the prediction of IT acceptance and use (Davis, 1986). A revised version of TAM, known as parsimonious TAM (Davis & Venkatesh, 1996), hypothesizes that IT use can be predicted by its perceived usefulness (PU) and perceived ease of use (PEOU), mediated by a subject's behavioral intention (BI). All factors in the TAM equation, except actual IT use, are therefore measured as one's perceptions regarding one's beliefs and intentions.

TAM as a Tool for Understanding Adoption

In practice, TAM has proven to be both powerful and parsimonious as a useful tool for understanding technology adoption through perceived characteristics (Cazier, Wilson, & Medlin, 2009). Lee, Kozar, and Larsen (2003), also reported support for the central relationships of TAM. Among the studies which assessed each specific relationship, 88% find PU influences BI, 71% find PEOU influences BI, 84% find PEOU influences PU, and 87% find BI influences IT use.

Lee et al. (2003) describe 25 additional factors that have been studied as contributors to TAM, ranging from measures of voluntariness of use to users' prior experiences with the technology. However, due to concerns regarding survey length and confounding due to a new domain, these additional factors will be reserved for study at a future time. We will be focusing on the core factors consisting of perceived usefulness, perceived ease of use, behavioral intentions, and actual use.

TAM in Agriculture

In a world where sustainability challenges arise, it becomes imperative for the IS community to educate others to build innovative IS solutions for a modern world (Watson, Boudreau, & Chen, 2010). TAM has proven to be a useful tool to understand technology adoption in agriculture. Adrian, Norwood, & Mask, (2005) used TAM to investigate the perception and attitudinal characteristics of farmers who planned to adopt technologies. Rezaei-Moghaddam & Salehi (2010) also explored agricultural worker's intentions toward precision agriculture technologies. By following TAM, they were able to determine that observability, trialability, and attitude to use positively affect intentions for

someone to adopt precision agriculture technologies.

3. RESEARCH MODEL

As shown in Figure 1, the research model follows the parsimonious TAM. The model is formed from the three core constructs found in the majority of TAM studies and a reflective construct of actual usage. Relationships among the constructs follow those commonly found in the literature and include tests for mediation.

Construct Definitions

The following constructs and associated definition are utilized in the model.

Perceived Ease of Use is "the degree to which a person believes that using a particular system would be free of effort" (Davis, 1989).

Perceived Usefulness is "the degree to which a person believes that using a particular system would enhance his or her job performance" (Davis, 1989).

Behavioral Intent is "a measure of the strength of one's intention to perform a specified behavior" (Davis, Bagozzi, & Warshaw, 1989).

Actual Usage is defined in this study as the number of actions executed by a user on the system of interest.

Construct Relationships

Relationships among the constructs are codified in the following hypotheses:

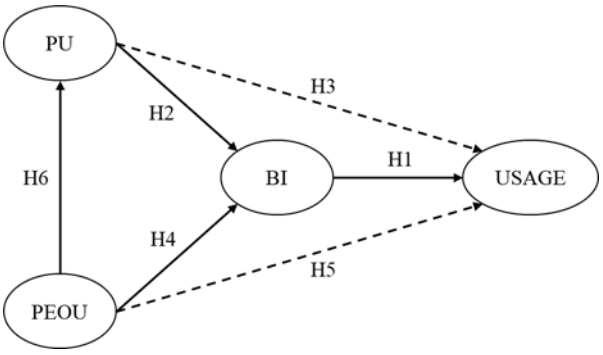

Figure 1. Research Model

H1: Increased intention to use apiary management software increases actual usage of apiary management software.

H2: Increased perceived usefulness of apiary management software increases the intention to use the apiary management software.

H3: An increase in perceived usefulness increases actual usage but is fully mediated by the user's behavioral intentions.

H4: Increased perceived ease of use of apiary management software increases the intention to use the apiary management software.

H5: An increase in perceived ease of use increases actual usage but is fully mediated by the user's behavioral intentions.

H6: Increased perceived ease of use of apiary management software increases the perceived usefulness of the apiary management software.

4. METHOD

The primary goal of the research is to validate that the TAM can be applied to the domain of beekeeping and, if so, to elicit improvements that can be made to the Hive Tracks software based on the model. To accomplish this task a survey is conducted of the Hive Tracks user base in conjunction with data extraction from the Hive Tracks database. Before undertaking the main study, a pilot study was conducted to find and resolve any issues with the survey instrument and extraction process.

Subjects

Subjects for the investigation are beekeepers registered as users of Hive Tracks software as of March 2018. Registered Hive Tracks users number over 19,000 individuals representing over 150 countries around the world (Hive Tracks, 2018). Users vary in both beekeeping experience level and number of hives managed. Additionally, experience level regarding use of the Hive Tracks software varies within the group.

Instrumentation

Two forms of instrumentation are utilized in the measurement model. A questionnaire is used to measure the antecedents of actual system use and collect demographic data. A measure of actual system usage is constructed from activity logs extracted from the Hive Tracks database.

Questionnaire

Items measuring the TAM constructs were chosen from previous research and adapted for the context of this investigation. For the constructs Perceived Ease of Use (PEOU), and

Perceived Usefulness (PU) four items were selected to measure each construct. The construct of Behavioral Intent (BI) was measured with three items. Table 1 shows the adapted items, their origin, and the associated construct.
Each item is measured using a seven-point Likert scale with endpoints labeled "Strongly Disagree" (Value = 1) to "Strongly Agree" (Value = 7). In addition, for each construct, a free-form question is included to allow participants to extrapolate on how to improve ratings on the construct. Additionally, participants are asked to provide basic demographic information including gender, year of birth, education level and number of hives managed.

The questionnaire is deployed utilizing the Qualtrics survey platform.

Actual Usage Measure
For the actual usage construct, four measures of user activity are extracted directly from the Hive Tracks database:

- The number of user logins.
- The number of user actions related to hive activities.
- The number of active hives associated with each user.
- The number of user actions related to non-hive activities such as reports and configuration.

The number of user actions related to hive activities is normalized by dividing it by the number of active hives found in the database to provide the average number of actions per registered hive. This measure in addition to the number of user logins and number of non-hive activities constitute the three operationalized measures of the actual use construct.

Data Collection
An invitation email was crafted containing a link for the online survey and sent to registered users by Hive Tracks management. The survey remained open for 30 days in which a user could elect to voluntarily respond. Following the close of the survey, actual usage data for responding users was extracted from the database for the following 30 days and for 11 months prior to the survey closing. Thus, allowing for examination of actual usage both before and after the survey.

Analysis
Confirmatory Factor Analysis and Structural Equation Modeling are accomplished utilizing SAS v9.4.

Measure	Questions
PEOU #1 (Premkumar & Bhattacherjee, 2008)	**Original:** Learning to use CBT is easy for me. **Adapted:** Learning to use Hive Tracks is easy for me.
PEOU #2 (Premkumar & Bhattacherjee, 2008)	**Original:** My interaction with CBT is clear and understandable. **Adapted:** My interaction with Hive Tracks is clear and understandable.
PEOU #3 (Lai & Li, 2005)	**Original:** It is easy to use Internet Banking to accomplish my banking tasks. **Adapted:** It is easy to use Hive Tracks to accomplish my beekeeping tasks.
PEOU #4 (Lai & Li, 2005)	**Original:** Overall, I believe Internet Banking is easy to use. **Adapted:** Overall, I believe Hive Tracks is easy to use.
PU #1 (Lai & Li, 2005)	**Original:** I can accomplish my banking tasks more quickly using Internet Banking. **Adapted:** I can accomplish my beekeeping tasks more quickly using Hive Tracks.
PU #2 (Lai & Li, 2005)	**Original:** Internet Banking enables me to make better decisions in utilizing banking services. **Adapted:** Hive Tracks enables me to make better decisions in beekeeping.
PU #3 (Lai & Li, 2005)	**Original:** Internet Banking enhances my efficiency in utilizing banking services. **Adapted:** Hive Tracks enhances my efficiency in beekeeping.
PU #4 (Lai & Li, 2005)	**Original:** Overall, I find Internet Banking useful. **Adapted:** Overall, I find Hive Tracks useful.
BI #1 (Lai & Li, 2005)	**Original:** I will use Internet banking on a regular basis in the future. **Adapted:** I will use Hive Tracks on a regular basis in the future.
BI #2 (Lai & Li, 2005)	**Original:** I will frequently use Internet banking in the future. **Adapted:** I will frequently use Hive Tracks in the future.
BI #3 (Lai & Li, 2005)	**Original:** I intend to continue using this software. **Adapted:** Overall, I will continue using Hive Tracks in the future.

Table 1. Original and adapted survey questions

5. RESULTS

Following the invitation to participate, 484 users responded by completing the survey. After preliminary examination of the responses, 49 cases were removed due to incomplete data or failure of bias/consistency check built into the instrument. Thus 435 usable responses are included in the analysis.

Item	Value	Percent
Gender	Female	26.46
	Male	71.43
	Prefer Not to Answer	2.1
Age (year)	< 31	3.04
	31-40	11.94
	41-50	18.50
	51-60	31.85
	61-70	22.95
	>71	4.68
	Missing	7.03
Education	Some Schooling	1.41
	High School Graduate, or the Equivalent	7.96
	Some college, no degree	17.56
	2-year degree	11.01
	Bachelor's Degree	28.10
	Graduate Degree	32.79
	Missing	1.17
Average Hives	Less Than 5	44.26
	5-10	29.51
	11-15	6.56
	16-20	4.22
	21-30	7.26
	31-40	2.34
	Over 40	5.85
Region	USA Midwest	19.44
	USA South	39.34
	USA West	14.99
	USA Northeast	12.65
	Other	13.59

Table 2. Participant Demographics

Participants

Although Hive Tracks has a good mix of users internationally, the majority of users are in the United States. Users are primarily hobbyist beekeepers with some sideline or part-time beekeepers. Respondents also have a variety of different experience levels in terms of beekeeping and using the software. Table 2 provides a demographic overview of the participants.

Measurement Model

To establish the unidimensionality of the scales a Confirmatory Factor Analysis utilizing the SAS 9.4 CALIS procedure was completed. To facilitate the analysis, a log transformation was applied to each of the actual use measurement items. No additional modifications were made to the measurement model.

Results indicated an acceptable measurement model (X2 = 224.75, df = 71, RMSEA = .07, CFI = .98). A Wald Test indicated all parameters are significant and thus none should be dropped. Table 2 shows the means, standard deviations, Cronbach's alpha, and intercorrelations for the outcome and antecedents variables. Additionally, the square root of the average variance extracted is included in bold on the diagonal to illustrate the discriminant validity of the scales.

Variable	Mean	SD	α	EOU	USF	BI	ACT
PEOU	5.25	1.33	.95	**.90**			
PU	5.01	1.39	.95	.79	**.91**		
BI	4.78	1.81	.98	.60	.69	**.97**	
USAGE	2.20	1.88	.96	.13	.11	.18	**.94**

Notes: Correlations are significant at the p<.05 level. Sqrt (Average Variance Extracted) indicated in bold.

Table 3. Scale Summarys and Correlations.

SEM Results

The estimated model indicated an acceptable fit with the data (X2 = 224.75, df = 71, RMSEA = .07, CFI = .98). Figure 2 shows the standardized effects for paths with p-values - dashed lines indicate insignificant paths.

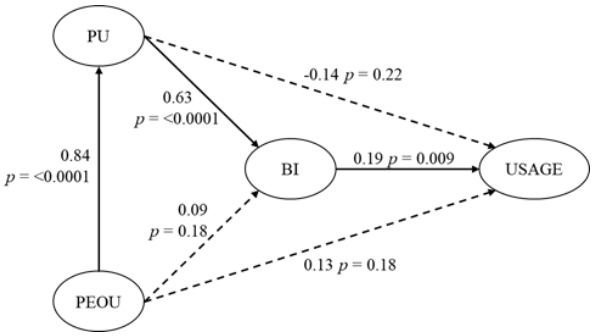

Figure 24. SEM Standardized Results

As indicated in Figure 2, support for hypothesis H1, H2, and H6 is demonstrated. An increase in PEOU increases PU ($\eta = 0.84$, $p < 0.0001$). An increase in PU increases BI ($\eta = 0.63$, $p < 0.0001$). Finally, an increase in BI increases USAGE ($\eta = 0.19$, $p = 0.009$).

No Support is found for the hypothesis H4 (PEOU => BI, p = 0.18). By association, no support is found for H5 as it relies, in part, on H4. A test for support of a direct effect of PEOU on USAGE is also unsupported (p = 0.18). Finally, further analysis demonstrates support for an indirect effect of PEOU on BI via PU (η = 0.53, p < 0.0001). Thus, an increase in PEOU increases BI and is fully mediated by PU.

Support for H3 is also demonstrated. With evidence for H2 and a direct effect of PU on USAGE not found (p = 0.22), further analysis demonstrates support for an indirect effect of PU on USAGE via BI (η = 0.53, p < 0.0001). Thus, an increase in PU increases USAGE and is fully mediated by BI.

From the estimated model it can be seen that perceived ease of use has a large effect on the perception of usefulness. Perceived usefulness, in turn, drives behavioral intent to use the software. The materialization of this intent can then be seen in actual usage as measured directly in the application.

Responses to the essay questions both confirm and supplement the model results. Suggestions to improve ease of use included: "allow me to edit fields like medication (not all of them on your list)"; "connect to QR reader on phone"; and "[add] voice commands." Thus, ease of use is a driving consideration for beekeepers. Likewise, suggestions to make the application more useful included feature requests such as "bulk import of supplies with a ccv would be nice" and "[add] ability to clone a yard."

Additional comments not directly related to the model are also noted. Many of these comments are related to pricing of the application. While germane to the question of how to improve user satisfaction, further research is needed to understand implications of cost on the model.

6. DISCUSSION

The results of this study are consistent with previous studies of TAM (King & He, 2006; Li et al., 2008; Schepers & Wetzels, 2007; Turner et al., 2010) and demonstrates support for the use of a parsimonious TAM in beekeeping. Cleary, how easy an application is to use and the useful features embodied in the application are a determinant of system usage.

However, the study also advances our understanding of what ease of use should mean. As one respondent noted: "My problem using the app is that I don't want to pull out my nice phone when I'm in the hives and my hands are covered with propolis." While another commented: "I need to be able to speak and have the notes automatically placed on the hive I'm talking to it about. I can't push buttons with my gloves on." This qualitative evidence suggests that explicitly accounting for the ease of integrating with physical tasks is needed in the model. Further research is needed to understand if this should be manifested as a facet of the ease of use construct or as an independent construct.

For providers of beekeeping software, this study clearly shows that while useful features are important, easy to use and useful features are even more important. Beekeeping involves several physical tasks, such as inspections, that are difficult to automate or reengineer. The physical nature of these tasks along with the environment in which they are performed must be considered in the design of software intended to support such tasks. With the recent advances in voice recognition and processing, the addition of such capabilities to beekeeping software would be a prudent design enhancement.

Limitations

The findings of this study should be interpreted with a degree of caution as only one beekeeping software application is examined. Additionally, the demographics of the participants may not be representative of the entire beekeeper population. The sample may be skewed as it is dominated by males, individuals of at least 51 years of age, and most of the participants hold at least a bachelor's degree. The actual distributions of these factors in the population is currently unavailable and such factors have proven valid in previous extensions of TAM.

The strong history of research surrounding TAM and its extensions is believed to mitigate such limitations as the core of the model has been supported by numerous studies in the past three decades. Additionally, the strength and significance of the relationships examined provide additional confidence in the findings.

Finally, since an online survey instrument is utilized, one must be vigilant of response bias. To mitigate this concern, establish scales and survey design found in existing TAM literature are employed. A pilot study was then conducted to ensure functionality of the instrument. Additionally, checks of the data for yea/nay-saying, acquiescence and extremity of answers were conducted.

7. CONCLUSIONS

This study demonstrates TAM as a viable lens for examining improvements is systems designed to support beekeepers and suggests an extension of the model for hazardous environments such as an apiary.

Beekeepers have been slow to adopt new technologies, especially information-based technologies, to help them better manage their colonies. Examining related industries, we can see several advantageous opportunities for impacting bee health including data collection, good data management, external data integration, and analysis of data. Maintaining healthy honeybee colonies requires intensive management by the beekeepers, so high-quality data collection will lead to effective understanding and optimization of the economic tradeoffs of Best Management Practices for the beekeepers.

8. ACKNOWLEDGEMENTS

We extend a special thank you to Hive Tracks for encouraging their user base to voluntarily participate in this research.

9. REFERENCES

Adrian, A. M., Norwood, S. H., & Mask, P. L. (2005). Producers' perceptions and attitudes toward precision agriculture technologies. *Computers and Electronics in Agriculture, 48*(3), 256-271. doi:https://doi.org/10.1016/j.compag.2005.04.004

Cazier, J., Wilson, E. V., & Medlin, B. D. (2009). The Impact of Privacy Risk Harm (RH) and Risk Likelihood (RL) on IT Acceptance: An Examination of a Student Information System. In H. Nemati (Ed.), *Techniques and Applications for Advanced Information Privacy and Security: Emerging Organizational, Ethical, and Human Issues* (pp. 211-224). Hershey, PA: IGI Global.

Chuttur, M. Y. (2009). Overview of the technology acceptance model: origins, developments and future directions. *Sprouts: Working Papers on Information Systems, 9*(37), 1–21.

Davis, F. D. (1986). *A technology acceptance model for empirically testing new end-user information systems: theory and results.* (PhD Doctoral dissertation). MIT Sloan School of Management, Cambridge, MA.

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly, 13*(3), 319–340. doi:10.2307/249008

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. *Management Science, 35*(8), 982-1003. doi:10.1287/mnsc.35.8.982

Davis, F. D., & Venkatesh, V. (1996). A critical assessment of potential measurement biases in the technology acceptance model: three experiments. *International Journal of Human-Computer Studies, 45*(1), 19-45. doi:10.1006/ijhc.1996.0040

Fishbein, M., & Ajzen, I. (1975). *Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research.* Reading, MA: Addison-Wesley.

Hive Tracks. (2018). Hive Tracks - About Us. Retrieved from https://hivetracks.com/about_us.php

King, W. R., & He, J. (2006). A meta-analysis of the technology acceptance model. *Information & Management, 43*(6), 740-755. doi:10.1016/j.im.2006.05.003

Lee, Y., Kozar, K. A., & Larsen, K. R. T. (2003). The Technology Acceptance Model: Past, Present, and Future. *Communications of the Association for Information Systems, 12*(50), 752–780. doi:10.17705/1cais.01250

Legris, P., Ingham, J., & Collerette, P. (2003). Why do people use information technology? A critical review of the technology acceptance model. *Information & Management, 40*(3), 191-204. doi:10.1016/s0378-7206(01)00143-4

Li, Y., Qi, J., & Shu, H. (2008). Review of Relationships Among Variables in TAM. *Tsinghua Science & Technology, 13*(3), 273-278. doi:https://doi.org/10.1016/S1007-0214(08)70044-0

Qingxiong, M., & Liping, L. (2004). The Technology Acceptance Model: A Meta-Analysis of Empirical Findings. *Journal of*

Organizational and End User Computing (JOEUC), 16(1), 59-72. doi:10.4018/joeuc.2004010104

Rezaei-Moghaddam, K., & Salehi, S. (2010). Agricultural specialists intention toward precision agriculture technologies: integrating innovation characteristics to technology acceptance model. *African Journal of Agricultural Research, 5*(11), 1191-1199.

Schepers, J., & Wetzels, M. (2007). A meta-analysis of the technology acceptance model: Investigating subjective norm and moderation effects. *Information & Management, 44*(1), 90-103. doi:https://doi.org/10.1016/j.im.2006.10.007

Sharp, J. H. (2007). Development, extension, and application: a review of the technology acceptance model. *Information Systems Education Journal, 5*(9), 1–11.

Turner, M., Kitchenham, B., Brereton, P., Charters, S., & Budgen, D. (2010). Does the technology acceptance model predict actual use? A systematic literature review. *Information and Software Technology, 52*(5), 463-479. doi:10.1016/j.infsof.2009.11.005

Watson, R. T., Boudreau, M.-C., & Chen, A. J. (2010). Information Systems and Environmentally Sustainable Development: Energy Informatics and New Directions for the IS Community. *MIS Quarterly, 34*(1), 23-38. doi:10.2307/20721413

An International Review and Study on Perceptions of Security, Adoption, and Implementation of Electronic Health Records

Faith Aliyu
Fa019@georgiasouthern.edu

Hayden Wimmer
hwimmer@georgiasouthern.edu

Georgia Southern University
Statesboro, GA 30460, USA

Loreen M. Powell
lpowell@bloomu.edu
Bloomsburg University of Pennsylvania
Bloomsburg, PA 17815, USA

Carl Rebman Jr.
carlr@sandiego.edu
University of San Diego
San Diego, CA 92110, USA

ABSTRACT

The transformation of the healthcare system has generated volumes of electronic data available to patients, healthcare workers, and researchers. Personal health information is paramount to the success of implementation of electronic health records (EHRs). Adoption and implementation of EHRs has been strong in the United States of America (USA) and yet varies among countries in the world. This study seeks to determine the factors that individuals, particularly those outside of the USA deem important when considering providing information for EHRs. Survey results indicated that intent of international citizens to provide personal health information depends on more on trust, risk, privacy, and perceived benefits. The outcome from this study can be helpful for other countries and organization seeking to create, establish, or augment an EHR system.

Keywords: International, electronic health records, security, healthcare information technology

1 INTRODUCTION

The aim of healthcare organizations is to provide patients with utmost patient care. EHRs have been proven to be interactive as it helps with proper documentation of medical history with up to date information. An EHR is the systematized collection of patient and population electronically-stored health information in a digital format. (Gunter 2005). More specifically EHRs can be defined as digitally stored healthcare information throughout an individual's lifetime with the purpose of supporting continuity of care, education, and

research (Ajami, 2013b). EHRs consists of multiple types and ranges of data such as lab reports, x-rays, photographs, chart, drugs administered, measurement of patient progress, and even audio of note dictations. Over the last 27 years huge advances in information technology (IT) and particularly in the areas of health, various forms of electronic records have been discussed, designed or implemented (Ajami 2013a, Cuk, Wimmer, Powell, Rebman, 2018).

The use of EHR's are now widespread, transforming the healthcare sector, delivering top notch service to patients with emphasis on patients' health and well-being. (Cuk et al, 2018; Wheatley 2013) The EHRs have 8 major functions which are; health information and data, result management, order management, decision support, electronic communication and connectivity, patient support, administrative processes and reporting population health (Woo 2013).

EHR technology made patients information easily accessible as records became portable and comprehensive. In addition, proponents proclaimed how EHR and Health IT would offer the following benefits; increased quality of healthcare, reduction of medication errors, improvement of patient health outcomes, reduction in health disparities, cost savings, improved patient safety, and augmented chronic disease management (Bowens 2010).

Patient care and satisfaction are what many health organizations offer their clients and EHR are one tool providers can use to achieve that goal (Cuk, Wimmer, Powell, 2017). Old EHR systems used manual paper-based data entry and were used to gather information for research and administrative purposes. The EHR records were not easily accessed and readily available to a large number of users. This caused a lot of delays in processing, updating, and utilization of records. In terms of security, these EHR systems were not that secure preventing unauthorized access and limiting scope of access (Cuk et al., 2018).

The implementation of EHR technology comes with challenges that include patient privacy and security, errors in data capture, errors in data interpretation, and legal and technology compatibility costs with current design of EMR systems (Palabindala et al., 2016, Sittig and Singh, 2011). Patients in general are wary of how their data is used and shared for research and development purposes (Bresnick 2018).

This study examines privacy, confidentiality and security with EHR systems and investigates patient's perceived security of online medical records, particularly of international patients. The format of this study is as follows. First is a discussion of a relevant literature followed by methodology discussion and test results. The manuscript concludes with results, limitations, and future research.

2. LITERATURE REVIEW

History and Evolution of Electronic Health Records

The concept of the EHR needs to be fully absorbed by physicians and the general public. The term EHR has been used in recent healthcare literature without a proper definition of its structure, usage and effect on the healthcare industries. Häyrinen et al. (2008) Investigated the structure of the EHRS as a whole. The objective was to find out how these records are being used, in what context and who has access to these records. A literature search was conducted on healthcare databases to discover the content of EHRS. The results showed that the EHRS consists of various types of data systems in various forms and they were used across all forms of healthcare from primary to tertiary. The information in the EHRS are recorded by different types of healthcare professionals and some by the patients but usually authorized by physicians.

Fragidis and Chatzoglou (2018) examined best practices for implementing EHR systems across 13 countries. Countries participating were largely European complimented by the US, New Zealand, and South Korea. Authors state it is important to consider each country's health system as well as their system of reimbursement and payment. The administrative and bureaucratic structure is also an important consideration. The primary contribution is valuable input from experts in the aforementioned countries who explain challenges and barriers to EHR adoption (Fragidis & Chatzoglou, 2018).

Types of Electronic Health records

Patient's perceived access to their personal health information is an issue of utmost concern, although effective communication between healthcare providers and patients lead to high – quality healthcare service, in the past this has been done in person or over the phone. But with the introduction of IT, patients can now communicate with healthcare providers electronically and hold meetings or schedule

appointments online (Baratam, Abdul, & Powell, 2020). Both patients and healthcare providers make use of a computer on a daily basis, either for personal or business reasons.

Hassol et al. (2004) Described the use of a linked web messaging which is linked to a patient's EHR to schedule appointments with healthcare providers, making a seamless communication between healthcare providers and patients. This study conducted an online survey of 4,282 members of the Geisinger health system who are registered users of the My Chart app, an app which makes patients communicate with their healthcare providers. A survey of focus groups was also conducted with 25 patients who are active users of the system. Age groups of users ranged from 18- 65 and older. Results were analyzed based on user satisfaction, ease of use, communication preferences and accuracy of patient EHR. On a scale of 1-100, most users indicated the system was easy to use with mean scores ranging from 78-85, users indicating how meaningful their medical records were ranged with mean scores between 65 -85. A small number of users were worried about the confidentiality of their health records or abnormality with their records. In conclusion, according to this study, patient's attitude towards the use of web messages and online EHR were positive.

There is on-going research on how patients view EHR and how this affects healthcare organizations. Many individuals have been affected by health information breaches over the years, and cases of data breaches in the healthcare industry and has a cost of about $5.6 billion per year (Millman 2014). According to the UIC health blog, 1 in 3 Americans have experienced some form of data breach in their healthcare records last year (Landi, 2020). Healthcare records can be assessed through the desktop and mobile devices making these records more vulnerable to attacks. These attacks came from various sources, hacking, theft, loss, improper/un-authorized access and un-professional data disposal. As most healthcare providers now adapt the EHR technology, the consumers (patients) are left with no choice than to familiarize themselves with the technology. But the worries that come with the security of patients' EHR make the acceptance of this technology slow. Most consumers have different fears about the security of their EHR, for example the fear of identity theft, personal information leaked online; especially for high dignitaries, the risk of employers knowing about their sensitive health issues amongst others may arise.

Reasons for disclosing Health records.
The need to disclose health records is important Bansal and Gefen (2010) Discussed the effect of patients disclosing their personal health information online with the sensitivity, privacy and trust concerns patients have towards their records being available online. All of which could be traced to personal characters and traits, information sensitivity, health status, experience and risk beliefs that fill in for trust. The unwillingness to provide health information by patients can hinder the implementation of online healthcare services. Most patients are concerned on how their health information is being used and accessed online. The privacy of their sensitive health information is of major fear as the internet can easily be accessed by anyone from anywhere. But on the other hand, patients must disclose their personal health information in order to receive proper care, the issue of this privacy might make some patients refuse healthcare in extreme cases.

Patients concern on Security of EHR's
The loss of information has been a problem due to changes in technology used in health care. A lot of analysis has been done to make health information accessible to various healthcare providers without conflicting patient's perception on confidentiality and autonomy. Cases of patient's information being stolen, lost, misplaced or released without authorization were reported in the UK, with 186 data breaches being reported at the department of health between July 2011 and June 2012 (Caldicott, 2013).

Papoutsi et al. (2015) examined views from patients and the public about information sharing and the concerns it raises about the security and privacy of EHR's used for providing healthcare. A cross sectional survey was conducted, with focus group discussions, the survey participants were gotten from primary and secondary care settings, a total of 5,331 participants were recruited but 2761 participants were used for final analysis in this research. Survey results showed that 79% of participants are worried or have concerns over the security of their health records if it was a national EHR system, 71% were of the opinion that the National Health System (NHS) cannot provide EHR safety at the time of the survey. The population sample that worried about the security of data supports the development of EHR, but 12% didn't support and 33% were

wavering. The issue of integrated EHR's raises worries on the security risks linked with the system, hence the need for a proper recognition of the EHR by the public and the creation of dependable security technique for sharing health data.

Agaku, Adisa, Ayo-Yusuf, and Connolly (2013) evaluated the perception of adults in the US towards the security of their health information. The need to protect patient's data is imminent as most patients express fear over the loss or mis-management of their health records, with the rise in data breaches being reported annually. This study examined the fourth wave in the first cycle of the health information national trends survey, this was done to determine respondents concern about personal health information breaches.

With the inception of new technological advancements, like cloud-based services, and file sharing apps, health information becomes more vulnerable and exposed to risks, due to the rate and volume at which information can be shared. At times, patients may authorize the disclosure of their health information unknowingly which makes them feel violated when they hear about it. Health laws like The Health Insurance Portability and Accountability Act of 1996 (*HIPAA*) have laws where a patient's health information should not be disclosed without proper authorization from patient. However, patient's notion about health records security has not been deeply explored, and its effect on relationship between their healthcare providers. From the survey, it was discovered that people had concerns about data breaches when there is a transfer of health records between healthcare providers, by fax 67%-69% , electronically 64.5- 67% and 12% -13 % did not disclose their Personal health information due to security , because they did not have an idea on how their records were being used. A multivariable logistic regression was used to evaluate the effect of security and privacy concerns on divulging personal health information to healthcare providers.

Fernández-Alemán, Señor, Lozoya, and Toval (2013) documented the findings from a systematic literature review on the security and privacy of EHR. The use of paper-based health records caused a lengthy paper trail, hence the need to move to EHR is inevitable. The benefits of an EHR are so numerous, especially when they are integrated, there is a huge reduction in costs, improved quality of care and an efficient record keeping. All of these benefits are based on the EHR's ability to meet some standard requirements, an effective EHR should be resilient to failure and be consistent with data integrity. The implementation of the EHR system has been hindered by patient's attitude, funding, organizational aspects and technology. A systematic literature review was carried out and data was extracted from 775 articles using a predefined search string, the data sources were from articles found in the ACM digital library, IEEE, science direct amongst others. The results showed that out of 49 articles selected, 26 used standards relating to privacy and security of EHR data. The HIPAA and the European data protection directive were the most widely used regulations, some articles discussed symmetric and asymmetric key schemes, 13 used a pseudo anonymity technique while 11 articles introduced the use of digital signature scheme rooted in public key infrastructure and 13 introduced a login/password with digital certificate or PIN for authentication. Some access control appeared to be role based as seen in 27 studies, 10 explained who should define HER system roles, and 11 discussed who provides access to EHR and some suggested access policies should be overridden in emergency situations.

3. RESEARCH QUESTIONS

Implementation and adoption of EHRs quite well in the United States and varies in other countries. This study sought to examine and determine what perceptions might have influence against adoption of EHR by international citizens. The factors of privacy, security, trust, and perceived benefits were examined.

Privacy was defined as the right individuals have to withhold information about themselves from being leaked to others. Clinical information is considered private and should be protected, it could be in form of treatment, test results, diagnoses that can be stored on various media where patient's identity cannot be confirmed. This data should be released only with the patient's permission or law, physicians can however gain access to this information for treatment and other administrative purposes. To preserve confidentiality, only authorized individuals should access this information.

Information security on the other hand is the safeguarding of data Confidentiality, integrity and availability. The HIPAA and HITECH also enforce the protection of health data, with serious consequences for violations. The need to secure EHR's are due to the increased use of

various technological devices most of which are mobile. Data exchange between different health organizations also poses a threat to EHR's It was discovered that healthcare providers often text other providers about work, the security of this messages are of huge concern; as the level of detail in this information exchange could be ambushed. An encryption of devices used to exchange health information is valuable, also awareness programs should take place to educate users of EHR's on the threats in the system. The use of audit trails to monitor those who have access to patient information.

Trust is the understanding that the data shared will be used for the intended purpose and by only those authorized individuals. Perceived benefits are the gain in individual health and quality of life by the sharing of an EHR.

4. METHODOLOGY

A structured survey was utilized, a questionnaire was used as a measure to gather data. The questionnaire had 12 sections. The first section was designed to collect demographic data from the participants, data relating to educational level, age, gender, and race and job field. The other sections had relating to the factors affecting an individual's perception on the security of EHRs and their intent to disclose sensitive health information to healthcare providers.

The survey had 42 questions and respondents could access via their phones or desktops. The survey is available upon request. The survey tool was acquired from existing constructs and was revised for the purpose of this paper. Six questions were used to measure Risk and the constructs were acquired from Bansal and Gefen (2010) and Malhotra et al. (2004). The questions measuring trust (one question) was adapted from Bansal and Gefen (2010) and (five questions) from Malhotra et al. (2004) to reveal the extent of trust patients have in EHRs. Also, the one (1) item measuring privacy was adapted from Bansal and Gefen (2010) , one (1) item used to measure intent to disclose was adapted from Bansal and Gefen (2010) to develop the construct Intention to disclose. To know how much individuals are concerned about the privacy of their information one item was adapted from Bansal & Gefen, 2010, to develop the construct Privacy. Most users are also concerned and want to know how beneficial the EHR's are to them, therefore 3 items were adapted from Ng, Kankanhalli, and Xu (2009) to develop the construct perceived benefits.

The population of this study was built for international students at Georgia Southern University. The survey was delivered via an online survey tool called Qualtrics. The study is an IRB approved study and a factor analysis was used to select significant variables that was used in the analysis.

On an average it took respondents 10 minutes to completely answer all questions in the survey. A total of 44 validated copies were collected, in the distribution 43% were male and 54% were female, and the age range of the participants were from 18-50 years old. Most of the respondents were graduate students and some were faculty members at the university. Most respondents held a bachelor's degree or above.

5. RESULTS

The dataset had 6 survey constructs and each construct had at least one question measuring patients perceived security of EHRs. These constructs are Demographics, Risk, Perceived Benefits, Trust, and Privacy. Intention to Disclose, which is a derivative of perception of EHRs (DEPENDENT VARIABLE) that depends on trust, risk, privacy, perceived benefits.

Linear regression was conducted with SPSS, with all the variables included. The intention to disclose, which the dependent variable, is had three questions which were all analyzed using the factor analysis, all variables that loaded more than 0.5 were considered significant and were used in the linear regression analysis. During the factor analysis, ITD2, had the highest score of 0.7 compared to ITD1 and ITD3. Factors that loaded less than 0.5 were removed from the analysis. For the independent variables used in this analysis, the following variables loaded more than 0.5 and were used in the analysis:
PB3, T1,T2,T3,T4,T5,PL1,PL2,R3,R4,R5.

The Output of a linear Regression Analysis is in SPSS produces some tables, but two tables are of major concern; The Model Summary and the Anova. During the first set of analysis all constructs were included with "Intention to disclose" (ITD2) as the dependent variable and the other constructs mentioned above as the independent variables. The analysis generated results that are explained as follows:

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.887[a]	.786	.573	.916

a. Predictors: (Constant), PL2, T3, PB1, R3, T1, PB3, R1, T5, P1, R6, T2, R5, R4, PL3, PL1, T4

The Model Summary – This table shows the R, R square and adjusted r –square and the standard error of the estimate. The R value shows the simple correlation between the observed and predicted values of the dependent variable. The R square which is known as the co-efficient of determination is 0.786, this explains that the regression modeled ITD2 (Intention to disclose 2) strongly as 78.6% of the variation in ITD2 is explained by the independent variables.

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	49.472	16	3.092	3.682	.006[b]
	Residual	13.437	16	.840		
	Total	62.909	32			

a. Dependent Variable: ITD2
b. Predictors: (Constant), PL2, T3, PB1, R3, T1, PB3, R1, T5, P1, R6, T2, R5, R4, PL3, PL1, T4

ANOVA shows the how well the regression equation fits the data, it predicts the dependent variable, in this analysis, the regression model predicts the dependent variable properly. Statistically a p value is a number between 0 and 1, and typically a p value that is < 0.05 indicates a strong evidence against a null hypothesis, but a p value of > 0.05 indicates a weak evidence against the null hypothesis. The p value here which is stated in the sig. section of the ANOVA table is 0.006, which indicates a strong support from our declared alpha value for the analysis.

6. DISCUSSION

The international survey respondents indicated that when risk was low, they were more willing to trust the EHR system, and thus more willing to disclose information. Thoughts and concerns for privacy were also quite high and when satisfied respondents were not as concerned with potential error loss. Survey participants did indicate a difference between health privacy and overall internet privacy. Respondents did not consider their own health condition to be a strong enough factor to adopt nor did they indicate they could be influenced by social norms. Perceived benefits of the EHR also played a strong role in willingness to disclose. Familiarity with EHRs neither detracted nor supported the respondent's decision.

This study had several limitations. First, it utilized international college students as surrogates for international decisions. Students are generally younger and in better health than the average international patient. The sample size was small and could not account for differences in countries. Many reports have indicated various levels of adoption among countries so results from this study might not be as generalizable. Future research should replicate this study with separate populations from specific countries. Additionally, future work will look at theoretical models which employ structural equation modeling.

7.CONCLUSION

The healthcare sector has experienced many advances in terms of patient's recordkeeping and welfare. This study sought to aid the health organizations on how to implement the EHR in health institutions. As most patients struggle with providing healthcare information due to increase in healthcare data breaches an analysis was been carried out and from results it can be concluded that what motivates patients to provide their sensitive health information to health providers are includes trust, risk, privacy and perceived benefits. Even though the EHR is fully utilized and mandatory in the US, it is still in its developing stage in countries like Nigeria, Bangladesh and India. This study can help countries who wish to pursue the adoption of EHR. The social implications point what factors influence trust and behavioral intentions to disclose information online. These factors can be considered when enlightening patients on the use of EHRs.

8. REFERENCES

Agaku, I. T., Adisa, A. O., Ayo-Yusuf, O. A., & Connolly, G. N. (2013). Concern about security and privacy, and perceived control over collection and use of health information are related to withholding of health information from healthcare providers. *Journal of the American Medical Informatics Association*, 21(2), 374-378.

Ajami, S., & Arab-Chadegani, R. (2013a). Barriers to implement electronic health records (EHRs). *Materia socio-medica*, 25(3), 213.

Ajami, S., & Bagheri-Tadi, T. (2013b). Barriers for adopting electronic health records (EHRs) by physicians. *Acta Informatica Medica*, 21(2), 129.

Bansal, G., & Gefen, D. (2010). The impact of personal dispositions on information sensitivity, privacy concern and trust in disclosing health information online. *Decision Support Systems*, 49(2), 138-150.

Baratam, S. S., Abdula, C. & Powell, L. (2020). health-related mobile applications for smart devices on insurance providers: aiding customer retention. *Issues of Information Systems*. 21(1), 30-36.

Bowens, F. M., Frye, P. A., & Jones, W. A. (2010). Health information technology: Integration of clinical workflow into meaningful use of electronic health records. *Perspectives in health information management*, *7*. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2966355/

Bresnick, J. (2018). Are Universal EHRs Key to Healthcare Value, Trust, and AI Adoption? *HealthIT Analytics*, Retrieved from https://healthitanalytics.com/news/are-universal-ehrs-key-to-healthcare-value-trust-and-ai-adoption

Caldicott, F. (2013). Information: To share or not to share. *Information Governance Review*, Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/192572/2900774_InfoGovernance_accv2.pdf

Cuk, S., Wimmer, H. & Powell, L. (2017). Problems associated with patient care reports and transferring data between ambulance and hospitals from the perspective of emergency medical technicians. *Issues in Information Systems.* 18(4), 16-26.

Culk, B., Wimmer, H. Powell, L. & Rebman Jr., C. (2018). Electronic emergency medical technician reports – testing perceptions of a prototype. *Issues in Information Systems*. 19(3), 81-91.

Fernández-Alemán, J. L., Señor, I. C., Lozoya, P. Á. O., & Toval, A. (2013). Security and privacy in electronic health records: A systematic literature review. *Journal of Biomedical Informatics*, 46(3), 541-562.

Fragidis, L. L., & Chatzoglou, P. D. (2018). Implementation of a nationwide electronic health record (EHR*). International Journal of Health Care Quality Assurance*, *31*(2), 116-130.

Gunter, T. D., & Terry, N. P. (2005). The emergence of national electronic health record architectures in the United States and Australia: models, costs, and questions. *Journal of Medical Internet Research*, 7(1), e3. Retrieved from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1550638/

Hassol, A., Walker, J. M., Kidder, D., Rokita, K., Young, D., Pierdon, S., Ortiz, E. (2004). Patient experiences and attitudes about access to a patient electronic health care record and linked web messaging. *Journal of the American Medical Informatics Association*, 11(6), 505-513.

Häyrinen, K., Saranto, K., & Nykänen, P. (2008). Definition, structure, content, use and impacts of electronic health records: A review of the research literature. *International Journal of Medical Informatics,* 77(5), 291-304.

Hur, H., Lee, S. R., Xuan, Y., Kim, Y. B., Lim, Y., Cho, Y. K., & Han, S. U. (2012). The effects of helicobacter pylori on the prognosis of patients with curatively resected gastric cancers in a population with high infection rate. *Journal of the Korean Surgical Society*, *83*(4), 203-211.

Kose, I., Rayner, J., Birinci, S., Ulgu, M. M., Yilmaz, I., & Guner, S. (2020). The adoption level of electronic health records in Turkish hospitals and the relation with hospital sizes. Retrieved from https://assets.researchsquare.com/files/rs-8500/v1/manuscript.pdf

Landi, H. (2020). Number of patient records breached nearly triples in 2019. *Fierce Healthcare.* Retrieved from https://www.fiercehealthcare.com/tech/number-patient-records-breached-2019-almost-tripled-from-2018-as-healthcare-faces-new-threats

Malhotra, N. K., Kim, S. S., & Agarwal, J. (2004). Internet users' information privacy concerns (IUIPC): The construct, the scale, and a causal model. *Information Systems Research*, 15(4), 336-355.

Millman, J. (2014). Health care data breaches have hit 30M patients and counting. *Washington Post*, Retrieved from https://www.washingtonpost.com/news/wonk/wp/2014/08/19/health-care-data-breaches-have-hit-30m-patients-and-counting

Ng, B.-Y., Kankanhalli, A., & Xu, Y. (2009). Studying users' computer security behavior: A health belief perspective. *Decision Support Systems,* 46(4), 815-825.

Palabindala, V., Pamarthy, A., & Jonnalagadda, N. R. (2016). Adoption of electronic health records and barriers. *Journal of Community Hospital Internal Medicine Perspectives*, 6(5), 32643.

Papoutsi, C., Reed, J. E., Marston, C., Lewis, R., Majeed, A., & Bell, D. (2015). Patient and public views about the security and privacy of electronic health records (EHRs) in the UK: Results from a mixed methods study. *BMC Medical Informatics and Decision Making*, 15(1), 86.

Sittig, D. F., & Singh, H. (2011). Legal, ethical, and financial dilemmas in electronic health record adoption and use. *Pediatrics*, 127(4), 1042-1047.

Thakkar, M., & Davis, D. C. (2006). Risks, barriers, and benefits of EHR systems: a comparative study based on size of hospital. Perspectives in Health Information Management, 3(5). Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2047303/

Wheatley, B. (2013). Transforming care delivery through health information technology. *The Permanente Journal,* 17(1), 81.

Wixom, B. H., & Todd, P. A. (2005). A theoretical integration of user satisfaction and technology acceptance. *Information Systems Research*, 16(1), 85-102.

Woo, H. E., & Pfeffer, M. A. (2013). A brief introduction to electronic health records and associated terminology. *Proceedings of UCLA Healthcare*, 17, 1-3.

Open Source Software: A Detailed Analysis of Contributions and Quality

Alan Peslak
arp14@psu.edu
Penn State University
Scranton, PA, USA

D. Scott Hunsinger
hunsingerds@appstate.edu
Appalachian State University
Boone, NC, USA

Abstract

Open source software has been an option for many applications since the dawn of computing. Simply put, open source is "software for which the original source code is made freely available and may be redistributed and modified." (Oxford dictionary). But with this free software, there often comes little support and sometimes perceived or actual questionable quality. Our study examines the current attitudes and participation among the developer community towards contributing to open source software as well as the present perceptions of quality among this group. Overall, we find that levels of participation are relatively low but do vary by demographic factors. Also, the perceived levels of quality remain below proprietary/closed source software, but again, demographics and country of origin show much variation.

Keywords: Open source software, quality, contribution, adoption

1. INTRODUCTION

Open source software (OSS) is defined by Oxford dictionary (2020) as "software for which the original source code is made freely available and may be redistributed and modified." Some people may prefer using open source software over proprietary software for multiple reasons, including control, training, security, stability, and community. Developers have more control with OSS because they can examine the code and make changes as desired. Since open source code is publicly accessible, people can study it to become better programmers. Some users perceive OSS to be more secure than proprietary software since updates and fixes can be done without asking permission. OSS may be considered more stable since it can still be updated even if the original developers cease working on the software. In addition. OSS often creates a strong community of users and developers (https://opensource.com/).

Many people use some type of open source software on a daily basis. Table 1 displays examples of some of the most popular open software that has been released.

Wordpress	Magento
Mozilla Firefox	Mozilla Thunderbird
FileZilla	GnuCash
Audacity	GIMP
OpenOffice	VLC
Handbrake	Pidgin
Freemind	Notepad++
7-Zip	Blender
PDFCreator	Calibre
TrueCrypt	Ubuntu

Table 1: Popular OSS Examples

The rest of our paper is arranged as follows: the Literature Review examines previous research in this area, including the usage of several theories including the Technology Acceptance Model (TAM) and Unified Theory of Acceptance and Use of Technology (UTAUT). Section 3 explains the methodology we used for this paper. Section 4 presents our results, while Section 5 provides discussion of these results along with conclusions.

2. LITERATURE REVIEW

A number of previous studies have examined the adoption of OSS. Some of these have used well-known theories such as the Technology Acceptance Model (TAM), Unified Theory of Acceptance and Use of Technology (UTAUT), along with additional constructs and theories, in order to better understand the factors influencing OSS adoption. Others have applied Grounded Theory, systemic literature reviews, frameworks, or other approaches.

Multiple researchers have used variations of TAM in their research regarding OSS adoption. Gallego et al. (2015) developed a research model based upon the Technology Acceptance Model, adding several constructs. The authors discovered that user training, user fit, technological complexity and trainers' support influence the adoption of OSS. Gwebu and Wang (2010) conducted an exploratory study of free open source software (FOSS) users' perceptions, using the Technology Acceptance Model along with other constructs. They identified potential barriers to FOSS adoption and provided recommendations that may increase adoption of FOSS. Gallego et al. (2007) identified the variables and factors that have a direct effect on individual attitude towards OSS adoption by using a variation of the technology acceptance model. Taha et al. (2018) examined the main factors affecting the adoption of OSS in the desktop environment. They administered over 340 questionnaires and found that quality, compatibility, support, and usability are the key factors that influence OSS adoption. Racero et al. (2020) examined students' behavioral intention to use OSS by combining the Technology Acceptance Model and Self-Determination Theory. They used the following constructs: Autonomy, competence, relatedness, perceived ease of use, perceived usefulness and behavioral intention. The results confirmed the positive influence of intrinsic motivations, autonomy and relatedness on the usefulness and ease of use and on behavioral intention to use Open Source Software.

Alrawashdeh et al. (2019) used an integrated model of OSS characteristics and UTAUT to survey 255 individuals working at public and private organizations. Software security, software interoperability, and software quality had a significant impact on performance expectancy. The authors concluded that effort expectancy, performance expectancy, self-efficacy, social influence, software cost, software interoperability, software quality, and software security are all important indicators in OSS acceptance and implementation.

Hauge, Ayala, and Conradi (2010) performed a systematic literature review of the adoption of open source software in software-intensive organizations. They identified 112 papers that provide empirical evidence on how organizations adopt OSS and created a classification framework consisting of six ways in which these organizations adopt OSS. The researchers found that existing research on OSS adoption does not sufficiently describe the context of the companies that are studied. Steinmacher et al. (2015) conducted a systematic literature review on the barriers faced by newcomers to open source software projects. They examined 291 studies using Grounded Theory to categorize the barriers into five groups: Social interaction, newcomers' previous knowledge, finding a way to start, documentation, and technical hurdles. They also classified the problems with regard to their origin into three categories: newcomers, community, or product. In order to examine OSS adoption in commercial firms, Thanasopon (2015) developed a framework consisting of four elements: external environment, organizational, technological, and individual contexts. The author found 14 factors that impact OSS adoption which fits into these four elements. Some of these factors encourage the adoption of OSS, while others are inhibitors.

Gwebu and Wang (2011) looked at the role of social identification in the adoption of OSS. The authors noted that most previous work had focused on OSS adoption at the organizational level; minimal work existed at the individual level. They found that social identification is a key driver of OSS adoption. Marsan, Pare, and Beaudry (2012) applied the socio-cognitive perspective of IT innovation adoption and the organizing vision theory by surveying 271 IT specialists in order to better understand the adoption of OSS in organizations. They classified specialists into two groups: Detractors and supporters. Detractors possess more years of experience but have less exposure to OSS than supporters. The perceptions of IT

specialists are positively associated with their company's openness to OSS adoption and the existence of an organizational policy that favors the adoption of OSS.

Katsamakas and Xin (2019) created a game-theoretic analytical model to explain when organizations adopt open source software applications and platforms and to explore the implications. Their analysis examines whether adoption patterns are socially beneficial. They found that open-source adoption depends upon organizational IT capabilities, network effects, and the fit of OSS with the company's application needs. Their results imply that open-source adoption can be socially inefficient.

Lopez et al. (2015) modeled OSS adoption strategies using a goal-oriented notation, examining objectives and dependencies to explore the consequences of adopting one strategy vs. another. They applied their approach to a large telecommunications company.

Sarrab and Rehman (2014) noted that governments and organizations are beginning to adopt OSS on a large scale. They conducted an empirical study of OSS adoption based upon software quality characteristics. Their research used additional internal quality characteristics for selecting OSS that were added to the dimensions of DeLone and McLean information systems' model. The authors organized the quality characteristics into a hierarchy, in which they list characteristics with three main dimensions of quality: information, service, and system.

Sbaia et al. (2018) mentioned that OSS is being adopted more by both organizations and individuals. They examined multiple OSS adoption models and used a case study approach to determine what information can be automatically retrieved from OSS platforms such as GitHub, SonarCloud, and StackExchange.

Silic and Back (2015) examined the influence of risk factors in the decision-making process for OSS adoption. They surveyed 188 IT decision-makers using an Open Source Risk Adoption Model to look at the perceived IT security risk relationship with the intention to adopt OSS. The authors found that IT security risk significantly influences OSS adoption intention.

Donga et al. (2019) suggest that innovation speed of OSS projects can influence users' interest in downloading and using the software. They used a large-scale panel data set from 7442 OSS projects on SourceForge between 2007 and 2010 and found inverted U-shaped relationships between initial release speed and user downloads, as well as between user downloads and update speed.

Most previous research has used much smaller data sets than what we use in our study.

3. METHODOLOGY

In order to study the current usage of open source software, we used the comprehensive 2019 Stack Overflow survey with over 88,000 respondents. Stack Overflow's annual Developer Survey is the largest and most comprehensive survey of people who code around the world. Each year, they field a survey covering everything from developers' favorite technologies to their job preferences. This year (2019) marked the ninth year they've published their annual Developer Survey results, and nearly 90,000 developers took the 20-minute survey earlier this year. (Stack Overflow, 2019). Despite our survey's broad reach and capacity in forming valuable conclusions, we acknowledge that our results don't represent everyone in the software community evenly. But, the 2019 survey had nearly 90,000 respondents and nearly 70,000 of those respondents were employed as software professionals. Our results include selected questions from the survey as well as detailed demographics available. The results were analyzed using IBM SPSS 26.

Our main research questions are focused on two areas participation and perceived quality:

RQ1 How active is the developer community in open source projects?
RQ1a Are there demographic and geographic differences in the developer community in participation in open source projects?
RQ2 What is the perceived quality of open source projects?
RQ2a Are there demographic and geographic differences in the developer community in perceived quality of open source projects?

4. RESULTS

Two specific questions in the survey asked whether respondents contributed to open source projects and also their opinion of the quality of open source software relative to proprietary or closed source software, basically commercial software:

How often do you contribute to open source?

1. Never
2. Less than once per year
3. Less than once a month but more than once per year
4. Once a month or more often;

and

How do you feel about the quality of open source software (OSS)?

1. OSS is, on average, of HIGHER quality than proprietary / closed source software
2. The quality of OSS and closed source software is about the same
3. OSS is, on average, of LOWER quality than proprietary / closed source software

Table 2 shows the overall means for these two questions based on the 86000+ responses. Overall, participation is slightly better than once a year. But this measure of central tendency is a bit misleading. The largest percentage of respondents did not contribute to open source software as shown in Table 3. Over 36% have NEVER contributed to open source. Overall, though this means that 64% have contributed at some time. 64% have either never or less than once per year but also 64% did at some point. We believe this shows an active participation among the developer community in open source projects. We must note of course that participation in this survey serves as somewhat of a bias and may not represent the entire developer population but we do believe that this result does indicate an active and significant force in the software field.

		OPEN SOURCE PARTICIPATION	OPEN SOURCE QUALITY
N	Valid	88883	86842
	Missing	1	2042
Mean		2.11674	2.32009

Table 2: Means for Open Source Participation and Quality

As noted, many researchers have suggested that open source software may be viewed by the population as of lesser quality than proprietary/closed/commercial software. The mean pf 2.32 suggests a perceived somewhat lower quality since 2 is equal and 3 is lesser. Quality of Open source software was seen by 42% as lower quality (Table 5) but 47% saw as same quality as Proprietary software. Only 11% saw as better. The fact that 47% saw open source software was seen as equal quality suggests that not all view open source software poorly. In fact, nearly half see as equal.

		Freq.	Percent	Valid Percent	Cum. Percent
Valid	Never	32295	36.3	36.3	36.3
	Less than Once per year	24972	28.1	28.1	64.4
	1 mon–1 year	20561	23.1	23.1	87.6
	More per mon.	111055	12.4	12.4	100.0
	Total	88883	100.0	100.0	
Miss.	System	1	.0		
Total		88884	100.0		

Table 3: Open Source Participation

(I)1=Male,2=Female,3=Other	(J)1=Male,2=Female,3=Other	Sig.
1	2	.000
	3	.001
2	1	.000
	3	.000
3	1	.001
	2	.000

Table 4: Post Hoc Analysis

		Freq.	Percent	Valid Percent	Cum. Percent
Valid	Higher Quality	8759	9.9	10.1	10.1
	Same	41527	46.7	47.8	57.9
	Lower Quality	36556	41.1	42.1	100.0
	Total	86842	97.7	100.0	
Miss.	System	2042	2.3		
Total		88884	100.0		

Table 5: Quality Analysis

The other area of our research questions was to explore whether there were demographic and/or geographic differences in participation and perceived quality. Variables available to us in the survey and used to explore for significant differences were Age, Gender, Years programming, Year started programming, Race, Professionals versus non-professionals, and Country. We examined all these variables.

1=Male, 2=Female, 3=Other	Mean	N	Std. Deviation
1	2.13563	77919	1.034646
2	1.82282	6344	.996989
3	2.19306	4439	1.097084
Total	2.11613	88702	1.038480

Table 6: Age and Gender ($p < .001$)

The first variable analyzed was gender. There were 8 categories of gender which we compressed to 3 categories because of low numbers in the neither male nor female identities. The results are in Table 6. Overall males were significantly more likely to participate in open source projects than females. This may suggest a gender bias in contributing to open source groups. There was no such difference with other gender participants. In fact, other gender were significantly higher than either male or female as shown in the Table 4. All variances were significant at $p < .01$.

OPEN SOURCE PARTICIPATION

Age Group	Mean	N	Std. Deviation
5-30	2.09011	47544	1.046185
31-50	2.16646	28722	1.013663
51+	1.98947	2943	.995003
Total	2.11405	79209	1.033548

Table 7: Age Group Analysis

Table 7 shows the open source participation by age group. The highest participation is by the 31-50 age group, followed by the 5-30. (Yes, there was a self-identified respondent age 5). It appears that older, more mature individuals in mid-career or age appear to have skills, time, and/or desire to participate in open source. For all these groups an ANOVA and Post hoc showed $p < .001$.

A between-subjects F test was performed with open source participation as the dependent variable and age group and gender as independent variables. The test showed no interaction effect for the two variables. (Table 8 in Appendix A)

Another research question examined was whether there were demographic differences with regard to perceived quality of open source software relative to closed source. Table 9 shows the results of perceived open source quality by gender. Surprisingly, there are no significant difference in quality based on gender.

OPEN SOURCE QUALITY

1=Male, 2=Female, 3=Other	Mean	N	Std. Deviation
1	2.32075	76422	.647843
2	2.31184	6064	.619027
3	2.31682	4179	.683969
Total	2.31993	86665	.647659

Table 9: Perceived Quality by Gender

Differences based on age group however do exist. The youngest group rates open source significantly more unfavorably than the 31-50 and in turn the 31-50 rate open source significantly more unfavorable than the 51+. The older you are the higher you rate open source quality. The reason for this is unclear. Perhaps older users have had more and longer term exposure to open source. A Between-subject F test showed no interaction effects between age group and gender for this variable.

OPEN SOURCE QUALITY

Age Group	Mean	N	Std. Deviation
5-30	2.33444	46483	.657144
31-50	2.30763	28294	.621326
51+	2.25552	2853	.620683
Total	2.32177	77630	.643239

Table 10: Comparison of Age Groups ($p < .001$)

Within our dataset, there were questions about how many years the respondent has been coding and at what age they started. We anticipated that the more years coding and the younger individuals started coding, the more likely they were to contribute to open source projects and the more years coding more likely to view open source favorably. Tables 11 and 12 analyze these two independent variables.

Coefficients

Model		Unstd. Coefficients B	Unstd. Coefficients Std. Error	Std. Coefficients Beta	t	Sig.
1	(Con.)	2.411	.014		172.633	.000
	Years Code	.007	.000	.061	17.504	.000
	Age 1st Code	-.024	.001	-.115	-32.648	.000

a. Dependent Variable: OPEN SOURCE PARTICIPATION

Table 11

Open source participation was significantly positively correlated with years coding. The

more years coding the higher the participation rate. And the more years coding, the more favorably open source is viewed (reverse scaled). When we examine age first coded however, we see slightly different results. The earlier a respondent started coding, the lower the quality perception. These apparently contradictory measures suggest the more you code but the older you start, the higher the participation and more favorable the perception. This suggests that perhaps open source is undervalued by the younger starters since they may be excluded from exposure at an early age.

Coefficients

Model		Unstd. Coefficients B	Unstd. Coefficients Std. Error	Std. Coefficients Beta	t	Sig.
1	(Con.)	2.370	.009		264.961	.000
	Years Code	-.002	.000	-.028	-7.946	.000
	Age 1st Code	-.002	.000	-.012	-3.395	.001

a. Dependent Variable: OPEN SOURCE QUALITY
*. The mean difference is significant at the .05 level.

Table 12

Professionals are statistically more likely to participate in open source projects but also view open source as of lesser quality. This quality difference could be related to the exposure professionals receive with closed source projects and company's implementing more closed source solutions. Non-professionals often have to utilize more open source products due to cost concerns.

OPEN SOURCE PARTICIPATION

DEV	Mean	N	Std. Deviation
No	2.03288	23204	1.072591
Yes	2.14636	65679	1.024607
Total	2.11674	88883	1.038538

p < .001

Table 13

OPEN SOURCE QUALITY

DEV	Mean	N	Std. Deviation
No	2.29953	22342	.665600
Yes	2.32721	64500	.641071
Total	2.32009	86842	.647579

p < .001

Table 14

The final area we examined was geographical differences. We excluded countries where there were less than 500 respondents to ensure we had a critical mass. The results are shown in Tables 15 and 16. The highest contributions came from some interesting sources with Iran, China, and Bangladesh at the top of the list. This was followed by several European countries, Switzerland, Germany, and the Netherlands. The United States is ranked number 22 in this list of participation rates.

Country	Mean	N	Std. Deviation
Iran	2.43902	738	1.029916
China	2.42018	664	1.044825
Bangladesh	2.31901	605	1.144691
Switzerland	2.29857	978	1.072985
Germany	2.25281	5866	1.027340
Netherlands	2.24298	1852	1.027619
Czech Republic	2.21728	764	1.076334
Austria	2.21216	839	1.054729
India	2.21046	9061	1.106954
Norway	2.17944	574	1.008271
Nigeria	2.16667	522	1.139502
Denmark	2.16370	617	1.026015
France	2.16060	2391	1.051066
Turkey	2.16017	949	1.016558
Australia	2.15554	1903	1.024915
Israel	2.15336	952	1.021128
Finland	2.14652	546	1.024720
Pakistan	2.12134	923	1.093468
Belgium	2.11692	727	1.030567
Sweden	2.09733	1274	0.998007
United States	2.09351	20949	1.024844
Russian Federation	2.08619	1694	1.021163
Hungary	2.08187	513	1.016045
New Zealand	2.08015	524	1.024213
United Kingdom	2.06240	5737	1.032733
Italy	2.06028	1576	0.996907
Spain	2.05611	1604	0.985532
Canada	2.03594	3395	1.008160
Brazil	2.03388	1948	0.972071

Greece	2.02878	556	1.000486
Ukraine	2.01843	868	0.999253
Bulgaria	2.01821	659	0.973651
Poland	2.01457	1922	1.020761
Argentina	1.99458	553	0.978930
Portugal	1.99429	525	0.981686
Ireland	1.97804	501	0.994745
Romania	1.96579	760	0.989477
Indonesia	1.95464	507	1.009791
South Africa	1.94896	627	0.984194
Mexico	1.85981	642	0.927465

Table 15: Highest contributions by country

Belief in the quality of open source by country reveals a much different list. The Russian federation has the highest regard for open source versus closed source software, followed by Ukraine, New Zealand, Poland, Bulgaria. With the exception of New Zealand, these are all former Soviet bloc countries and may reflect their limited resources or lesser trust in "Western" sources.

Country	Mean	N	Std. Deviation
Russian Federation	2.13983	1652	0.641158
Ukraine	2.17882	850	0.664442
New Zealand	2.25832	511	0.641539
Poland	2.26223	1880	0.654016
Bulgaria	2.27315	648	0.679409
South Africa	2.27406	613	0.666615
Mexico	2.27473	637	0.645404
Canada	2.28468	3330	0.619625
Pakistan	2.29385	895	0.746143
Romania	2.29690	741	0.650034
Portugal	2.29709	515	0.619944
Israel	2.29803	916	0.650881
Turkey	2.30270	925	0.714145
Czech Republic	2.30470	745	0.619826
United Kingdom	2.31502	5625	0.619127
Finland	2.31648	534	0.602551
United States	2.31850	20543	0.615852
China	2.32258	651	0.727755
Austria	2.32278	821	0.626122
Iran	2.32489	711	0.735373
Greece	2.32532	541	0.678701
Italy	2.32751	1545	0.643964
Sweden	2.33008	1230	0.617079
Australia	2.33030	1871	0.625600
Switzerland	2.33090	958	0.604844
Belgium	2.33616	708	0.594642
Germany	2.34642	5756	0.609972
Denmark	2.35585	607	0.614909
Bangladesh	2.35986	578	0.762133
Brazil	2.36137	1926	0.657617
Norway	2.36348	564	0.628635
Netherlands	2.36900	1813	0.598996
Nigeria	2.37126	501	0.699923
India	2.38019	8772	0.715249
France	2.39966	2342	0.608188
Spain	2.40280	1569	0.604737
Argentina	2.43203	537	0.622599

Table 16

5. DISCUSSION AND CONCLUSIONS

The study of the use, acceptance, and adoption of open source software has mainly focused on limited datasets. Though our study has limited specificity on reasons for adoption, it is the first comprehensive review on practitioner attitudes and quality perceptions of open source software. In this way, it extends and supports some of the conclusions of prior research. Past research on open source software has focused on relatively small datasets or limited sample population.

Our study found that overall quality perception for open source software is significantly less than closed source. Gallego et al. (2015) and Taha et al. (2018) suggest training and support are key variables in acceptance of OSS. The lack of support inherent in open source may be a key factor in its perceived quality shortfall.

Racero et al. (2020) suggested that intrinsic motivation plays a key role intention to use OSS. Positive exposure to the software may be a path to a higher perceived quality.

Hagu, Ayala, and Conradi did not find clear reasons for lack of OSS adoption. The lack of received quality we found appears to be a fundamental reason.

Participation in OSS was studied by Gwebu and Wang and they found social identification as a key driver of adoption. This social aspect may be missing in many OSS projects and can be addressed.

One of the key influencers of OSS perception is security risk according to Silic and Back (2015). This may be a key underlying factor in our discovered quality shortfall.

There is much to be gained by use of open source software. Cost savings, transparent logic, and worldwide community input all serve as motivators to implement OSS solutions. Many developers are already engaged in contributing to OSS projects. But the numbers are not as robust as they could be. Further research is needed to more fully understand why OSS is not viewed as favorably as closed source software and practices and platforms need to be further refined so that more individuals can contribute to OSS. Wikipedia and its open source knowledge base has replaced many sources of general information. The potential exists for OSS to do likewise for business and personal software.

6. REFERENCES

Adewumi, A., Misra, S., Omoregbe, N., Crawford, B., & Soto, R. (2016). A systematic literature review of open source software quality assessment models. SpringerPlus. 5(1936). doi:10.1186/s40064-016-3612-4

Akatsu, S., Fujita, Y., Kato, T., & Tsuda, K. (2018). Structured analysis of the evaluation process for adopting open-source software. Procedia Computer Science. 126. 1578-1586. doi:10.1016/j.procs.2018.08.131

Alrawashdeh, T., Elbes, M., Almomani, A., ElQirem, F., & Tamimi, A. (2019). User acceptance model of open source software: An integrated model of OSS characteristics and UTAUT. Journal of Ambient Intelligence and Humanized Computing. https://doi.org/10.1007/s12652-019-01524-7

Andersen-Gott, M., Ghinea, G., & Bygstad, B. (2012). Why do commercial companies contribute to open source software? International Journal of Information Management. 32. 106-117. doi:10.1016/j.ijinfomgt.2011.10.003

Antin, J., & Shaw, A. (2012, May 5-10). Social desirability bias and self-reports of motivation: a study of amazon mechanical turk in the US and India. Crowd Sourcing & Peer Production II, Austin, TX, United States.

Boricha, M. (2020, Jun. 28). 18 best open source softwares of 2020. TechArrival. https://techrrival.com/best-open-source-softwares/

Colazo, J., & Fang, Y. (2010). Following the Sun: Temporal Dispersion and Performance in Open Source Software Project Teams. Journal of the Association for Information Systems. 11(11/12). 684-707.

Daniel, S., Maruping, L., Cataldo, M., & Herbsleb, J. (2018). The Impact of Ideology Misfit on Open Source Software Communities and Companies. MIS Quarterly. 42(4). 1069-1096. doi:10.25300/MISQ/2018/14242

Dong, J., Wu, W., & Zhang, Y. (2019). The faster the better? Innovation speed and user interest in open source software. Information & Management. 56. 669-680. https://doi.org/10.1016/j.im.2018.11.002

Edison, H., Carroll, N., Conboy, K., & Morgan, L. (2018). An investigation into inner source software development: Preliminary findings from a systematic literature review. ACM. doi: 10.1145/1235

Ellis, C., Parker, J., (2020, Feb. 27). The best open source software 2020: Free, open software for home of business. Techradar. https://www.techradar.com/best/best-open-source-software

Frej, M., Bach, C., Shock, R., & Desplaines, E. (2015). Open-source software: Adoption and Challenges. American Society for Engineering Education.

Frequently Asked Questions regarding Open Source Software (OSS) and the Department of Defense (DoD). (n.d.). Department of Defense Chief Information Officer. https://dodcio.defense.gov/Open-Source-Software-FAQ/

Frey, K., Luthje, C., & Haag, S. (2011). Whom Should Firms Attract to Open Innovation Platforms? The Role of Knowledge Diversity and Motivation. Long Range Planning, 44. 397-420. doi:10.1016/j.lrp.2011.09.006

Gallego, M., Bueno, S., Racero, F., & Noyes, J. (2015). Open source software: The effects of training on acceptance. Computers in Human Behavior. 49. 390-399. http://dx.doi.org/10.1016/j.chb.2015.03.029

Gallego, M., Luna, P., & Bueno, S. (2008). User acceptance model of open source software. Computers in Human Behavior. 24. 2199-2216. doi:10.1016/j.chb.2007.10.006

Gwebu, K., & Wang, J. (2010). Seeing eye to eye? An exploratory study of free open source software users' perceptions, The Journal of Systems and Software. 83. 2287-2296. doi:10.1016/j.jss.2010.07.011

Gwebu, K., & Wang, J. (2011). Adoption of Open Source Software: The role of social identification. Decision Support Systems. 51. 220-229. doi:10.1016/j.dss.2010.12.010

Iskoujina, Z., & Roberts, J. (2015). Knowledge sharing in open source software communities: motivations and management. Journal of Knowledge Mangement. 19(4). 791-813. doi:10.1108/JKM-10-2014-0446

Katsamakas, E., Mingdi, X. (2019). Open source adoption strategy. Electronic Commerce Research and Applications. 36. 100872. https://doi.org/10.1016/j.elerap.2019.100872

Kaufmann, N., Schulze, T., & Veit, D. (2011). More than fun and money. Worker Motivation in Crowdsourcing – A Study on Mechanical Turk. AMCIS 2011 Proceedings-All Submissions. 340. http://aisel.aisnet.org/amcis2011_submissions/340

Krogh, G., Haefliger, S., Spaeth, S., & Wallin, M. (2012). Carrots and rainbows: Motivation and social practice in open source software development. MIS Quarterly. 36(2). 649-676.

Li, Y., Tan, C., & Yang, X. (2013). It is about what we have: A discriminant analysis of organizations' decision to adopt open source software. Decision Support Systems. 56. 56-62. http://dx.doi.org/10.1016/j.dss.2013.05.006

Lopez, D., Pablos, C., & Santos, R. (2010). Profiling F/OSS adoption modes: An interpretive approach. International Federation for Information Processing. 319. 354-360.

Lopez, L., Costal, D., Ayala, C. Franch, X. Annosi, M., Glott, R., & Haaland, K. (2015). Adoption of OSS components: A goal-oriented approach. Data & Knowledge Engineering. 99. 17-38. http://dx.doi.org/10.1016/j.datak.2015.06.007

Marsan, J. & Pare, G. (2013). Antecedents of open source software adoption in health care organizations: A qualitative survey of experts in Canada. International Journal of Medical Informatics. 82. 731-741. http://dx.doi.org/10.1016/j.ijmedinf.2013.04.001

Marsan, J., Pare, G., & Beaudry, A. (2012). Adoption of open source software in organizations: A socio-cognitive perspective. Journal of Strategic Information Systems. 21. 257-273. http://dx.doi.org/10.1016/j.jsis.2012.05.004

Maruping, L., Sherae, D., Cataldo, M. (2019). Attracting, Onboarding, and Retaining Newcomer Developers in Open Source Software Projects. MIS Quarterly. 43(2). 951-976.

Moqri, M., Mei, X., Qiu, L., & Bandyopadhyay, S. (2018). Effect of "Following" on Contributions to Open Source Communities. Journal of Management Information Systems. 35(4). 1188-1217. https://doi.org/10.1080/07421222.2018.1523605

"Nguyen-Duc, A., Cruzes, D., & Conradi, R. (2015). The impact of global dispersion on coordination, team performance and software quality – A systematic literature review. Information and Software Technology. 57. 277-294.

http://dx.doi.org/10.1016/j.infsof.2014.06.002"

Oram, A. (2011). Promoting Open Source Software in Government: The Challenges of Motivation and Follow-Through. Journal of Information Technology & Politics. 8. 240-252. doi:10.1080/19331681.2011.592059

Oyvind, H., Ayala, C., & Conradi, R. (2010). Adoption of open source software in software-intensive organizations - A systematic literature review. Information and Software Technology. 52. 113-1154. doi:10.1016/j.infsof.2010.05.008

Pickett, P. (2019, Nov. 20). How open-source software works. The Balance Careers. https://www.thebalancecareers.com/what-is-open-source-software-2071941

Qu, W. Yang, Z., & Wang, Z. (2011). Multi-level framework of open source software adoption. Journal of Business Research. 64. 997-1003. doi:10.1016/j.jbusres.2010.11.023

Racero, F., Bueno, S., & Gallego, M. (2020). Predicting students' behavioral intention to use open source software: A combined view of the technology acceptance model and self-determination theory. Applied Sciences. 10. 2711. doi:10.3390/app10082711

Sarrab, M., & Rehman, O. (2014). Empirical study of open source software selection for adoption, based on software quality characteristics. Advances in Engineering Software. 69. 1-11. http://dx.doi.org/10.1016/j.advengsoft.2013.12.001

Sbai, N., Lenarduzzi, V. Taibi, D., Sassi, S., &Ghezala, H. (2018). Exploring information from OSS repositories and platforms to support OSS selection decisions. Information and Software Technology. 104. 104-108. https://doi.org/10.1016/j.infsof.2018.07.009

Silic, M., & Back, A. (2015). The Influence of Risk Factors in Decision-Making Process for Open Source Software Adoption. International Journal of Information Technology & Decision Making. 1(35). https://www.worldscientific.com/worldscinet/ijitdm

Spaeth, S., & Krogh, G. (2015). Perceived Firm Attributes and Intrinsic Motivation in Sponsored Open Source Software Projects. Information Systems Research. 26(1). 224-237. http://dx.doi.org/10.1287/isre.2014.0539

Spinellis, D., Giannikas, V. (2012). Organizational adoption of open source software. The Journal of Systems and Software. 85. 666-682. doi:10.1016/j.jss.2011.09.037

Steinmacher, I. Silva, M. Gerosa, M. & Redmiles, D. (2015). Information and Software Technology. 59. 67-85. http://dx.doi.org/10.1016/j.infsof.2014.11.001

Steinmacher, I., Gerosa, M., & Redmiles, D. (2014). Attracting, Onboarding, and Retaining Newcomer Developers in Open Source Software Projects. ACM.

Taha, A., Abbood, A., Abdulrazzaq, A., & Albahri, A. (2018). Identifying the affecting factors of open source software in IT community. Journal of Engineering and Applied Sciences. 13(14). 5771-5780. doi: 10.3923/jeasci.2018.5771.5780

Thanasopon, B. (2015). Open source software (OSS) adoption in commercial firms: A literature review. KMTL Information Technology Journal. http://journal.it.kmitl.ac.th

Van Loon, A., & Toshkov, Dimiter. (2015). Adopting open source software in public administration: The importance of boundary spanners and political commitment. Government Information Quarterly. 32. 207-215. http://dx.doi.org/10.1016/j.giq.2015.01.004

Walker, T. (2017, Jan. 15). 20 Most popular open source software ever. Tripwire Magazine. https://www.tripwiremagazine.com/20-most-popular-open-source-software-ever-2/

Weiling, K., & Zhang, P. (2010). The Effects of Extrinsic Motivations and Satisfaction in Open Source Software Development. Journal ot the Association for Information Systems. 11. 784-808.

What is open source?. (2020). OpenSource. https://opensource.com/resources/what-open-source

Appendices and Annexures

Appendix A – Table 8

Tests of Between-Subjects Effects

Dependent Variable: OPEN SOURCE PARTICIPATION

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	Hypothesis	16320.922	1	16320.922	321.623	.003
	Error	102.699	2.024	50.745[a]		
Age Group	Hypothesis	29.394	2	14.697	23.570	.000
	Error	25.682	41.187	.624[b]		
Gender	Hypothesis	140.635	2	70.317	89.992	.000
	Error	124.202	158.954	.781[c]		
Age Group * Gender	Hypothesis	1.307	4	.327	.309	.872
	Error	83693.914	79036	1.059[d]		

a. .717 MS(GEN2) + .283 MS(Error)
b. .595 MS(AgeGroup * GEN2) + .405 MS(Error)
c. .379 MS(AgeGroup * GEN2) + .621 MS(Error)
d. MS(Error)