

https://dspace.org/

AJIC Issue 25, 2020 - Full Issue - print-on-demand version

2020-06-30
LINK Centre, University of the Witwatersrand (Wits), Johannesburg
https://hdl.handle.net/10539/29198
The African Journal of Information and Communication (AJIC). (2020). Issue 25. Retrieved from http://wiredspace.wits.ac.za/handle/10539/29189

Downloaded from a DSpace repository. https://dspace.org/

THE AFRICAN JOURNAL OF INFORMATION AND COMMUNICATION (AJIC)

ISSUE 25, 2020



RESEARCH ARTICLES

The Digitalised Terrorism Ecology: A Systems Perspective
Nixon Muganda Ochara, Nancy Achieng Odhiambo and Armstrong Kadyamatimba

Application of Machine Learning Classification to Detect Fraudulent E-wallet Deposit Notification SMSes
Fillemon S. Enkono and Nalina Suresh

A Supplementary Tool for Web-archiving Using Blockchain Technology
John E. de Villiers and André Calitz

Factors Impacting Tanzanian Rural and Peri-urban Drug Dispensaries' Perceived Benefits from Using an mHealth Reporting System
Alistair Elias and Joel S. Mtebe

Pilot Testing of an Information Extraction (IE) Prototype for Legal Research
Brenda Scholtz, Thashen Padayachy and Oluwande Adewoyin

Cyber-Threat Information-Sharing Standards: A Review of Evaluation Literature
Nenekazi N. P. Mkuzangwe and Zubeida C. Khan

PUBLICATION REVIEW

Book Review: International Telecommunications Law and Policy
Pontian N. Okoli

Published by the LINK Centre
University of the Witwatersrand (Wits)
Johannesburg, South Africa
https://www.wits.ac.za/linkcentre

ISSN 2077-7213 (online version)
ISSN 2077-7205 (print version)

AJIC 25 - 5-July-20 - 20h14.indd 1 7/5/2020 8:22:45 PM

The African Journal of Information and Communication (AJIC)

Issue 25, 2020

Published by the LINK Centre, School of Literature, Language and Media (SLLM)
Faculty of Humanities, University of the Witwatersrand (Wits)
Johannesburg, South Africa
https://www.wits.ac.za/linkcentre/ajic

Published since 2000, *The African Journal of Information and Communication (AJIC)* is a peer-reviewed, interdisciplinary, open access academic journal focused on information and communication ecosystems in Africa, elsewhere in the developing world, and at global level. Accredited by the South African Department of Higher Education and Training (DHET), *AJIC* pursues its open access objective by publishing online, free to the user, under a Creative Commons licence, and by not imposing article processing charges on its contributors.

Editorial Advisory Board

The journal is supported by an international editorial advisory board, comprising:
Lucienne Abrahams, University of the Witwatersrand, Johannesburg, South Africa
Olufunmilayo Arewa, Temple University, Philadelphia, US
Erik de Vries, HAN University of Applied Sciences, Nijmegen, The Netherlands
Nagy K. Hanna, author and international development strategist, Washington, DC, US
Tawana Kupe, University of Pretoria, South Africa
Manoj Maharaj, University of KwaZulu-Natal, Durban, South Africa
Gillian Marcelle, Resilience Capital Ventures, Washington, DC, US
Uche M. Mbanaso, Nasarawa State University, Keffi, Nigeria
Caroline B. Ncube, University of Cape Town, South Africa
Nixon Muganda Ochara, University of Venda, Thohoyandou, South Africa
Tunji Oloapa, Ibadan School of Government and Public Policy (ISGPP), Nigeria
Carlo M. Rossotto, World Bank Group, Washington, DC, US
Ewan Sutherland, University of the Witwatersrand, Johannesburg, South Africa
Hossana Twinomurinzi, University of South Africa, Pretoria, South Africa

Editors

Managing Editor: Tawana Kupe, Vice-Chancellor, University of Pretoria, South Africa, tawana.kupe@up.ac.za

Corresponding Editor: Lucienne Abrahams, Director, LINK Centre, Faculty of Humanities, University of the Witwatersrand, PO Box 601, Wits 2050, Johannesburg, South Africa, ajic.submissions@gmail.com

Publishing Editor: Chris Armstrong, Research Associate, LINK Centre, University of the Witwatersrand, Johannesburg, South Africa, chris.armstrong@wits.ac.za

AJIC 25 - 5-July-20 - 20h14.indd 2-1 7/5/2020 8:22:45 PM

Peer-Reviewing
AJIC acknowledges with gratitude the following peer reviewers of submissions published, or considered for publication, in this issue: Jason Cohen, Barry Dwolatzky, Warren Hero, Edgar Jembere, Willard Munyoka, Loyiso Nongxa, Brett van Niekerk, Kiru Pillay, Trishana Ramluckan, Ewan Sutherland and Barend Taute.

Production
Sub-editing: LINK Centre
Proofreading: Linda Van de Vijver
Desktop-publishing: LINK Centre

This work is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) licence: http://creativecommons.org/licenses/by/4.0

ISSN 2077-7213 (online version)
ISSN 2077-7205 (print version)

AJIC is published by the LINK Centre, School of Literature, Language and Media (SLLM), Faculty of Humanities, University of the Witwatersrand (Wits), PO Box 601, Wits 2050, Johannesburg, South Africa. The LINK Centre is located at the Wits Tshimologong Digital Innovation Precinct, 41 Juta St., Braamfontein, Johannesburg, https://www.tshimologong.joburg

Past issues of *AJIC*, and its precursor *The Southern African Journal of Information and Communication (SAJIC)*, are available at https://www.wits.ac.za/linkcentre/ajic and https://www.wits.ac.za/linkcentre/sajic

CONTENTS

RESEARCH ARTICLES

The Digitalised Terrorism Ecology: A Systems Perspective
Nixon Muganda Ochara, Nancy Achieng Odhiambo and Armstrong Kadyamatimba

Application of Machine Learning Classification to Detect Fraudulent E-wallet Deposit Notification SMSes
Fillemon S. Enkono and Nalina Suresh

A Supplementary Tool for Web-archiving Using Blockchain Technology
John E. de Villiers and André Calitz

Factors Impacting Tanzanian Rural and Peri-urban Drug Dispensaries' Perceived Benefits from Using an mHealth Reporting System
Alistair Elias and Joel S. Mtebe

Pilot Testing of an Information Extraction (IE) Prototype for Legal Research
Brenda Scholtz, Thashen Padayachy and Oluwande Adewoyin

Cyber-Threat Information-Sharing Standards: A Review of Evaluation Literature
Nenekazi N. P. Mkuzangwe and Zubeida C. Khan

PUBLICATION REVIEW

Book Review: International Telecommunications Law and Policy
Pontian N. Okoli

AJIC 25 - 5-July-20 - 20h14.indd 2-3 7/5/2020 8:22:45 PM

RESEARCH ARTICLES

AJIC 25 - 5-July-20 - 20h14.indd 4-5 7/5/2020 8:22:46 PM

The Digitalised Terrorism Ecology: A Systems Perspective

Nixon Muganda Ochara
Research Professor, School of Management Sciences, University of Venda, Thohoyandou, South Africa; and Research Associate, LINK Centre, University of the Witwatersrand (Wits), Johannesburg
https://orcid.org/0000-0001-5736-7901

Nancy Achieng Odhiambo
Doctoral Candidate, Department of Business Information Systems, School of Management Sciences, University of Venda, Thohoyandou, South Africa
https://orcid.org/0000-0002-3123-7305

Armstrong Kadyamatimba
Dean and Professor, School of Management Sciences, University of Venda, Thohoyandou, South Africa
http://orcid.org/0000-0002-9638-0858

Abstract
This study uses a systematic review methodology to interpret existing literature on the digital dimensions of contemporary terrorism and counter-terrorism. Using the theory of synergetics as a guiding analytical framework, the study conducts meta-synthesis of relevant literature, including application of soft systems methodology (SSM), in order to generate conceptualisation of a digitalised terrorism ecology. This ecology comprises five interacting sub-systems: open digital infrastructure; digital information ecology; digital terrorism enactment; digital capabilities; and digital enslavement.

Keywords
digital technologies, digitalisation, terrorism, counter-terrorism, synergetics, soft systems methodology (SSM), sociomateriality, digitalised terrorism ecology, open digital infrastructure, digital information ecology, counter power, digital terrorism enactment, digital enslavement

DOI: https://doi.org/10.23962/10539/29196

Recommended citation
Ochara, N. M., Odhiambo, N. A., & Kadyamatimba, A. (2020). The digitalised terrorism ecology: A systems perspective. *The African Journal of Information and Communication (AJIC), 25*, 1-19. https://doi.org/10.23962/10539/29196

AJIC 25 - 5-July-20 - 20h14.indd 6-1 7/5/2020 8:22:46 PM

This article is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) licence:
https://creativecommons.org/licenses/by/4.0

1. Introduction

Countering terrorism remains a priority for national governments throughout the world. The Global Terrorism Index (GTI) report of 2018 reported that deaths from terrorism events fell by 27% between 2016 and 2017, a third consecutive year of decline. However, the same GTI report stated that every region of the world had recorded a higher average impact of terrorism in 2017 compared to 2002. Evidence also suggests that the nature of terrorism is becoming increasingly complex. For instance, while the dominant view of global terrorism is that it is largely fuelled by Muslim extremism, the 15 March 2019 attack on two Muslim mosques in New Zealand pointed to the increasing prominence of far-right movements as a source of terrorism (Hawi, Osborne, Bulbulia, & Sibley, 2019). Such instances involving right-wing extremists suggest the changing nature of terrorism, with drivers anchored not only in religion, but also in culture, the economy, politics, globalisation, and various forms of media (Sirgy, Estes, El-Aswad, & Rahtz, 2019).

While the nature of the terrorism ecosystem has been changing, research and counter-terrorism responses have retained a strong focus on Al-Qaeda and jihadist terrorism generally; have remained event-driven; and have under-emphasised the roles of digital technologies, state-sponsored terrorism, and right-wing extremists (Asongu, Nnanna, Biekpe, & Acha-Anyi, 2019; Schuurman, 2019). This under-emphasis on the actual current dynamics of terrorism provided the justification for this study, which sought to develop a nuanced understanding of the contemporary terrorism ecology, with particular attention to the ecology's digital dimensions.

McLuhan, writing in the 1960s (see McLuhan, Gordon, Lamberti, & Scheffel-Dunand, 2011), foresaw that technological ecosystems are not passive containers, and are, rather, active processes that reshape people and technologies alike. This insight is particularly meaningful in our contemporary world, in which, in July 2019, it was estimated that more than 4.33 billion people (approximately 56% of the world's population) were active internet users (Statista, 2019). Such pervasiveness of digital technologies cannot help but have a reshaping dynamic in human life.

Digital ecology approach

We adopted a digital ecology approach for this study (see García-Marco, 2011), which is an approach that can unravel and capture the complexity of the various interacting sub-systems and processes that characterise terrorism's macro-systems. The digital ecology approach evolved from the notion of information ecology, and is a systems approach that seeks to understand the ways in which societies and their knowledge and communication are being shaped by digital technologies. The approach has been used, inter alia, to analyse the evolution of the World Wide Web (Huberman, 2003), digital libraries (García-Marco, 2011), social communities on the internet (Finin et al., 2008), and e-government (Ochara, 2014). Our study applied the digital ecology approach in order to develop an understanding of how digitalisation is influencing the structuring of terrorism ecologies.

Theory of synergetics

We used the theory of synergetics (Haken, 1984) as a structuring device for our study. Synergy, in the systems theory context, denotes a conceptual or mathematical product of causes (or factors), and is used in many sciences as a general model to account for non-linear change (Schmitt, Eid, & Maes, 2003). According to Haken (1984), the five core properties of the theory are:

- order parameters (macroscopic patterns);
- control parameters;
- internal and external system constraints;
- internal and external parameters and system elements; and
- environment.

As will be seen below, we used these five core properties of synergetics as the main tools for structuring our research and findings.

2. Research methodology

We used the systematic literature review method in order to identify the most relevant literature on how digitalisation interweaves with terrorism and counter-terrorism. Both peer-reviewed sources (e.g., journal articles, scholarly books), and non-peer-reviewed sources (non-scholarly books, media articles, blogs, reports, websites and videos) were used as data sources for the systematic literature review. We used scholarly electronic databases such as Elsevier, ProQuest, SAGE Publications and Web of Science for peer-reviewed content, and Google for non-peer-reviewed sources. Table 1 presents the key search terms that were used.

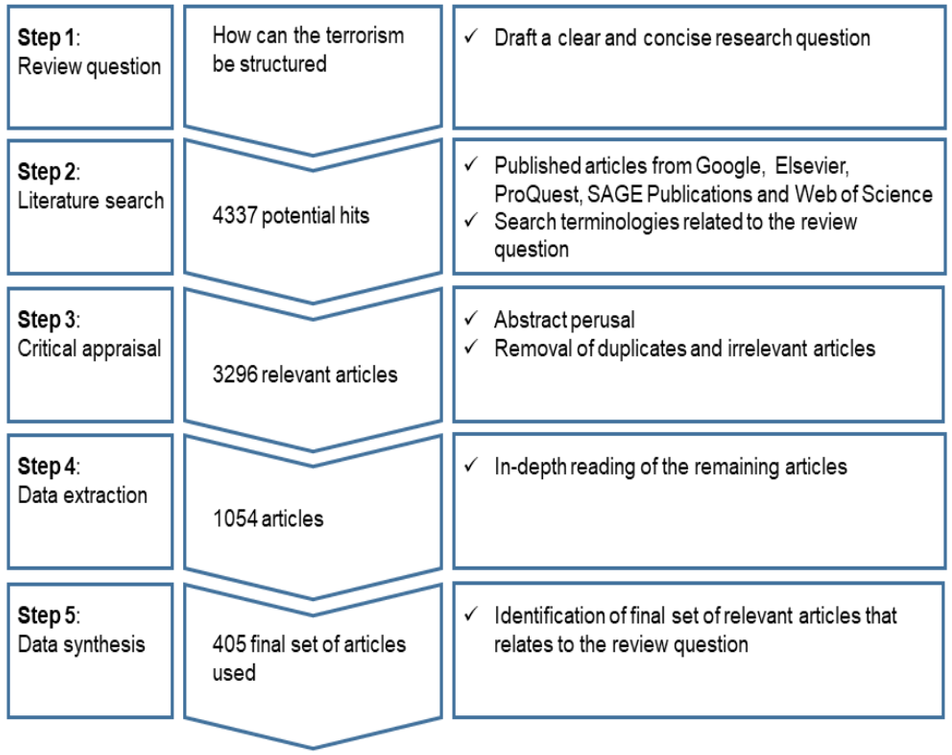
AJIC 25 - 5-July-20 - 20h14.indd 2-3 7/5/2020 8:22:46 PM

Table 1: Key search terms

Terrorism	Extremism	Ideology AND Terrorism
Counter-Terrorism	Radicalization	Definition of Terrorism
Religion AND Terrorism	Policies AND Terrorism	Technology AND Terrorism
Root Causes of Terrorism	Forms of Terrorism	Counter-Terrorism AND Technology
Radicalization AND Terrorism	Poverty AND Terrorism	Terrorism AND Facebook
Terrorism AND modus operandi	Governance AND Terrorism	Terrorism AND Twitter
Terrorism AND Internet	Terrorism AND YouTube	Terrorism AND Social Media
Counter-Terrorism AND Twitter	Counter-Terrorism AND Training	Counter-Terrorism AND Social Media
Research AND Terrorism	Europe AND Terrorism	Terrorism AND USA
Africa AND Terrorism	Human Rights and Terrorism	Terrorism Organizations
Terrorism AND ISIS	Terrorism AND Al-Qaeda	Terrorism AND Osama
Terrorism AND Al-Shabaab	Terrorist Groups	Terrorism AND Propaganda
Terrorism AND Recruitment	Terrorism AND Training	Counter-Terrorism AND Organisational Culture
Counter-Terrorism AND Collaboration	Terrorism AND Media	Counter-Terrorism AND Institutional Rivalry
Counter-Terrorism AND Organisational Structure	Counter-Terrorism AND Policy	Counter-Terrorism AND Human Rights
Community Policing AND Terrorism	Terrorism and Boko Haram	Counter-Terrorism AND Boko Haram
Counter-Terrorism AND Media	Counter-Terrorism AND Facebook	Counter-Terrorism AND YouTube

The search focused mainly on English-language content published since 11 September 2001, the date of the "9/11" terrorist attacks in the United States that fundamentally changed both academic and non-academic treatments of the topic of terrorism. Our search produced 4,337 documents. We applied an inclusion and exclusion process through which we removed duplicates, scanned through the titles and abstracts of the identified documents, and checked through the references sections of the identified documents to identify whether more documents could be included in the list. This process resulted in a final dataset of 405 documents (see Figure 1).

Figure 1: Process followed to identify the 405 documents



Meta-synthesis

We then evaluated the articles via meta-synthesis, a qualitative methodology for synthesising outcomes of several studies that are similar in topic or outcome measure (see Park & Gretzel, 2007). Informed by Ochara's (2013) guidelines for application of theory in research, our meta-synthesis was focused on dividing the 405 documents according to their relevance to the five components of the theory of synergetics (Haken, 1984) outlined above: order parameters (macroscopic patterns); control parameters; internal and external system constraints; internal and external parameters and system elements; and environment.

Thematic analysis

All 405 documents were then subjected to thematic analysis, a qualitative technique that consists of delineating the data according to its thematic patterns and then assigning codes to the themes (see Alhojailan, 2012). For this thematic analysis, we used the NVivo software, and applied the CATWOE soft systems methodology (SSM). CATWOE is a mnemonic representation of six elements that, when

AJIC 25 - 5-July-20 - 20h14.indd 4-5 7/5/2020 8:22:46 PM

considered, can aid conceptualisation of problems, and responses to problems (Checkland, 1981, pp. 224–225):

- customers;
- actors;
- transformation;
- worldview;
- owners; and
- environmental constraints.

We chose an SSM methodology because such methodologies are suitable for addressing real-world, complex, and poorly-structured problems—and terrorism is a problem of that sort. Terrorism is often associated with, inter alia, fundamentalism and radicalisation (see Hafez & Mullins, 2015), religious intolerance (see Doosje et al., 2016), and political and civil instability (see Kagwanja, 2006).

Rittel and Webber (1973) refer to certain problems that are unstructured or poorly structured, and that are irregular, complex, novel and adaptive in nature, as "wicked". The problem of terrorism fits this category. Such problems require that they first be clearly defined and structured before a resolution can be proposed (Dunn, 2015), and SSM tools help in the process of defining and structuring problems.

3. Findings

Meta-synthesis results

Table 2 shows the distribution of the 405 documents when we delineated them according to their relevance to the five components of the theory of synergetics.

Table 2: Data categorisation in terms of theory of synergetics

Components	First-order themes	No. of articles
Order parameters (macroscopic patterns)	✓ dominant ideology ○ philosophy ○ root causes ✓ policies ✓ open models of information-sharing ○ collaboration ○ sharing economy ✓ open technology platforms	94
Control parameters	✓ regional economic imbalance ○ poverty and grievances ○ extremism ✓ poor governance ✓ technology (media and internet)	73
Internal and external system constraints	✓ institutional rivalry ✓ organisational perspectives ○ norms and values ○ roles and responsibilities	18
Internal and external parameters and system elements	✓ complexity of inter-organisational cognitive behaviour ✓ complexity of cross-organisational competencies ✓ collective intelligence ✓ open governance models ○ collaboration ○ transparency ✓ participation	71
Environment	✓ definitions of terrorism ✓ multiple agencies ✓ changing modes of attack ✓ improved access to, and availability of, technology	149

The distribution of these articles was considered sufficient to allow for the SSM approach to be applied.

Thematic analysis results

Braun, Clarke, and Terry (2014) call for three steps to be followed when conducting thematic analysis of data: familiarisation, generation of initial codes, and theme search. We applied the six elements (presented above) of the CATWOE SSM in performing these three steps. We produced a thematic word cloud visualisation (Figure 2) based on the initial thematic "nodes" generated by the NVivo qualitative data analysis software.

Figure 2: Thematic word cloud



AJIC 25 - 5-July-20 - 20h14.indd 6-7 7/5/2020 8:22:46 PM

Our thematic analysis also produced a table summarising high-level themes (Table 3) in terms of the six CATWOE categories, with the focus of the thematic analysis being the digitalised terrorism ecosystem.

Table 3: Second-order themes identified in terms of CATWOE

C	Customers	government, state, community, individuals, terrorists, Al-Qaeda, ISIS, people, public, group, organisations, society, companies, human, university, international, countries
A	Actors	technology, social media, Twitter, Facebook, computer and internet, community, government, websites, organisations, video, system, digital, political, police, press, human, university, media, social, computer, people, individuals, society, community
T	Transformation	policy, strategy, communication, frameworks, analysis, content, activities, technologies, data, work
W	Worldview	Islamic, political, national, privacy, rights, counter-terrorism, surveillance, terrorism
O	Owners	technology, social media, Twitter, Facebook, computer and internet, government, community, organisations, videos, journalists, companies, system, digital, police, press, university, media, information, social, online, computer
E	Environmental constraints	support, attacks, material, violence, services, intelligence, effective, strategy, threat, techniques, news, approach, information, terror, privacy, enforcement, program, cyber terrorism, war, AI, struggle, propaganda

In a CATWOE SSM analysis, *customers* are the beneficiaries or victims affected by the system under consideration (Checkland & Scholes, 1990). What becomes apparent in Table 3 above is the wide range of individuals and groups who are either beneficiaries or victims within the digitalised terrorism ecosystem. In CATWOE, *actors* are the agents who carry out the main activities of a system (Checkland & Scholes, 1990). In Table 3, what is apparent in this category is the dominance of digital technology themes.

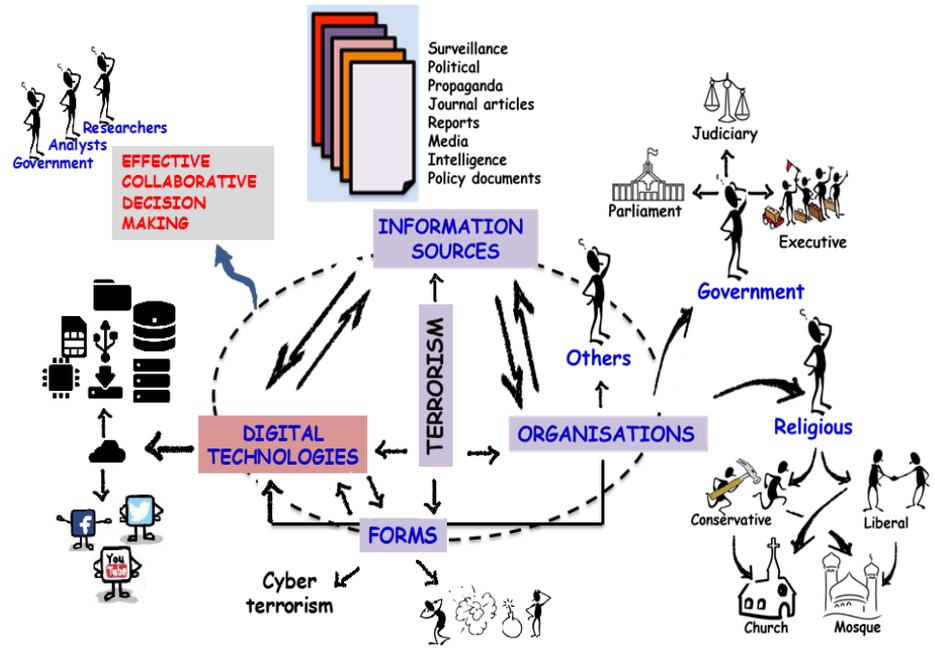
The third aspect of CATWOE, *transformation*, focuses on purposeful activities undertaken by the actors with the aim of solving a problem (Checkland & Scholes, 1990). In the context of the digitalised terrorism ecosystem, we treated transformation as a process, or set of artefacts, that influences the evolution, nature and forms of interactions among the multiple stakeholders in the ecosystem.

The *worldview* in CATWOE is connected to an individual's worldview and beliefs, and gives transformation meaning (Bergvall-Kåreborn, Mirijamdotter, & Basden, 2004). We see in Table 3 that themes related to the ethos of religion, politics, nationalism, digital rights, and responses to the terrorism problem were prevalent

in the 405 documents analysed. The *owners* are those with the power to either stop the transformations or allow them to take place (Checkland & Scholes, 1990). *Environmental constraints* are considered as the internal or external limitations that can hinder transformation (Bergvall-Kåreborn, Mirijamdotter & Basden, 2004). In Table 3, we see a multitude of constraints identified.

This SSM-oriented thematic analysis also produced the "rich picture" shown in Figure 3, which we used as a precursor to the conceptual model based on the theory of synergetics (Figure 4).

Figure 3: Rich picture



In Figure 3, we see, from a problem-structuring perspective, the pervasive role of digital technology in human systems—a finding that resonates with the sociomateriality literature (see Cecez-Kecmanovic et al., 2014), and with how technology affordances and capabilities play a critical role in agency (Weißenfels, Ebner, Dittes, & Smolnik, 2016). For instance, Daniel, Hartnett, and Meadows (2017) forcefully argue that social media's democratising affordances are evident in the transformation of power structures from top-down to much more bottom-up power.

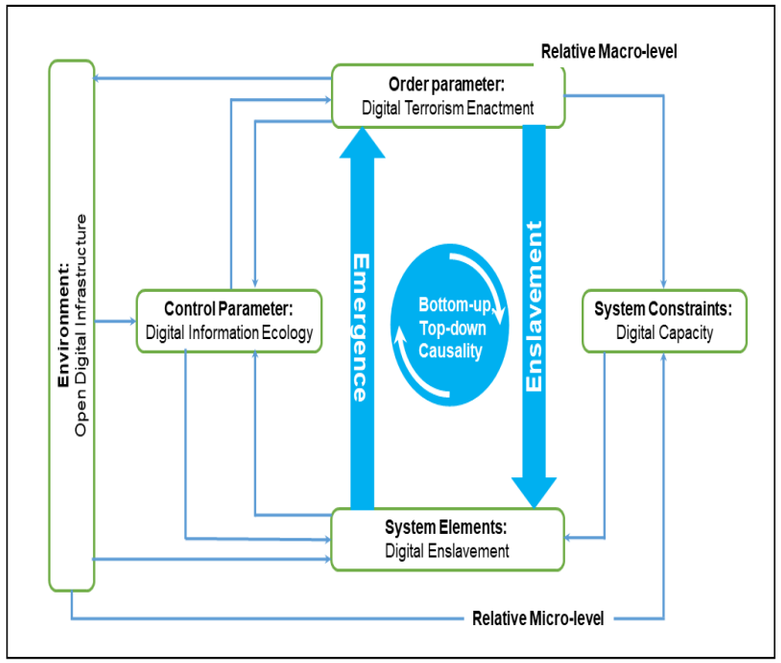
AJIC 25 - 5-July-20 - 20h14.indd 8-9 7/5/2020 8:22:46 PM

Figure 3 also accentuates the importance of an understanding of information sources as essential for the development of state understandings of new, varied forms of terrorism, and of evolving state policies and strategies for countering terrorism. With increased online access to information, terrorists can develop counter-strategies to governments' counter-terrorism efforts. In addition, increased online information dissemination has empowered terrorist groupings such as ISIS and Al-Qaeda in their efforts to advance propaganda and recruit new members.

Conceptual model of the digitalised terrorism ecology

Grounded in the theory of synergetics, Figure 4 depicts the *digitalised terrorism ecology* model that we derived from the thematic analysis, with the model set out in terms of the five core properties, as explained earlier, of the theory of synergetics.

Figure 4: Digitalised terrorism ecology



Viewed through the synergetics lens, the documents provided evidence that the structuring process of terrorism starts with an external activation from what we term, in Figure 4, the *digital information ecology*. This digital information ecology context kindles a behavioural change process among terrorism genres at the level of system elements, which in turn leads, or may lead, to the emergence of new terrorism policies (e.g., community policing, counter-terrorism strategies) at the macroscopic level of the order parameter. These new terrorism policies result, in turn, in what we term, in

Figure 4, *digital enslavement* of the system elements (e.g., individual terrorists who feel compelled to adopt particular patterns of behaviour).

4. Analysis

We now analyse the findings of the systematic literature review in terms of components from the theory of synergetics deployed in Figure 4, in the following order:

- environment;
- control parameter;
- order parameter;
- system constraints; and
- system elements.

Environment: Open digital infrastructure

A critical underlying principle for digitalised self-organisation is the openness of systems, which allows for power to be added throughout the environment. In our analysis, the *open digital infrastructur*e is the core construct that undergirds the emergence of terrorism in its current form. Even a casual analysis reveals that in all aspects (severity, magnitude, operations), the "structure" of terrorism has become more complex since its emergence in the 1960s and 1970s. We attribute this dramatic change in the nature of terrorism to the advent of the internet infrastructure (Denning, 2010; Benson, 2014), due in large part to its characteristic openness, which enhances networking and increases terrorism capability.

The identification above of the central role of digital technologies confirms the dominant view that the emergence of wider digital sociomaterial infrastructures such as the internet are reshaping society (Merali, 2006), with implications for increased global terrorism (Gillespie, Osseiran, & Cheesman, 2018). Therefore, we view the current sociomaterial digital infrastructure as the environmental nexus that continues to "structure" the terrorism ecology.

The open nature of digital infrastructure is recognised as underpinning newer digitalisation paradigms such as cloud computing, big data, transparent computing, and nomadic computing (Odero, Ochara, & Quenum, 2017). To understand the nature of the influence of open digital infrastructure on terrorism ecologies, we adopt the perspective of technology affordance to reify the potentials of the emerging features of such newer digital technologies and infrastructures. From a relational perspective, based on social interactions that shape and are shaped by the technology and the context (Carugati, Fernández, Mola, & Rossignoli, 2018), the evolving open digital infrastructure for terrorism emphasises the fact that the focus should not be on the features that digital technologies possess, but rather on how actors' goals and capabilities can be related to the inherent potential offered by the features (Autio, Nambisan, Thomas, & Wright, 2018).

AJIC 25 - 5-July-20 - 20h14.indd 10-11 7/5/2020 8:22:46 PM

Thus, the locus of the environmental aspect of the open digital infrastructure places, at the fore, how open and flexible affordances of digital technologies enable "newer" manifestations of the digital information ecology (the control parameter, as explained below), which occasion change in behaviour within terrorism ecologies. The dominant affordances that facilitate the engagement of terrorism ecologies with the open digital infrastructure as an enabler of the digital information ecology are social media vehicles such as Facebook, Twitter, and Instagram (see the CATWOE and rich picture analyses above), empowered by social software and algorithms, and enabled by smart devices (Tuten & Mintu-Wimsatt, 2018). Certainly, prior research entrenches social media vehicles as the core proxy for open digital infrastructure (Paganini, 2016; Blaker, 2015; Plantin, Lagoze, Edwards, & Sandvig, 2018), with characteristics of the sharing economy and open digital technology platforms influencing contemporary information ecologies of terrorists. The sharing economy, or gig economy, is seen as the most predominant contemporary organising principle influencing how human beings share knowledge and assets, enabled by the increased connectivity, scale, speed, and transparency of the internet (Burtch, Carnahan, & Greenwood, 2018).

Control parameter: Digital information ecology and digital counterpower

The environmental influence of open digital infrastructure enables the control parameter: the *digital information ecology*. The evolution of the digital information ecology, as a control parameter, can be explained using the concept of digital "counterpower" as used by Castells. According to Castells (2011, p. 773), "[…] wherever there is power, there is counterpower, enacting the interests and values of those in subordinate positions". We use the notion of digital counterpower to interpret the link between the first-order themes that emerged from application of the theory of synergetics and the second-order themes that emerged from application of CATWOE.

The first-order themes that were identified (Table 2) emphasised poor governance and regional economic imbalances that fuel and entrench poverty, grievances, and extremism. Furthermore, technology (particularly digital media and the internet) emerged as a forceful equaliser, empowering billions with access to information. The second-order emergent themes (identified via CATWOE) were strongly linked to the worldview component, which was characterised by "Islamic, political, national, privacy, rights, counter-terrorism, surveillance, terrorism". While the first-order themes were related to "economic imbalances and poor governance", thus pointing to the "root causes" of terrorism (Newman, 2006), the second-order themes pointed to emergent "worldviews" that are increasingly mediated by digital technologies (Bertram, 2016).

The control parameter can thus be viewed through the lens of *digital counterpower*, as a legitimisation tool that links the emergence of terrorism genres to the forceful

mediation of internet sources of information that shape the worldview of the terrorist. As Hajj, McEwan, and Turkington (2019) found out, internet usage influences the information ecology of individuals, thus shaping their worldviews. In our analysis, the terrorist emerges as part of a digital counterpower movement—as a form of political violence, predominantly shaped by the information ecology of the internet as an alternative voice to mainstream worldviews.

Order parameter: Digital terrorism enactment

The first-order themes that emerged from the meta-synthesis for order parameters (Table 2) captured three dominant concepts: dominant ideology (philosophy, root causes); policies; and open models of information-sharing (collaboration, sharing economy, and open technology platforms). The second-order themes (identified via CATWOE) revolved around the notion of transformation. From the methodological perspective of SSM, the notion of transformation focuses on purposeful activities undertaken by actors within a particular ecosystem with the aim of solving a problem (Checkland & Scholes, 1990).

While the control parameter themes, linked to the digital information ecology, were found to be positioned as contributing to a legitimation process in which the notion of digital counterpower is used to rationalise terrorism as an alternative form of political violence, the order parameter themes appear to capture *digital terrorism enactment*, i.e., the mobilisation activities of terrorism and counterterrorism. Mobilisation, in the manner used by Swanson and Ramiller (1997), serves the dynamic function of activating, motivating and structuring the forces that emerge to support the material realisation of new acts of terrorism. Mobilisation implies that actors (both terrorists and counter-terrorism agencies) look to the terrorism ecosystem for the resources required to realise their agendas. For instance, both terrorists and counter-terrorism actors are focused on manipulating algorithms that underpin digital information ecology to spread propaganda and enact responses (Ammar, 2019). The emergence, explained above, of digital counterpower as part of the digital information ecology control parameter could be seen to explicate the root causes of terrorism. At the same time, the emergence of alternative worldviews could be seen as structuring the digital terrorism enactment order parameter. It is this latter perspective that envisions the emergence of terrorism and counter-terrorism through various transformation activities objectified in dominant terrorism ideologies, policies, and strategies.

System constraints: Digital capabilities

Under the CATWOE analysis, we likened the constraining nature of various external and internal systems to "owners", in terms of the Checkland and Scholes (1990) characterisation of owners as those with the power to either stop transformations or allow them to take place. The role of digital technologies in the CATWOE analysis points to the visible role of the relational view of power (Doolin & McLeod, 2012)

AJIC 25 - 5-July-20 - 20h14.indd 12-13 7/5/2020 8:22:46 PM

as enabling or constraining the digital capabilities of players within the terrorism ecology.

The differentiated *digital capabilities* of counter-terrorism agencies allow, on one hand, a category of these agencies to be efficient in responding to terrorism incidents because they have developed the critical infrastructure, the weaponisation of information, and the targeted use of social media messaging (Omand, 2018). On the other hand, institutional separation in the security sector is deeply entrenched, and is linked to the separation of roles, responsibilities and (digital) resource endowments, which foments institutional competition (Campana & Légaré, 2010). From a terrorism perspective, successful terrorism incidents are in part shaped by their digital capabilities, for instance, their utilisation of the latest communication devices to recruit, share and distribute information, and to mobilise support, particularly through the internet (Carty & Barron, 2019). Thus, for both terrorists and counter-terrorism agencies, digital capability remains a fundamental source of power and counter-power.

System elements: Digital enslavement

The theory of synergetics envisages a dominant order parameter emerging from competing order parameters to "enslave" systems elements (see Figure 4). As this dominant order parameter emerges, the "enslaving" process entrenches system elements through which individuals and collectives (e.g., security agencies, terrorism cells) are both enabled and constrained, by digital capabilities, to participate in an ecology that gives rise to terrorism and counter-terrorism activities. In the case under examination here, the order parameter (digital terrorism enactment) determines the behaviour, via digital enslavement, of the numerous individual elements. For instance, at the macroscopic level of the order parameter, various scholars have claimed that religion, particularly Islamist ideology, is used to foment global terrorism, with the counter-terrorism response undergirded by the belief that Islamist ideology is a root cause of terrorism (Bartolucci, 2019; Schuurman, 2019).

Thus, at the macroscopic order parameter level, a view of terrorism as emanating from Islamic philosophy has stabilised. And we saw in the analysis earlier in this article that the formation of this dominant view of the root causes of terrorism is influenced by open models of information-sharing (collaboration, sharing economy, and open technology platforms). This stabilised narrative of the nature of terrorism structures and influences the behaviour of individual systems elements (such as terrorism cells, counter-terrorism agencies) at the microscopic level, by means of the process we regard as *digital enslavement*.

5. Conclusions

Our systematic literature review has demonstrated that viewing terrorism as separate from the broader digitalisation of society would not only be imprudent but would also constrain societies' ability to counter terrorism effectively. Rather, viewing terrorism as a self-organising digital ecology, comprising a complex web of interacting agents, users, and technologies, allows for better conceptualisation.

Our review resulted in the identification of five key sub-systems interacting within the digitalised terrorism ecology:

- open digital infrastructure;
- digital information ecology and digital counterpower;
- digital terrorism enactment;
- digital capabilities; and
- digital enslavement.

These five sub-systems act to define the sociomaterial nature of the terrorism ecology, influencing the emergence of terrorism and counter-terrorism responses. Viewed in terms of these sub-systems, the digitalised terrorism ecology comprises a complex coexistence of the five sub-systems, centred on the open digital infrastructure, with this infrastructure typified by the "sharing economy" whose modalities are at play within the digital platforms that foment terrorism (Teigland, Holmberg, & Felländer, 2019). This sharing economy has inspired the emergence of the digital information ecology for individuals and collectives, empowering them to consider and appropriate previously unknown worldviews and notions of digital counterpower. The emergence of dominant worldviews, through the process of enactment, then structures and enslaves the emergence of various terrorism and counter-terrorism genres. While the prevailing view is that there is an increase in lone-actor terrorist acts (Sela-Shayovitz & Dayan, 2019), the influence of the digitalised terrorism ecology on individuals and collectives over time needs to be the prime foundation for understanding terrorism.

This study has also provided methodological and theoretical insights. The methodological approach that was adopted demonstrated the possibility of using SSM to inform qualitative thematic analysis. Furthermore, analytical theorising based on synergetics, a theory based on self-organisation, allowed for the unpacking of the complex problem of terrorism from a systems perspective. This combination of SSM and synergetics was found to be of value in furthering our understanding of how the contemporary digital terrorism ecology has emerged. While the claims made in this article are of necessity tentative due to the limits of a meta-synthesis, further empirical research can focus on testing the insights that have emerged.

AJIC 25 - 5-July-20 - 20h14.indd 14-15 7/5/2020 8:22:46 PM

References

Alhojailan, M. I. (2012). Thematic analysis: A critical review of its process and evaluation. *West East Journal of Social Sciences*, *1*(1), 39–47.

Ammar, J. (2019). Cyber gremlin: Social networking, machine learning and the global war on Al-Qaida- and IS-inspired terrorism. *International Journal of Law and Information Technology, 27*(3), 238–265. https://doi.org/10.1093/ijlit/eaz006

Asongu, S. A., Nnanna, J., Biekpe, N., & Acha-Anyi, P. N. (2019). Contemporary drivers of global tourism: Evidence from terrorism and peace factors. *Journal of Travel & Tourism Marketing*, *36*(3), 345–357. https://doi.org/10.1080/10548408.2018.1541778

Autio, E., Nambisan, S., Thomas, L. D. W., & Wright, M. (2018). Digital affordances, spatial affordances, and the genesis of entrepreneurial ecosystems. *Strategic Entrepreneurship Journal*, *12*(1), 72–95. https://doi.org/10.1002/sej.1266

Bartolucci, V. (2019). The discourse on terrorism of Donald Trump. In J. Kowalski (Ed.), *Reading Donald Trump: A parallax view of the campaign and early presidency* (pp. 127–147). Cham, Switzerland: Springer. https://doi.org/10.1007/978-3-319-93179-1_7

Benson, D. C. (2014). Why the internet is not increasing terrorism. *Security Studies, 23*(2) 293–328. https://doi.org/10.1080/09636412.2014.905353

Bergvall-Kåreborn, B., Mirijamdotter, A., & Basden, A. (2004). Basic principles of SSM modeling: An examination of CATWOE from a soft perspective. *Systemic Practice and Action Research*, *17*(2), 55–73. https://doi.org/10.1023/b:spaa.0000018903.18767.18

Bertram, L. (2016). Terrorism, the internet and the social media advantage: Exploring how terrorist organizations exploit aspects of the internet, social media and how these same platforms could be used to counter violent extremism. *Journal for Deradicalization*, *7,* 225–252.

Blaker, L. (2015). The Islamic State's use of online social media. *Military Cyber Affairs*, *1*(1), 4. https://doi.org/10.5038/2378-0789.1.1.1004

Braun, V., Clarke, V., & Terry, G. (2014). Thematic analysis. In P. Rohleder & A. C. Lyons (Eds.), *Qualitative research in clinical and health psychology* (pp. 95–114). Basingstoke, UK: Palgrave Macmillan.

Burtch, G., Carnahan, S., & Greenwood, B. N. (2018). Can you gig it? An empirical examination of the gig economy and entrepreneurial activity. *Management Science*, *64*(12), 5497–5520. https://doi.org/10.1287/mnsc.2017.2916

Campana, A., & Légaré, K. (2010). Russia's counterterrorism operation in Chechnya: Institutional competition and issue frames. *Studies in Conflict & Terrorism*, *34*(1), 47–63. https://doi.org/10.1080/1057610x.2011.531458

Carty, V., & Barron, F. G. R. (2019). Social movements and new technology: The dynamics of cyber activism in the digital age. In B. Berberglu (Ed.), *The Palgrave handbook of social movements, revolution, and social transformation* (pp. 373–397). Cham, Switzerland: Springer. https://doi.org/10.1007/978-3-319-92354-3_16

Carugati, A., Fernández, W., Mola, L., & Rossignoli, C. (2018). *My choice, your problem ? Mandating IT use in large organisational networks*. *Information Systems Journal, 28*(1), 6–47. https://doi.org/10.1111/isj.12120

Castells, M. (2011). *The rise of the network society: The information age: Economy, society, and culture, volume 1* (2nd ed.). Hoboken, NJ: John Wiley & Sons.

Cecez-Kecmanovic, D., Galliers, R. D., Henfridsson, O., Newell, S., & Vidgen, R. (2014). The sociomateriality of information systems: Current status, future directions. *MIS Quarterly*, *38*(3), 809–830. https://doi.org/10.25300/misq/2014/38:3.3

Checkland, P., & Scholes, J. (1990). *Soft systems methodology in action*. Hoboken, NJ: John Wiley & Sons.

Checkland, P. (1981). *Systems thinking, systems practice*. Hoboken, NJ: John Wiley & Sons.

Daniel, E., Hartnett, E., & Meadows, M. (2017). Don't throw rocks from the side-lines: A sociomaterial exploration of organizational blogs as boundary objects. *Information Technology & People*, *30*(3), 542–561. https://doi.org/10.1108/itp-02-2015-0036

Denning, D. E. (2010). Terror's web: How the internet is transforming terrorism. In Y. Jewkes, & M. Yar, (Eds.), *Handbook of internet crime* (pp. 194–213). Milton, UK: Willan Publishing.

Doolin, B., & McLeod, L. (2012). Sociomateriality and boundary objects in information systems development. *European Journal of Information Systems*, *21*(5), 570–586. https://doi.org/10.1057/ejis.2012.20

Doosje, B., Moghaddam, F. M., Kruglanski, A. W., De Wolf, A., Mann, L., & Feddes, A. R. (2016). Terrorism, radicalization and de-radicalization. *Current Opinion in Psychology*, *11*, 79–84. https://doi.org/10.1016/j.copsyc.2016.06.008

Dunn, W. N. (2015). *Public policy analysis: An integrated approach* (6th ed.). Abingdon, UK: Routledge.

Finin, T., Joshi, A., Kolari, P., Java, A., Kale, A., & Karandikar, A. (2008). The information ecology of social media and online communities. *AI Magazine*, *29*(3), 77–92. https://doi.org/10.1609/aimag.v29i3.2158

García-Marco, F. J. (2011). Libraries in the digital ecology: Reflections and trends. *Electronic Library*, *29*(1), 105–120. https://doi.org/10.1108/02640471111111460

Gillespie, M., Osseiran, S., & Cheesman, M. (2018). Syrian refugees and the digital passage to Europe: Smartphone infrastructures and affordances. *Social Media + Society*, *January-March,* 1–12. https://doi.org/10.1177/2056305118764440

Hafez, M., & Mullins, C. (2015). The radicalization puzzle: A theoretical synthesis of empirical approaches to homegrown extremism. *Studies in Conflict & Terrorism*, *38*(11), 958–975. https://doi.org/10.1080/1057610x.2015.1051375

Hajj, N., McEwan, P. J., & Turkington, R. (2019). Women, information ecology, and political protest in the Middle East. *Mediterranean Politics*, *24*(1), 62–83. https://doi.org/10.1080/13629395.2017.1380116

Haken, H. (1984). Can synergetics be of use to management theory? In H.-U. Gilbert, & J. B. Probst (Eds.), *Self-organization and management of social systems: Insights, promises, doubts, and questions* (pp. 33–41). Berlin: Springer. https://doi.org/10.1007/978-3-642-69762-3_3

Hawi, D., Osborne, D., Bulbulia, J., & Sibley, C. G. (2019). Terrorism anxiety and attitudes toward Muslims. *New Zealand Journal of Psychology*, *48*(1), 80–89.

Huberman, B. A. (2003). *The laws of the web: Patterns in the ecology of information*. Cambridge, MA: MIT Press.

Kagwanja, P. (2006). Counter-terrorism in the Horn of Africa: New security frontiers, old strategies. *African Security Studies*, *15*(3), 72–86. https://doi.org/10.1080/10246029.2006.9627608

McLuhan, M., with Gordon, W. T., Lamberti, E., & Scheffel-Dunand, D. (2011). *The Gutenberg galaxy: The making of typographic man*. Toronto: University of Toronto Press.

Merali, Y. (2006). Complexity and information systems: The emergent domain. *Journal of Information Technology*, *21*(4), 216–228. https://doi.org/10.1057/palgrave.jit.2000081

AJIC 25 - 5-July-20 - 20h14.indd 16-17 7/5/2020 8:22:46 PM

Newman, E. (2006). Exploring the root causes of terrorism. *Studies in Conflict & Terrorism*, *29*(8), 749–772. https://doi.org/10.1080/10576100600704069

Ochara, N. M. (2014). Towards a regional ontology for e-participation: An ecological view. In P. M. Sebina, K. H. Moahi & K. J. Bwalya (Eds.), *Digital access and e-government: Perspectives from developing and emerging countries* (pp. 60–72). Hershey, PA: IGI Global. https://doi.org/10.4018/978-1-4666-5868-4.ch005

Ochara, N. M. (2013). Linking reasoning to theoretical argument in information systems research. *AIS Electronic Library (AISeL)*, 1–11. Retrieved from http://aisel.aisnet.org/cgi/viewcontent.cgi?article=1501&context=amcis2013

Odero, K., Ochara, N. M., & Quenum, J. (2017). Towards big data-driven logistics value chains for effective decision making and performance measurement. In *Proceedings of the 11th European Conference on Information Systems Management (ECISM 2017), Genoa, Italy, 14-15 September*. https://doi.org/10.2139/ssrn.2960510

Omand, D. (2018). The threats from modern digital subversion and sedition. *Journal of Cyber Policy*, *3*(1), 5–23. https://doi.org/10.1080/23738871.2018.1448097

Paganini, P. (2016). The role of technology in modern terrorism. Madison, WI: InfoSec Institute.

Park, Y. A., & Gretzel, U. (2007). Success factors for destination marketing web sites: A qualitative meta-analysis. *Journal of Travel Research*, *46*(1), 46–63. https://doi.org/10.1177%2F0047287507302381

Plantin, J.-C., Lagoze, C., Edwards, P. N., & Sandvig, C. (2018). Infrastructure studies meet platform studies in the age of Google and Facebook. *New Media & Society*, *20*(1), 293–310. https://doi.org/10.1177/1461444816661553

Rittel, H. W., & Webber, M. M. (1973). 2.3 planning problems are wicked problems. In N. Cross (Ed.), *Developments in design methodology.* Chichester, UK: John Wiley & Sons.

Schmitt, M., Eid, M., & Maes, J. (2003). Synergistic person x situation interaction in distributive justice behavior. *Personality and Social Psychology Bulletin*, *29*(1), 141–147. https://doi.org/10.1177/0146167202238379

Schuurman, B. (2019). Topics in terrorism research: Reviewing trends and gaps, 2007–2016. *Critical Studies on Terrorism*, *12*(3), 463–480. https://doi.org/10.1080/17539153.2019.1579777

Sela-Shayovitz, R., & Dayan, H. (2019). Female Palestinian terrorists: The role of the Intifada period and the terrorism context. *Studies in Conflict & Terrorism*, 1–18. https://doi.org/10.1080/1057610x.2019.1575027

Sirgy, M. J., Estes, R. J., El-Aswad, E.-S., & Rahtz, D. R. (2019). Proposed response: Counterterrorism strategies focusing on the demand side of the terrorism market. In *Combatting Jihadist terrorism through nation-building: A quality-of-life perspective* (pp. 149–173). Cham, Switzerland: Springer. https://doi.org/10.1007/978-3-030-17868-0_8

Statista. (2019). Global digital population as of July 2019 [Web page]. Retrieved from https://www.statista.com/statistics/617136/digital-population-worldwide/

Swanson, E. B., & Ramiller, N. C. (1997). The organizing vision in information systems innovation. *Organization Science*, *8*(5), 458–474. https://doi.org/10.1287/orsc.8.5.458

Teigland, R., Holmberg, H., & Felländer, A. (2019). The importance of trust in a digital Europe: Reflections on the sharing economy and blockchains. In A. B. Engelbrekt, N. Bremberg, A. Michalski, & L. Oxelheim (Eds.), *Trust in the European Union in challenging times* (pp. 181–209). Cham, Switzerland: Springer. https://doi.org/10.1007/978-3-319-73857-4_9

Tuten, T., & Mintu-Wimsatt, A. (2018). Advancing our understanding of the theory and practice of social media marketing: Introduction to the special issue. *Journal of Marketing Theory and Practice*, *26*(1–2), 1–3. https://doi.org/10.1080/10696679.2018.1393277

Weißenfels, S., Ebner, K., Dittes, S., & Smolnik, S. (2016). Does the IS artifact matter in sociomateriality research? A literature review of empirical studies. In *2016 49th Hawaii International Conference on System Sciences (HICSS)*, Koloa, 5–8 January. https://doi.org/10.1109/hicss.2016.252

AJIC 25 - 5-July-20 - 20h14.indd 18-19 7/5/2020 8:22:47 PM

Application of Machine Learning Classification to Detect Fraudulent E-wallet Deposit Notification SMSes

Fillemon S. Enkono
Technologist, Department of Aeronautical Science, School of Military Science, University of Namibia
https://orcid.org/0000-0002-0891-4654

Nalina Suresh
Lecturer, Information Technology Department, School of Computing, University of Namibia
https://orcid.org/0000-0002-9846-7199

Abstract
Fraudulent e-wallet deposit notification SMSes designed to steal money and goods from m-banking users have become pervasive in Namibia. Motivated by an observed lack of mobile applications to protect users from such deceptions, this study evaluated the ability of machine learning to detect the fraudulent e-wallet deposit notification SMSes. The naïve Bayes (NB) and support vector machine (SVM) classifiers were trained to classify both ham (desired) SMSes and scam (fraudulent) e-wallet deposit notification SMSes. The performances of the two classifier models were then evaluated. The results revealed that the SVM classifier model could detect the fraudulent SMSes more efficiently than the NB classifier.

Keywords
m-banking, e-wallets, short message service messages (SMSes), deposit notification, fraud, ham SMSes, scam SMSes, detection, machine learning, classifiers, naïve Bayes (NB), support vector machine (SVM), classification accuracy (CA), feature extraction, feature selection

DOI: https://doi.org/10.23962/10539/29195

Recommended citation
Enkono, F. S., & Suresh, N. (2020). Application of machine learning classification to detect fraudulent e-wallet deposit notification SMSes. *The African Journal of Information and Communication (AJIC), 25*, 1-13.
https://doi.org/10.23962/10539/29195

This article is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) licence:
https://creativecommons.org/licenses/by/4.0

AJIC 25 - 5-July-20 - 20h14.indd 1-2
7/5/2020 8:22:47 PM

1. Introduction and research problem

In the past decade, there has been significant growth of spam (unwanted messages) in email and short message service messages (SMSes), and a contemporaneous growth in the capabilities of mobile banking (m-banking) (Almeida et al., 2011; Shaikh & Karjaluoto, 2015). Among its capabilities, m-banking allows users to utilise mobile phones to make payments from their bank accounts to other users' electronic wallet (e-wallet) accounts. It also allows payment recipients to receive notification SMSes that acknowledge payments into their e-wallets. The widespread use of SMS notifications to acknowledge e-wallet deposits has served to establish such notifications as a trusted means of proof-of-payment for m-banking users. This trust has tended to make e-wallet recipients complacent, to the extent that they may neglect to verify the legitimacy of a payment upon receipt of a deposit notification SMS. This habit of non-verification has fostered the emergence of fraudulent SMSes that use false e-wallet deposit notifications in an attempt to deceive and defraud m-banking users (Arde, 2012; *Erongo*, 2016). It has been reported that m-banking users in Namibia, its neighbour South Africa, and other developing countries have suffered substantial losses, of both money and goods, through falling victim to fraudulent e-wallet deposit notification SMSes (Arde, 2012; Christopher & Kar, 2018; *Erongo*, 2016; Nagel, 2015).

This study was motivated by an observed lack of mobile applications for detecting fraudulent e-wallet deposit notification SMSes in order to safeguard m-banking users from the associated frauds. Our study, conducted in Namibia, evaluated the application of two machine learning classifiers to distinguish between ham (legitimate) SMSes and scam (fraudulent e-wallet deposit notification) SMSes. The classifiers tested were the naïve Bayes (NB) and support vector machine (SVM) models. The evaluation aimed to determine which of the two models could detect the fraudulent e-wallet deposit notification SMSes more efficiently. The ultimate aim was to establish which model would be a good candidate for implementation as an application for detecting fraudulent notifications on users' mobile devices.

2. E-wallet deposit notification SMS fraud

In the typical scenario, the fraudster first obtains the m-banking user's mobile number from a source such as a website, a Facebook notice, or an advertisement. The fraudster then forges an e-wallet deposit notification SMS that purports to acknowledge the deposit of a certain amount and sends it to the m-banking user. The fraudster follows up with a call or SMS, and claims to have mistakenly deposited the amount specified in the bogus notification SMS into the user's e-wallet. The fraudster then asks the targeted user to refund the money via an e-wallet payment. The fraudster may seek to make the refund more palatable to the user by asking for only a portion of the amount supposedly deposited, and may also use other social engineering tricks to lure the user into falling for the fraud.

If the targeted user does not verify that the funds that the fraudster claims to have sent have actually been paid into their e-wallet, the user may fall for the fraud and make a payment from their account to the fraudster's e-wallet. In cases where the targeted user falls for the fraud, the fraudster simply withdraws the money and discards the SIM card used. Fraudsters use similar tricks in respect of goods, e.g., sending bogus e-wallet deposit notification SMSes to salespersons pretending to have paid for the goods, thus seeking to obtain goods without actually paying for them.

Since such frauds are relatively easy to detect and avoid under normal circumstances, fraudsters take advantage of specific situations in order to increase the target's vulnerability—e.g., communicating with someone who has had a death in the family, or with an online seller—as seen in Figure 1, which provides examples of fraud in Namibia and South Africa. Among the factors that create an enabling environment for e-wallet fraudsters are low SIM card costs and ready access to e-wallet services by anyone with a SIM.

Figure 1: Namibian and South African online postings on e-wallet payment notification fraud



AJIC 25 - 5-July-20 - 20h14.indd 3-4 7/5/2020 8:22:47 PM

3. Literature review

SMS spam, smishing, and machine learning (ML) classification

Approaches to combatting spam email and SMSes generally involve the use of machine learning classification (see Akbari & Sajedi, 2015; Choudhary & Jain, 2017). Initial machine learning classification approaches to detecting SMS spam largely treated spam as a generalised set of undesired SMSes, without detection and delineation of specific types of spam (Abdulhamid et al., 2017). As malicious SMS phishing, a form of cyber-fraud known as "smishing", has become more prominent, approaches to spam detection have had to become more focused. The emergence of smishing has led to investigation of machine learning classifiers designed to specifically detect smishing SMSes (Goel & Jain, 2018; Jain & Gupta, 2018).

Further complicating the picture has been the emergence of legitimate SMSes with characteristics that overlap with those of spam and that could easily be erroneously treated as spam by existing spam detection or filtering systems (Reaves, Blue, Tian, Traynor, & Butler, 2016). These legitimate SMSes include those sent for purposes of advertising and promotion, and SMSes for verification codes and for password-reset codes, both of which users can receive from sender numbers that they could not have known beforehand (Reaves et al., 2016). This makes targeted approaches to detecting specific types of SMS spam even more necessary, especially for the types of spam, such as the e-wallet deposit notification frauds, that have the potential to cause users significant losses. The problem of fraudulent e-wallet deposit notification SMSes is relatively new, despite its reported pervasive manifestations, and this tends to explain why our literature survey did not manage to identify any work done on the application of machine learning classification to specifically detect fraudulent e-wallet deposit notification SMSes.

SMS datasets, feature extraction, and machine learning classification

Studying or employing machine learning classification of SMSes requires access to appropriate datasets of ham SMSes and spam SMSes (Abdulhamid et al., 2017). Methods often used to obtain the required dataset include obtaining SMSes from public databases (Ahmed et al., 2014), extracting SMSes from public web-based sources (Almeida et al., 2011), and collecting SMSes directly from users (Shahi & Yadav, 2014).

A study by Cormack (2008) explains that prior to applying machine learning to classify textual content, the content must first be represented as a collection of features derived from the text or from extrinsic information related to the text. Feature extraction is often employed to capture textual features for classifying texts, and the feature extraction process frequently produces multi-dimensional feature sets. It then becomes necessary to employ feature selection, so as to eliminate the features that are less significant in the classification of the text.

Various machine learning classifiers have been employed to classify SMSes for filtering or detecting spam. Some of the classifiers that are extensively used include naïve Bayes (NB), Random Forest (RF) and support vector machine (SVM) (Ahmed et al., 2014; Hedieh et al., 2016; Nagwani & Sharaff, 2017; Nuruzzaman et al., 2011). The NB classifier is widely used for SMS classification due to its simplicity and speed, while the common use of SVM and RF tends to be motivated by their high classification accuracy (CA)—often reported to be in ranges above 90% (Nagwani & Sharaff, 2017). For our study, we chose to test and compare the accuracy of the NB and SVM classifiers.

4. Methodology

The study employed an experimental research design, and used the Weka open source data mining software platform for the experiments.

The SMS dataset

We collected a dataset of 240 unique SMSes from Namibian m-banking users: 184 ham (i.e., normal and legitimate) SMSes and 56 scam (i.e., fraudulent) e-wallet deposit notification SMSes. The ham SMSes included legitimate e-wallet deposit notification SMSes. The ham SMSes, and some of the scam e-wallet deposit notification SMSes, were solicited from volunteers via invitations sent out on Facebook. The majority of the scam e-wallet deposit notification SMSes were extracted from user posts on public Facebook group (M-banking users habitually share examples of scam e-wallet deposit notification SMSes in online fora in order to warn other users). We then represented the 240 raw SMSes in terms of three attributes and a class specification, as outlined in Table 1.

Table 1: Attributes used to represent sample of raw SMSes

Attribute or Class	Description
senderNumLen	The length (total number of digits) of the mobile number of the SMS sender. (Banking institutions use short SMS codes to send legitimate e-wallet deposit notification SMSes, while fraudsters use normal (e.g., 10-digit) mobile numbers.)
content	The string of content in the body of the SMS.
contentLen	The length (number of characters) of the SMS body. Ham SMSes tend to have fewer characters than deposit notification SMSes (both legitimate and fraudulent).
smsClass	Whether the SMS is in the ham class or the scam e-wallet deposit notification class.

AJIC 25 - 5-July-20 - 20h14.indd 5-6 7/5/2020 8:22:47 PM

Feature extraction

We found that the three initial attributes used to define the raw SMSes—*senderNumLen*, *content* and *contentLen*—were insufficient to effectively classify the ham SMSes and the scam e-wallet deposit notification SMSes. We also found that the SMSes' contents contained numerous textual features that could be used to improve the classification. Hence, we employed Weka's StringToWordVector unsupervised filter in order to extract normalised word and term features from the SMSes' contents, for use, in addition to the *senderNumLen* and *contentLen* features, in classifying the SMSes.

Optimal classification features

A total of 1,223 features were extracted from the contents of the 240 SMSes in the dataset. With the addition of the *senderNumLen* and *contentLen* features, the set of features numbered 1,225. Weka's information gain (IG) feature selection algorithm was then used to select a subset of the 1,225 features that allowed optimal classification of the SMSes. The optimal classification features were determined by applying feature selection using different IG threshold (IG_{thsbld}) values. Table 2 shows the number of features selected using different IG_{thsbld} in the [0.0, 1.0] range, and the two classifier models' respective average CA for 10-fold cross-validation.

Table 2: IG_{thsbld}, number of selected features, and the classifiers' average CA

IG_{thsbld}	No. of selected features	Average CA	
		NB	SVM
0.000	119	0.9711	0.9876
0.025	48	0.9711	0.9917
0.050	25	0.9545	0.9628
0.075	22	0.9545	0.9628
0.100	21	0.9545	0.9628

Table 2 shows that a subset of 48 features, selected with IG_{thsbld} = 0.025, allowed the NB and SVM models to classify the SMSes with optimal CA. These 48 features allowed the NB classifier model to maintain CA = 0.9711, its highest observed CA; and the SVM model to achieve CA = 0.9917, its highest recorded CA.

Table 3 shows the 48 word and term features that allowed the NB and SVM models to optimally classify the SMSes, along with their IG values, which indicate how much information each feature contributed towards the correct classification of SMSes that contain them.

Table 3: Features and their IG values

Feature	Feature IG value	Feature	Feature IG value	Feature	Feature IG value
e-wallet	0.6416	f	0.365	with	0.0368
na	0.6158	expires	0.3431	on	0.0368
dial	0.615	at	0.2779	your	0.035
*140*392#	0.6134	for	0.2682	of	0.035
sent	0.6028	you	0.132	will	0.035
contentLen	0.5421	362626	0.1281	can	0.0331
00	0.5392	the	0.0822	am	0.0331
select	0.5372	i	0.0686	me	0.0313
fnb	0.5309	valid	0.0626	*140*999#	0.029
expired	0.5233	it	0.0519	are	0.0277
press	0.5233	senderNumLen	0.0475	or	0.0277
proceed	0.5233	queries	0.0437	2350	0.026
services	0.5233	06129922	0.0437	1600	0.026
atm	0.498	and	0.0424	14	0.026
pin	0.498	to	0.0401	please	0.0259
new	0.4754	16hrs	0.0393	in	0.0259

Classifier models' training and evaluation

The dataset of the 240 SMSes, having been defined using the 48 optimal features, was then used for training and evaluating the NB and SVM classifier models. Supervised learning was employed, following a 10-fold cross-validation approach. The evaluation gave most weight to the models' capability to detect fraudulent e-wallet deposit notification SMSes, and was based on the following metrics adopted from works by Abdulhamid et al. (2017), Mahmoud and Mahfouz (2012), and Hedieh et al. (2016):

True positives (TP): The number of scam e-wallet deposit notification SMSes that are correctly classified.

True negatives (TN): The number of ham SMSes that are correctly classified.

False positives (FP): The number of ham SMSes that are falsely classified.

False negatives (FN): The number of scam e-wallet deposit notification SMSes that are falsely classified.

AJIC 25 - 5-July-20 - 20h14.indd 7-8 7/5/2020 8:22:47 PM

False positives rate (FPR): The rate of ham SMS misclassification.

$$FPR = \frac{FP}{FP+TN} FPR = \frac{FP}{FP+TN} \quad \text{equation (1)}$$

False negatives rate (FNR): The rate of scam e-wallet deposit notification SMSes misclassification.

$$FNR = \frac{FN}{FN+TP} FNR = \frac{FN}{FN+TP} \quad \text{equation (2)}$$

Classification accuracy (CA): The ratio of correctly classified SMSes to the total number of input SMSes.

$$CA = \frac{TP+TN}{TP+FP+TN+FN} CA = \frac{TP+TN}{TP+FP+TN+FN} \quad \text{equation (3)}$$

Precision: The proportion of SMSes classified as scam e-wallet deposit notifications that are correctly classified.

$$Precision = \frac{TP}{TP+FP} Precision = \frac{TP}{TP+FP} \quad \text{equation (4)}$$

Recall: The proportion of the actual scam e-wallet deposit notification SMSes that are correctly classified.

$$Recall = \frac{TP}{TP+FN} Recall = \frac{TP}{TP+FN} \quad \text{equation (5)}$$

F1-measure: The harmonic mean of precision and recall.

$$F1\ measure = \frac{2 \times Precision \times Recall}{Precision + Recall} F1\ measure = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad \text{equation (6)}$$

5. Comparative evaluation results

Table 4 presents the two classifier models' performances in terms of TP, TN, FP, FN, FPR and FNR evaluation metrics.

Table 4: Classification performance in terms of TP, TN, FP, FN, FPR and FNR

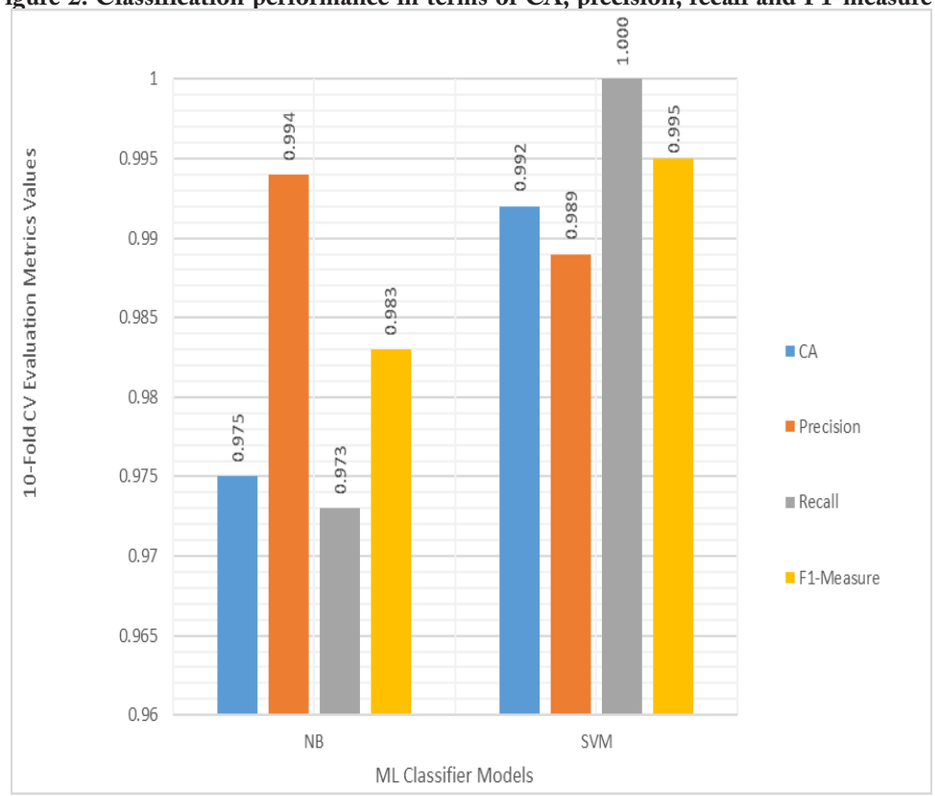
Metrics	10-fold cross-validation average results	
	NB	SVM
TP	17.700	18.200
TN	5.700	5.600
FP	0.100	0.200
FN	0.500	0.000
FPR	0.017	0.034
FNR	0.027	0.000

The results in Table 4 show that, on average, the SVM model correctly classified 18.2 scam e-wallet deposit notification SMSes (i.e., TP = 18.2) and misclassified 0.0 (i.e., FN = 0.0) compared to the NB's 17.7 and 0.5 respectively. This reveals that the SVM classifier model was more efficient than the NB model with respect to detecting scam e-wallet deposit notification SMSes. However, the results indicate the contrary about the models' capability to detect the ham SMSes, with the NB having correctly classified an average of 5.7 ham SMSes and misclassified 0.1, compared to the SVM's 5.6 and 0.2 respectively. The two models' FPR and FNR averages conform with the aforementioned results.

Because the emphasis of the study was on the models' capability to efficiently detect scam e-wallet deposit notification SMSes, we gave performance in terms of this criterion more weight than performance in respect of the detection of ham SMSes. Thus, the SVM classifier model was found to be superior to the NB model.

The graph in Figure 2 depicts the two classifier models' performance in terms of CA, precision, recall and F1-measure.

Figure 2: Classification performance in terms of CA, precision, recall and F1-measure



AJIC 25 - 5-July-20 - 20h14.indd 9-10 7/5/2020 8:22:47 PM

The results shown in Figure 2 indicate that the SVM model correctly classified more SMSes than the NB model, producing the highest CA. The two classifier models demonstrated contrasting performances in terms of precision and recall, with SVM achieving the highest recall while NB produced the highest precision. The differences between precision and recall, in terms of their respective equations (4) and (5) provided in the previous section, is represented by FP and FN. The two classifier models had contrasting FP and FN values due to their dissimilar performances with respect to correctly classifying the ham SMSes and the scam e-wallet deposit notification SMSes, as highlighted in the previous section. This caused the observed models' contrasting precision and recall values. The harmonic mean of precision and recall (i.e., F1-measure) helped to remove any ambiguity regarding which of the two models made a better overall classifier for both ham SMSes and scam e-wallet deposit notification SMSes, with the SVM model's F1-measure = 0.995 being superior to the NB model's F1-measure = 0.983.

6. Conclusions

In an attempt to contribute to solution of the problem of fraudulent e-wallet deposit notification SMSes, we trained NB and SVM classifier models to classify ham SMSes and scam e-wallet deposit notification SMSes, following which their performances were evaluated. The evaluation results indicate that the SVM classifier model can detect fraudulent e-wallet deposit notification SMSes more efficiently than the NB model. The SVM model's strength was highlighted by its FNR = 0.000, CA = 0.992, recall = 1.000 and F1-measure = 0.995, making it the more efficient model. Our envisaged future work is to extend this study by developing a mobile application or applications making use of the SVM classifier model to detect fraudulent e-wallet deposit notification SMSes on a user's mobile device.

References

Abdulhamid, S. M., Abd Latiff, M. S., Chiroma, H., Osho, O., Abdul-Salaam, G., Abubakar, A. I., & Herawan, T. (2017). A review on mobile SMS spam filtering techniques. *IEEE Access*, *5*, 15650–15666. https://doi.org/10.1109/ACCESS.2017.2666785

Ahmed, I., Guan, D., & Chung, T. C. (2014). SMS classification based on naïve Bayes classifier and apriori algorithm frequent itemset. *International Journal of Machine Learning and Computing*, *4*(2), 183–187. https://doi.org/10.7763/IJMLC.2014.V4.409

Akbari, F., & Sajedi, H. (2015). SMS spam detection using selected text features and boosting classifiers. In *2015 7th Conference on Information and Knowledge Technology (IKT)* (pp. 1–5). Institute of Electrical and Electronics Engineers (IEEE). https://doi.org/10.1109/IKT.2015.7288782

Almeida, T. A., Gómez, J. M., & Yamakami, A. (2011). Contributions to the study of SMS spam filtering: New collection and results. In *2011 ACM Symposium on Document Engineering* (pp. 259–262). Association for Computing Machinery (ACM). https://doi.org/10.1145/2034691.2034742

Arde, A. (2012, January 8). EFT scammers fake their proof of payment to dupe you, the seller. *IOL*. Retrieved from https://www.iol.co.za/personal-finance/eft-scammers-fake-their-proof-of-payment-to-dupe-you-the-seller-1209136

Choudhary, N., & Jain, A. K. (2017). Towards filtering of SMS spam messages using machine learning based technique. In D. Singh, B. Raman, A. K. Luhach, & P. Lingras (Eds.), *Advanced Informatics for Computing Research: First International Conference, ICAICR 2017* (pp. 18–30). https://doi.org/10.1007/978-981-10-5780-9_2

Christopher, N., & Kar, S. (April 24, 2018). Mobile wallet scam: As next-gen spenders go cashless, e-wallet scamsters too are getting creative. *The Economic Times*. Retrieved from https://tech.economictimes.indiatimes.com/news/internet/as-next-gen-spenders-go-cashless-e-wallet-scamsters-too-are-getting-creative/63889870

Cormack, G. V. (2008). Email spam filtering: A systematic review. *Foundations and Trends in Information Retrieval*, *1*(4), 335–455. https://doi.org/10.1561/1500000006

Crime in the City NAMIBIA (n.d.). [Facebook group]. Retrieved from https://www.facebook.com/groups/616819348371856/permalink/2462372400483199/

Erongo. (2016, December 1). Crooks out to empty wallets. Retrieved from http://www.erongo.com.na/news/crooks-out-to-empty-wallets/?

FNB Namibia Classic Clashes. (n.d.). [Facebook group]. Retrieved from https://www.facebook.com/FNBNamClassicClashes

Goel, D., & Jain, A. K. (2018). Mobile phishing attacks and defence mechanisms: State of art and open research challenges. *Computers & Security*, *73*, 519–544. https://doi.org/10.1016/J.COSE.2017.12.006

Govpage. (2017). Fake payment SMS & email confirmation. Retrieved from https://www.govpage.co.za/general-scams-warnings/fake-payment-sms-email-confirmation

Hedieh, S., Parast, G. Z., & Akbari, F. (2016). SMS spam filtering using machine learning techniques: A survey. *Machine Learning Research*, *1*(1), 1–14. https://doi.org/10.11648/j.mlr.20160101.11

IOL. (2019, March 25). More than R12.8bn sent from eWallet transactions in six months. Retrieved from https://www.iol.co.za/business-report/economy/more-than-r128bn-sent-from-ewallet-transactions-in-six-months-20071814

Jain, A. K., & Gupta, B. B. (2018). Rule-based framework for detection of smishing messages in mobile environment. *Procedia Computer Science*, *125*, 617–623. https://doi.org/10.1016/J.PROCS.2017.12.079

Mahmoud, T., & Mahfouz, A. (2012). SMS spam filtering technique based on artificial immune system. *International Journal of Computer Science Issues*, *9*(2), 589–597.

Nagel, E. (2015, June 5). Watch out for these common payment scams [Blog post]. *Gumtree*. Retrieved from https://blog.gumtree.co.za/watch-out-for-these-common-payment-scams/

Nagwani, N. K., & Sharaff, A. (2017). SMS spam filtering and thread identification using bi-level text classification and clustering techniques. *Journal of Information Science*, *43*(1), 75–87. https://doi.org/10.1177/0165551515616310

Nuruzzaman, M. T., Lee, C., & Choi, D. (2011). Independent and personal SMS spam filtering. In *2011 IEEE 11th International Conference on Computer and Information Technology* (pp. 429–435). Institute of Electrical and Electronics Engineers (IEEE). https://doi.org/10.1109/CIT.2011.23

AJIC 25 - 5-July-20 - 20h14.indd 11-12 7/5/2020 8:22:47 PM

Reaves, B., Blue, L., Tian, D., Traynor, P., & Butler, K. R. B. (2016). Detecting SMS spam in the age of legitimate bulk messaging. In *Proceedings of the 9th ACM Conference on Security & Privacy in Wireless and Mobile Networks - WiSec '16* (pp. 165–170). Association for Computing Machinery (ACM). https://doi.org/10.1145/2939918.2939937

Shahi, T. B., & Yadav, A. (2014). Mobile SMS spam filtering for Nepali text using naive Bayesian and support vector machine. *International Journal of Intelligence Science*, *04*(01), 24–28. https://doi.org/10.4236/ijis.2014.41004

Shaikh, A. A., & Karjaluoto, H. (2015). Mobile banking adoption: A literature review. *Telematics and Informatics*, *32*(1), 129–142. https://doi.org/10.1016/J.TELE.2014.05.003

Wagiet, R. (2012). SMS banking scam exposed. *Eyewitness News*. Retrieved from https://ewn.co.za/2012/08/15/Latest-SMS-banking-scam

A Supplementary Tool for Web-archiving Using Blockchain Technology

John E. de Villiers
Master's Student, Department of Computing Sciences, Nelson Mandela University, Port Elizabeth, South Africa
https://orcid.org/0000-0002-9146-8180

André P. Calitz
Professor, Department of Computing Sciences, Nelson Mandela University, Port Elizabeth, South Africa
https://orcid.org/0000-0002-2555-9041

Abstract
The usefulness of a uniform resource locator (URL) on the World Wide Web is reliant on the resource being hosted at the same URL in perpetuity. When URLs are altered or removed, this results in the resource, such as an image or document, being inaccessible. While web-archiving projects seek to prevent such a loss of online resources, providing complete backups of the web remains a formidable challenge. This article outlines the initial development and testing of a decentralised application (DApp), provisionally named Repudiation Chain, as a potential tool to help address these challenges presented by shifting URLs and uncertain web-archiving. Repudiation Chain seeks to make use of a blockchain smart contract mechanism in order to allow individual users to contribute to web-archiving. Repudiation Chain aims to offer unalterable assurance that a specific file and its URL existed at a given point in time—by generating a compact, non-reversible representation of the file at the time of its non-repudiation. If widely adopted, such a tool could contribute to decentralisation and democratisation of web-archiving.

Keywords
web-archiving, blockchain, trusted timestamping, non-repudiation, cryptographic hash functions, Merkle trees, DApps, smart contracts

DOI: https://doi.org/10.23962/10539/29194

Recommended citation
De Villiers, J. E., & Calitz, A. (2020). A supplementary tool for web-archiving using blockchain technology. *The African Journal of Information and Communication (AJIC), 25*, 1-14. https://doi.org/10.23962/10539/29194

This article is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) licence: https://creativecommons.org/licenses/by/4.0

AJIC 25 - 5-July-20 - 20h14.indd 13-1 7/5/2020 8:22:47 PM

1. Introduction

The ubiquity of the internet-enabled World Wide Web has allowed for its use as a global storage platform for data and information (Berners-Lee & Fischetti, 2001). The preservation of web content is of immeasurable value, and tools and services that can aid in this preservation need to be continually researched and explored. Web resources are linked to uniform resource locators (URLs), but the permanence of URL citations is uncertain and dependent on several factors, including websites' management policies and their archiving and hosting agreements. When a URL changes or is deleted, the associated resource, such as a web page, a document, an image, a photo or a video, will become either difficult to access or completely inaccessible.

Web-archiving initiatives seek to prevent losses of online resources due to URL changes or deletions, thus playing an important role in the preservation of data and information for future researchers and the broader public. However, it is misguided to assume the completeness and perpetual existence of web archives, such as the Wayback Machine (Habibzadeh, 2013; Internet Archive, n.d.; Murphy, Hashim, & O'Connor, 2008). Challenges facing web archives include the apparent lack of a complete, unbiased web archive across the entire web, and the archives' obligations to remove copyright-infringing content.

The purpose of our study was to perform the initial development and testing of a decentralised application (DApp) that could be made publicly available as a non-repudiation tool for use by individuals and entities seeking to participate in web-archiving. The DApp that we began development of, which we have named Repudiation Chain, is designed to provide a means of storing pertinent metadata relating to any given URL, by using the Ethereum cryptocurrency's blockchain smart contract tool. Repudiation Chain seeks to provide non-repudiation properties, i.e., guarantees that, at a specific point in time, the integrity or contents of a file linked to a URL can be established as an immutable representation of that file's existence. Repudiation Chain seeks to provide proof of the origin and integrity of the data linked to a URL; to prevent participants from denying responsibility for actions they take, such as altering a file linked to a URL; and to preserve information on the entity requesting the immutable representation.

Repudiation Chain is designed in such a way that, upon request by an entity, such as an internet user, the tool is required to store, in a blockchain, a time-stamped, immutable, compact, and public representation of a file that has been given a URL. This stored representation of the file must be associated with the entity that originally made the request for storage, as well as the URL that points to the file location. Repudiation Chain must also be able to determine whether there is a difference between the file that has been stored in the blockchain and the file identified by the URL.

Repudiation Chain is thus designed to be a supplementary web-archiving tool that can be used in conjunction with a web-archive initiative. It is intended for use by those who need access to a free notary-like tool, and those who wish to track whether a specific file hosted online has changed over time. Repudiation Chain thus aims to supplement, and address some flaws inherent in, existing internet-archiving services, as described in the introduction. It also aims to help decentralise and democratise web-archiving, by allowing individual users and small entities to participate in, and increase the accuracy of, web-archiving.

2. Literature review

Web-archiving

Web-archiving can be divided into two categories:

- *micro-archiving*, performed on a small scale (including by individuals) for day-to-day purposes and aiming to preserve an object for further study; and
- *macro-archiving*, performed on a large scale by organisations with specialised technical expertise in the field for preservation purposes (e.g., of societal and cultural heritage) and to serve as primary sources of historic data and information.

Web-archiving services are provided by trusted third parties, and the services' basic operations are (ITU, 2000) storing of documents, and issuing of signed copies of stored documents, when requested by an authorised entity, with inclusion of the documents' dates of registration.

Distributed ledger systems and blockchains

A distributed ledger system is a database distributed across multiple hosts on a network. It is typically public and requires some consensus among contributors operating on the system so that it can synchronise, share, and replicate itself. One of the more popular types of distributed ledger system is a blockchain (UK Government Chief Scientific Adviser, 2016). A blockchain platform is an open, decentralised ledger that can record transactions between parties across a peer-to-peer network, without the need for a central certifying authority. Blockchain applications are increasingly being developed for services such as fund transfers, smart contracts, e-voting, and efficient supply chain management. Blockchains can now be found deployed in numerous sectors, including finance, healthcare, law, trade, and real estate.

A blockchain is a data structure consisting of a list of data records, called blocks, which are cryptographically and sequentially linked. A block contains and stores the following data (Conte de Leon, Stalick, Jillepalli, Haney, & Sheldon, 2017):

- *Block data:* the set of messages or transactions, analogous to a collection of records;

AJIC 25 - 5-July-20 - 20h14.indd 2-3 7/5/2020 8:22:47 PM

- *Chaining hash values:* the hash values of the previous block in the blockchain; and
- *Block hash*: a representation of a given block's content as well as its chaining hash.

The existence and authenticity of a block are reliant on the block's cryptographic signature. This signature is composed of a timestamp, the hash value of the given block, and the hash value of the preceding block—unless the block is the genesis block (the first block in a blockchain).

Blockchains provide certain strengths and advantages over other systems, including the following:

- they do not require a central authority and thus they move trust requirements to the underlying technology (Conte de Leon et al., 2017; Walch, 2017);
- in a publicly available distributed ledger, all participants can make transactions, thus providing transparency (UK Government Chief Scientific Adviser, 2016); and
- all blocks added are validated and distributed across the network to form an emergent blockchain (and not all participants will have the same blockchain, thus making the emergent blockchain the one that the majority of the network participants agree upon at any given point in time) (Conte de Leon et al., 2017).

Two notable weaknesses of blockchains are:

- the security and immutability of a blockchain rely on the strength of the cryptographic hash function used (Conte de Leon et al., 2017); and
- they currently lack sufficiently widespread adoption.

Trusted timestamping and non-repudiation

A digital trusted timestamp, analogous to a traditional physical timestamp, serves as proof that the timestamped data or information existed at the time indicated by the stamp. A trusted timestamp is one that cannot be tampered with and, accordingly, is one provided by a service provider whose validity is trusted by all involved parties (Truu, 2010; Adams, Pinkas, Cain, & Zuccherato, 2001). In trusted timestamping, the creation and modification times are stored. This requires the use of publicly available, trusted, timestamp management infrastructure in the form of a timestamping authority. Trusted timestamps are particularly important for cryptocurrency ledgers such as Bitcoin, as well as for anything where reliable proof is needed of data or information's existence in a particular form at a particular time (Nakamoto, 2008).

A topic closely related to trusted timestamping is non-repudiation. A general definition of a system that provides non-repudiation is a system where a user cannot deny actions or refuse accountability (Negus, 2020). In the context of a cryptographic

service, non-repudiation is a form of authentication that provides proof of the integrity and origin of a given file, with proof being unforgeable and verifiable by any third party at any time (McCullagh & Caelli, 2000).

In Repudiation Chain, the smart contract used can be extended to provide a non-repudiation service, but the primary property it seeks to provide, in the tool's current stage of development, is immutable proof that a given URL-linked file existed at a specific point in time in a specific form.

Cryptographic hash functions and Merkle trees

A hash function maps an input of arbitrary length to an output of fixed length (Goodrich & Tamassia, 2014). Hash functions have many functional properties, which have led to their widespread use as one of the underlying means of security for computing devices, online communication, and the internet (Merkle, 1989; Preneel, 2011). Many different hash functions have emerged, including the family of secure hash algorithms (SHAs), which are extensively used in blockchain implementations and other cryptographically sensitive applications. The contents of a typical blockchain can be altered only if enough computational power has reached consensus on the alteration, and the computational power required is heavily dependent on the strength of the hash function (the hash function's resilience to different types of attacks) used by the blockchain.

A Merkle tree is a data structure that is dependent on hash functions (Berman, Karpinski, & Nekrich, 2007). Merkle trees can be used to provide a representation of data, and to provide a secure means of verifying the contents of a data file (Nakamoto, 2008; Singhal, Dhameja, & Panda, 2018). Every leaf value in a Merkle tree can be identified with respect to a publicly known root and the authentication path of that leaf. A Merkle tree is built from its leaves to its root so that the root can be used as a representation of the entire data file. Repudiation Chain uses a Merkle tree to represent data files of varying length in a compact manner.

Ethereum DApps and smart contracts

Ethereum, which supports the Ether cryptocurrency, uses a blockchain, and it provides a platform for DApps and smart contracts (Buterin, 2014). DApps provide an interface between smart contracts and the Ethereum Virtual Machine (EVM) where smart contracts are executed. A smart contract is a self-executing, distributed code. Executing a smart contract accrues a "gas" cost, which represents the amount of computational work required. The price of gas is proportional to the priority that the EVM gives to executing a smart contract, based on block "miners" prioritising contracts that are seen as profitable. A miner is a person using computing technologies, such as the Antminer S9 mining hardware, to process blocks of transactions, such as

AJIC 25 - 5-July-20 - 20h14.indd 4-5 7/5/2020 8:22:47 PM

Bitcoin transactions, and add the blocks to the blockchain (Hertig, 2017). Table 1 provides approximate gas costs, in 2020, for some operations (Wood, 2020).

Table 1: Prices of some operations in an Ethereum smart contract

Operation	Gas
adding two numbers	3 gas
multiplying two numbers	5 gas
secure hash algorithm 3 (SHA3) operation	30 gas
cost per word of input for SHA3 operation	6 gas
storing one byte of non-zero transaction data	68 gas

3. Initial development and testing of Repudiation Chain

To conduct the initial development work on Repudiation Chain, we worked with the aforementioned EVM smart contract tool provided by Ethereum. The metadata components that needed to be stored were:

- an immutable representation of the URL-linked file;
- a timestamp that is separate from the one provided by Ethereum and is expensive to alter, thus providing additional protection against alteration;
- the URL in question; and
- an association to the entity making use of the service.

More trustworthy timestamping

Several different kinds of attacks exist for Ethereum smart contracts. One is based on an unmined block's timestamp. If a smart contract uses a block's timestamp, it is possible for a miner who holds some incentive to manipulate the given contract's functionality to tamper with the timestamp for a block that the miner mines. This vulnerability has been reduced to a few seconds, but it still presents a potential weakness (Atzei, Bartoletti, & Cimoli, 2017).

Accordingly, since Repudiation Chain relies on a trusted timestamp, it was determined that the native timestamp provided for each block by Ethereum was inadequate. An alternative means of providing a trusted timestamp, or of providing a fail-safe method to detect possible tampering, needed to be identified. To this end, we designed Algorithm 1. Algorithm 1 provides a simple means of identifying whether a list of timestamps deviates by more than some predetermined length of time in a list of timestamps. The acceptable time deviation can be altered to conform to the specifics of the required scenario. This determination of an acceptable time deviation would need to be based on a consideration of the amount of time it takes to get a timestamp, the size of the list of timestamps, whether getting a timestamp can be parallelised (i.e., via multiple timeservers operating in a near-simultaneous fashion), the amount of time it takes for a contract to be executed, the transaction fee, and the source of the timestamps.

Algorithm 1: Detect suspicious timestamp

```
function isSuspiciousTimestamp(ListofTimestamps, TimeDeviation) : returns Boolean
    Long averageTimestamp ← Mean(ListofTimestamps))
    for each timestamp in ListofTimestamps do
        if ((timestamp + TimeDeviation) > averageTimestamp) or
                ((timestamp – TimeDeviation) < averageTimestamp) then
                    return true
        end if
    end for
    return false
end function
```

If Algorithm 1 returns a value of false, then the average timestamp of the list of timestamps can be used in the smart contract. This provides an emergent, fail-safe timestamp. This fail-safe timestamp indicates that the participants external to the miner came to consensus regarding a time. Thus, this fail-safe can be used in conjunction with the block's timestamp, with any significant difference between the two indicating that either the miner or the provider of the list of timestamps tampered with the timestamps provided.

Algorithm 1 only provides a means of detecting whether a timestamp was tampered with. To further obscure the possibility or risk of tampering, the list of timestamps should be from multiple, geographically dispersed sources, i.e., to reduce the possibility of collusion between possible timestamp providers. The derived average timestamp should be presented to the smart contract only once it satisfies Algorithm 1. This average timestamp is presented to the smart contract only when consensus is reached on the client's side. This results in the miner having a diminished window in which to tamper, without being detected, with the block's timestamp or the underlying smart contract's functionality. This process leads to the improvement of a smart contract's usefulness as a means of providing trusted timestamping. Examples of possible sources for lists of timestamps include the NTP Pool Project and Google's Public NTP.

Ensuring secure hash algorithms and application

Repudiation Chain is designed to use hash functions to provided file identification and integrity—as the hash value of any file can be used to identify that file, and any changes made to that file will result in a change to the file's hash value (Mackall, 2006).

Smart contracts on Ethereum should not be used to find multiple hash values, as hashing is an expensive operation in Ethereum (Wood, 2014). An alternative to using a smart contract to find the hash value of a file is to find the hash value on the client's side (i.e., on the side of the entity using Repudiation Chain). But finding

AJIC 25 - 5-July-20 - 20h14.indd 6-7 7/5/2020 8:22:47 PM

the hash value of a large file can fail due to hardware limitations, as the entirety of a large file cannot "fit" in the memory of the machine when calculating the hash value. Algorithm 2 was designed to overcome this practical limitation by partitioning the given file into a list of *m*-partitions, iterating over the list, and calculating the hash value of the entire file in a linear manner. Algorithm 3 provides a generalisation of Algorithm 2 for application to a list.

Algorithm 2: Naïve file hash algorithm used by Repudiation Chain

```
function NaiveFileHash(File, HashFunction, BufferSize) : returns FileHash
    FileHash HashValue
    while (entire File not read) do
        curBuffer ← File.readNextBuffer(BufferSize)
        HashValue ← HashFunction(output + curBuffer)
    end while
    return HashValue
end function
```

Algorithm 3: Data structure devised from Algorithm 2 for direct comparison to Merkle trees

```
function HashLinkedChain(ListofItems) : returns HashValue
    HashValue current
    for each Items in ListofItems do
        current ← Hash(current + Items)
    end for
    return current
end function
```

We conducted an ANOVA statistical comparison between Algorithm 3 and Merkle trees (using a binary Merkle tree), whereby the null hypothesis was set as: *the mean performance of either algorithm is identical*. This comparison was based on the time required to generate a representation of files of varying lengths when using either algorithm. Algorithm 3 was found to perform significantly more quickly than the Merkle tree, at a p-value less than 0.00000001. (In lay terms, the p-value is the probability of obtaining the observed results when the associated null hypothesis is true, with accordingly smaller p-values being preferable (see Mendenhall & Sincich, 2012).) This difference in performance was likely due to Algorithm 3 calculating roughly half as many hash values as a Merkle tree for the same file. Algorithm 3 does not have the same authentication path properties as a Merkle tree, but due to Algorithm 3's superior performance for the desired use case, it is the preferred means of generating hash value representations of files. The sacrifice Algorithm 3 makes in order to achieve greater speed than a Merkle tree is a loss of functionality that was not needed by Repudiation Chain.

4. Comparison between Repudiation Chain and Bitcoin's Proof of Existence

To test Repudiation Chain, we compared its efficacy to that of Bitcoin's Proof of Existence non-repudiation service, which operates as a web application (Crosby, Nachiappan, Pattanayak, Verma, & Kalyanaraman, 2016; Proof of Existence, n.d.). Proof of Existence provides non-repudiation by storing the cryptographic hash of a file as part of a transaction on Bitcoin's blockchain, with each transaction having an associated timestamp (Crosby et al., 2016). Users of Proof of Existence upload the file they wish to non-repudiate and pay the associated fee of 0.00025 Bitcoins (BTC), which, at the time of the finalisation of this article in June 2020, had a value of USD2.34 (based on an exchange rate, calculated via CoinMarketCap, of BTC1=USD9,349.75 and ETC1=228.80 (CoinMarketCap, 2020)). Proof of Existence is a centralised web DApp, unlike Repudiation Chain, which is a smart contract that interacts with a client-side application. Unlike Proof of Existence, Repudiation Chain facilitates the transfer of trust to the client's hardware as opposed to a centralised third-party's hardware. The cost to use Proof of Existence is set at a flat flee, and this fee is larger than the cost to store the appropriate information, with this disparity presumably being necessary to pay for infrastructure costs and to maintain a profit. Repudiation Chain does not share the same cost burdens and possible profit motivations.

In order to compare this USD2.34 cost to use Proof of Existence with the cost to use Repudiation Chain, we derived a function that could be used to estimate the gas cost of Repudiation Chain use, as follows:

> For an adjusted R^2 value of 0.9997806 for a p-level less than 0.00001, the gas cost to use the RC is given by: $(695.81*x_1) + 210598.31$; where x_1 is the length of the URL. The lower bound of a URL's length is 13 characters, the mean is 35 characters and the upper bound is 425 (Ducut, Liu, & Fontelo, 2008).

Table 2 provides a summary of the expected costs associated with using Repudiation Chain.

Table 2: Repudiation Chain gas cost estimates

URL length	Gas cost	Gas price (Wei)	Transaction fee (Ether)	USD value (June 2020)
13	219643.84	4.92e+15	4.92e-03	USD1.13
35	234951.66	5.26e+15	5.26e-03	USD1.20
425	506317.56	1.13e+16	1.13e-02	USD2.59

AJIC 25 - 5-July-20 - 20h14.indd 8-9 7/5/2020 8:22:48 PM

Table 3 compares Repudiation Chain and Proof of Existence in terms of cost, how files are received, information that is stored, and anonymity.

Table 3: Comparison between Repudiation Chain and Proof of Existence

Factors	Repudiation Chain	Proof of Existence
Cost of service	Dependent on length of URL (see Table 2)	Flat fee of BTC0.00025
Approximate cost to add non-repudiated representation of a file	USD1.13 – USD2.59	USD1.97
How files are received	Download	Upload
Information stored	two distinct and independent timestamps hash of file hash of cargo address of the adder URL of file	timestamp hash of file
Anonymity	Derived from Ethereum's and device's anonymity	Derived from Bitcoin's, browser's and device's anonymity

As seen in Table 3, our proposed Repudiation Chain was found to provide more functionality and information storage than Proof of Existence, and to be cheaper to use for many URL lengths.

5. Directions for further testing and development of Repudiation Chain

The next key step will be extensive testing of the functionality and scalability of Repudiation Chain and its underlying system on the main Ethereum network. Additionally, there are several possible additional elements of development. Repudiation Chain could utilise available processing cores more effectively through the parallelisation of Algorithms 1 and 2 for timestamping and hash representation, potentially resulting in shorter allowed time deviations. Repudiation Chain could also be redesigned to incorporate a "token" or alt-currency system, which would allow for donation of funds (through a publicly accessible wallet) in support of Repudiation Chain's functioning, thus extending its democratisation. Such functionality would, however, likely require investigation of cost-effective spam-filtering.

The functionality of Repudiation Chain could also be extended to provide additional front-end options for users, in the form of a web application or an application for smartphone devices (e.g., DApp, which would be very similar to a standard mobile application, but with some important computations and storage functionality handled by the Ethereum blockchain). Moreover, additional layers of anonymity could also be explored, in order to further protect the interests and privacy of users. This might include devising a means of "laundering" transactions between Ethereum smart contracts and Repudiation Chain, i.e., by obfuscating the origins of a given transaction or interaction. Consideration could also be given to the addition of a means to manipulate the speed at which information is stored by Repudiation Chain—e.g., by allowing users to define the gas price relative to the current market, thus improving the user experience.

References

Adams, C., Pinkas, D., Cain, P., & Zuccherato, R. J. (2001). *Internet X.509 public key infrastructure time-stamp protocol (TSP)*. The Internet Society. https://doi.org/10.17487/rfc3161

Ainsworth, S. G., AlSum, A., SalahEldeen, H., Weigle, M. C., & Nelson, M. L. (2011). How much of the web is archived? In *Proceedings of the 11th Annual International ACM/IEEE Joint Conference on Digital Libraries* (pp. 133–136). Association for Computing Machinery (ACM). https://doi.org/10.1145/1998076.1998100

AlNoamany, Y., AlSum, A., Weigle, M. C., & Nelson, M. L. (2014). Who and what links to the internet archive. *International Journal on Digital Libraries, 14*, 101–115. https://doi.org/10.1007/s00799-014-0111-5

Atzei, N., Bartoletti, M., & Cimoli, T. (2017). A survey of attacks on Ethereum smart contracts. In M. Maffei & M. Ryan (Eds.), *Principles of Security and Trust: 6th International Conference (POST 2017)* (pp. 164–186). https://doi.org/10.1007/978-3-662-54455-6_8

Bayardo, R. J., & Sorensen, J. (2005). Merkle tree authentication of HTTP responses. In *WWW '05: Special interest tracks and posters of the 14th International Conference on World Wide Web* (pp. 1182–1183). Association for Computing Machinery (ACM). https://doi.org/10.1145/1062745.1062929

Becker, G. (2008). *Merkle signature schemes, Merkle trees and their cryp*tanalysis. Ruhr-Universität Bochum, Germany. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.392.7879&rep=rep1&type=pdf

Berman, P., Karpinski, M., & Nekrich, Y. (2007). Optimal trade-off for Merkle tree traversal. *Theoretical Computer Science, 372*, 26–36. https://doi.org/10.1016/j.tcs.2006.11.029

Berners-Lee, T., & Fischetti, M. (2001). *Weaving the Web: The original design and ultimate destiny of the World Wide Web.* New York: HarperCollins. https://doi.org/10.5860/choice.37-3934

Buterin, V. (2014). Ethereum white paper: A next generation smart contract & decentralized application platform. Retrieved from https://cryptorating.eu/whitepapers/Ethereum/Ethereum_white_paper.pdf

Chohan, U. W. (2017). Cryptocurrencies: A brief thematic review. SSRN. https://dx.doi.org/10.2139/ssrn.3024330

AJIC 25 - 5-July-20 - 20h14.indd 10-11 7/5/2020 8:22:48 PM

CoinMarketCap. (2020). All cryptocurrencies. Retrieved 25 June 2020 from https://coinmarketcap.com/all/views/all/

Conte de Leon, D., Stalick, A. Q., Jillepalli, A. A., Haney, M. A., & Sheldon, F. T. (2017). Blockchain: Properties and misconceptions. *Asia Pacific Journal of Innovation and Entrepreneurship*, *11*(3) 286–300. https://doi.org/10.1108/apjie-12-2017-034

Crosby, M., Nachiappan, Pattanayak, P., Verma, S., & Kalyanaraman, V. (2016). Blockchain technology: Beyond Bitcoin. *Applied Innovation Review*, *2*, 6–19.

Ducut, E., Liu, F., & Fontelo, P. (2008). An update on uniform resource locator (URL) decay in MEDLINE abstracts and measures for its mitigation. *BMC Medical Informatics and Decision Making*, *8*, 23. https://doi.org/10.1186/1472-6947-8-23

Duranti, L. (1992). Origin and development of the concept of archival description. *Archivaria 35*, 47–54.

Dworkin, M. J. (2015). *SHA-3 standard: Permutation-based hash and extendable output functions.* Federal Information Process Standards Publication 202, NIST. https://doi.org/10.6028/nist.fips.202

Etherscan. (2020). Ethereum average gas price chart. Retrieved 25 June 2020 from https://etherscan.io/chart/gasprice

Everett, S., Calitz, A. P., & Greyling, J. (2017). The case for a 'sovereign' distributed securities depository for securities settlement. *Journal of Securities Operations & Custody*, *9*(3), 269–292.

Goodrich, M. T., & Tamassia, R. (2014). *Algorithm design and applications.* Hoboken, NJ: John Wiley & Sons.

Haber, S., & Stornetta, W. S. (1991). How to time-stamp a digital document. *Journal of Cryptology*, *3*, 99–111. https://doi.org/10.1007/bf00196791

Habibzadeh, P. (2013). Decay of references to web sites in articles published in general medical journals: Mainstream vs small journals. *Applied Clinical Informatics*, *4*(4), 455–464. https://doi.org/10.4338/aci-2013-07-ra-0055

Hertig, A. (2017). Ethereum 101: Chapter 5: How do Ethereum smart contracts work? CoinDesk. Retrieved from https://www.coindesk.com/learn/ethereum-101/ethereum-smart-contracts-work

Hevner, A. R. (2007). A three cycle view of design science research. *Scandinavian Journal of Information Systems, 19*(2), 4.

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *Management Information Systems Quarterly*, *28*(1), 75–105. https://doi.org/10.2307/25148625

Internet Archive. (n.d.). [Website]. Retrieved from https://archive.org/

International Telecommunication Union (ITU). (2000). X. 842: Information technology: Security techniques: Guidelines for the use and management of trusted third party services. ITU-T Study Group 7, Telecommunication Standardisation Sector. https://doi.org/10.3403/02617626u

Masanés, J. (2006). Web archiving: Issues and methods. In *Web archiving* (pp. 1–53). Heidelberg: Springer. https://doi.org/10.1007/978-3-540-46332-0_1

Mackall, M. (2006). Towards a better SCM: Revlog and mercurial. In *Ottawa Linux Symposium 2* (pp. 83–90).

Manuel, S., & Peyrin, T. (2008). Collisions on SHA-0 in one hour. In Nyberg K. (Ed.), *Fast software encryption: FSE 2008* (pp 16-35). Lecture Notes in Computer Science, Vol. 5086. Berlin: Springer. https://doi.org/10.1007/978-3-540-71039-4_2

McCullagh, A., & Caelli, W. (2000). Non-repudiation in the digital environment. *First Monday, 5*(8). https://doi.org/10.5210/fm.v5i8.778

Mendenhall, W., & Sincich, T. (2012). *A second course in statistics: Regression analysis* (7th ed.). Saddle River, NJ: Prentice Hall.

Merkle, R. C. (1989). A certified digital signature. In G. Brassard (Ed.), *Advances in cryptology — CRYPTO' 89 proceedings* (pp. 218–238). Lecture Notes in Computer Science, Vol. 435. New York: Springer.

Murphy, J., Hashim, N. H., & O'Connor, P. (2008). Take me back: Validating the Wayback Machine. *Journal of Computer-Mediated Communication*, *13*(1), 60–75. https://doi.org/10.1111/j.1083-6101.2007.00386.x

Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system. Retrieved from https://bitcoin.org/bitcoin.pdf

Naor, M., & Yung, M. (1989). Universal one-way hash functions and their cryptographic applications. In *STOC '89: Proceedings of the Twenty-first Annual ACM Symposium on Theory of Computing* (pp. 33–43). Association for Computing Machinery (ACM). https://doi.org/10.1145/73007.73011

Negus, C. (2020). *Linux bible* (10th ed.). Indianapolis: John Wiley & Sons.

Niels, B. (2011). Web archiving – Between past, present, and future. In C. Ess, & M. Consalvo (Eds.), *The handbook of internet studies* (pp. 24–42). Chichester, UK: John Wiley & Sons. https://doi.org/10.1002/9781444314861.ch2

Offermann, P., Levina, O., Schönherr, M., & Bub, U. (2009). Outline of a design science research process. In *Proceedings of the 4th International Conference on Design Science Research in Information Systems and Technology.* Association for Computing Machinery (ACM). https://doi.org/10.1145/1555619.1555629

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology. *Journal of Management Information Systems*, *24*(3), 45–77. https://doi.org/10.2753/mis0742-1222240302

Preneel, B. (2011). Hash functions. In H. C. van Tilborg, & S. Jajodia (Eds.), *Encyclopedia of cryptography and security* (pp. 543–545). Berlin: Springer Science & Business Media.

Proof of Existence. (n.d.). [Website]. Retrieved from https://proofofexistence.com

Seo, J.-W., Kim, D.-K., Kim, H.-C., & Chung, J.-W. (2007). The algorithm of sharing incomplete data in decentralized P2P. *International Journal of Computer Science and Network Security*, *7*(8), 149–153.

Singhal, B., Dhameja, G., & Panda, P. S. (2018). *Beginning blockchain: A beginner's guide to building blockchain solutions.* Berlin: Apress.

Suzuki, K., Tonien, D., Kurosawa, K., & Toyota, K. (2006). Birthday paradox for multi-collisions. In M.S. Rhee, & B. Lee (Eds.), *Information security and cryptology – ICISC 2006* (pp. 29–40). Lecture Notes in Computer Science, Vol. 4296. Berlin: Springer. https://doi.org/10.1007/11927587_5

Thomsen, S. S., & Knudsen, L. R. (2009). *Cryptographic hash functions.* Technical University of Denmark, Kongens Lyngby.

Truu, A. (2010). *Standards for hash-linking based time-stamping schemes.* University of Tartu, Estonia.

Tsudik, G. (1992). Message authentication with one-way hash functions. In *IEEE INFOCOM '92: The Conference on Computer Communications* (pp. 2055–2059). Institute of Electrical and Electronics Engineers (IEEE). https://doi.org/10.1109/infcom.1992.263477

AJIC 25 - 5-July-20 - 20h14.indd 12-13 7/5/2020 8:22:48 PM

UK Government Chief Scientific Adviser. (2016). *Distributed ledger technology: Beyond block chain.* London: Government Office for Science.
Vega-Redondo, F. (2003). *Economics and the theory of games.* Cambridge, UK: Cambridge University Press.
Walch, A. (2017). Open-source operational risk: Should public blockchains serve as financial market infrastructures? In D. L. Chuen, & R. H. Deng, (Eds.), *Handbook of blockchain, digital finance, and inclusion, volume 2* (pp. 243–269). Cambridge, MA: Academic Press. https://doi.org/10.1016/b978-0-12-812282-2.00011-5
Waldrop, M. M. (2016). The chips are down for Moore's law. *Nature*, *530*(7589), 144-147.
Wall, L. D. (2018). *Some blockchain challenges.* Atlanta: Federal Reserve Bank of Atlanta.
Wood, G. (2014). *Ethereum: A secure decentralised generalised transaction ledger.* Retrieved from https://gavwood.com/paper.pdf
Wood, G. (2020). *Ethereum: A secure decentralised generalised transaction ledger: Petersburg version.* Retrieved from https://ethereum.github.io/yellowpaper/paper.pdf

Factors Impacting Tanzanian Rural and Peri-urban Drug Dispensaries' Perceived Benefits from Using an mHealth Reporting System

Alistair Elias
Master's Student, Department of Computer Science and Engineering, University of Dar es Salaam
https://orcid.org/0000-0001-9852-3405

Joel S. Mtebe
Senior Lecturer, Department of Computer Science and Engineering, University of Dar es Salaam
https://orcid.org/0000-0003-2760-7673

Abstract
This study examines the factors impacting Tanzanian drug dispensaries' perceptions of the potential benefits of using a mobile health (mHealth) reporting system. Since 2003, the Government of Tanzania has implemented its Accredited Drug Dispensing Outlet (ADDO) programme in support of dispensaries in under-served rural and peri-urban areas. A core element of this ADDO programme is the dispensary use of an mHealth application for reporting on the drugs they are dispensing and the medical conditions that are being treated. This study canvassed the views of ADDO programme dispensaries using the mHealth reporting system in the Iringa, Mbeya, and Njombe Regions, through a survey questionnaire completed by 318 dispensaries and focus group discussions with 38 dispensary personnel. The data revealed that four factors—system quality, information quality, service quality, and price value—are all having a statistically significant impact on dispensaries' perceptions of the potential benefits of using the mHealth system.

Keywords
mobile health (mHealth), online reporting, drug dispensaries, drug dispensing outlets, information systems, perceived benefits, updated D&M IS success model, extended unified theory of acceptance and use of technology (UTAUT2), system quality, information quality, service quality, facilitating conditions, price value, Accredited Drug Dispensing Outlet (ADDO) programme, Tanzania, Iringa, Mbeya, Njombe

Acknowledgement
The authors acknowledge with gratitude the statistical advice provided by an anonymous reviewer.

DOI: https://doi.org/10.23962/10539/29193

AJIC 25 - 5-July-20 - 20h14.indd 14-1 7/5/2020 8:22:48 PM

Recommended citation
Elias, A., & Mtebe, J. S. (2020). Factors impacting Tanzanian rural and peri-urban drug dispensaries' perceived benefits from using an mHealth reporting system. *The African Journal of Information and Communication (AJIC), 25*, 1-22.
https://doi.org/10.23962/10539/29193



This article is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) licence:
https://creativecommons.org/licenses/by/4.0

1. Introduction

The use of mobile phone technology has rapidly expanded in developing countries due to decreases in the cost of connectivity and increases in network coverage (Chib et al., 2014). Researchers and governments have been taking advantage of these developments in the search for innovative ways to support healthcare provision (Aranda-Jan et al., 2014). Consequently, many mobile health (mHealth) projects have been implemented, which are aimed at: improving delivery processes, supply chain management of essential medical commodities, assisting policymakers in planning and implementing different programmes, providing remote access to healthcare facilities, facilitating the training of health workers, and enabling remote monitoring and surveillance (Chib et al., 2014; Leon, Schneider, & Daviaud, 2012).

In Tanzania, the CommCare mHealth system helped home-based care providers to provide social support for patients with clinical symptoms associated with adverse drug events in Dar es Salaam (Bogan et al., 2009). The system was found to have improved the timeliness of home-based care visits by 86%, thereby enhancing providers' performance (DeRenzi et al., 2012).

Similarly, the SMS for Life reporting system managed to prevent stock-outs of essential malaria drugs in nearly 5,000 Tanzanian health facilities (Mikkelsen-Lopez et al., 2014). Users sent structured text messages to district-level supervisors at the end of each week to report their facilities' stock levels of anti-malaria drugs. Using this system, on average, only 29% of health facilities were completely stocked out of the first-line anti-malarials over a period of 15 months. Other examples of mHealth systems adopted in Tanzania include the systems developed for reporting of malaria in north-eastern Tanzania (Francis et al., 2017), for surveillance of malaria in Zanzibar (Barclay, Smith, & Findeis, 2012), and for reducing maternal mortality in rural areas (Mbaruku et al., 2018).

More recently, the Ministry of Health and Social Welfare, through the Tanzania Food and Drugs Authority (TFDA) and a technical partner, Management Sciences for Health, implemented an mHealth reporting system to enable Accredited Drug Dispensing Outlet (ADDO) dispensaries to transmit information, on

drugs dispensed and conditions treated, to the Ministry and the TFDA (Rutta et al., 2015). The ADDO programme aims to improve access to affordable, quality pharmaceutical services and drugs in rural areas with few or no registered pharmacies. This has been done by training retail drug store dispensary personnel to provide some prescription drugs (Embrey et al., 2016). However, it was found that many dispensaries were keeping data manually. Thus the data could not be easily reported to the regulators and the government authorities at the district, regional, and national Ministry level (Dillip et al., 2017). It was, therefore, difficult for the Ministry to assess ADDO dispensaries' performance and compliance with existing regulations. To address this problem, the mHealth system was implemented and integrated into the District Health Information Software 2 (DHIS2) platform to track and monitor data from dispensaries from various regions in Tanzania. Using this system, data on the movement of drugs that are being dispensed can easily be monitored, since dispensaries can send reports easily using their mobile devices. The system has been implemented in the Iringa, Mbeya, and Njombe Regions.

The study described and discussed in this article: (1) investigated the factors that impact dispensaries' perceptions of the potential benefits of using the ADDO programme's mHealth reporting system; and (2) generated recommendations for how the Tanzanian Government can improve dispensaries' experiences of using the system. The use factors explored by the research were drawn from the updated D&M IS success model (DeLone & McLean, 2003), and the extended unified theory of acceptance and use of technology (UTAUT2) (Venkatesh et al., 2003).

2. Literature review

Tanzania's Accredited Drug Dispensing Outlet (ADDO) programme

Traditionally in Tanzania, prescription drugs could be sold only by an authorised pharmacy, with a non-authorised retail drug store (known in Kiswahili as *duka la dawa baridi (DLDB)*) permitted to sell only non-prescription drugs. An increase in the number of health care seekers and private clinics (Chalker et al., 2015) triggered an increase in drug stores that illegally sold prescription drugs, especially in rural and peri-urban areas not served by authorised pharmacies. An estimated 10.5% of the lowest-income quartile of the country, representing 75% of Tanzanians, live in rural and peri-urban areas (Rutta et al., 2015).

To alleviate the problem, the Tanzania Food and Drug Authority (TFDA) launched a public-private programme called the Accredited Drug Dispensing Outlet (ADDO) programme, which aimed to train non-authorised retail drug store dispensaries to provide some prescription drugs (Rutta et al., 2009). Those stores would then acquire ADDO status (known in Kiswahili as *duka la dawa muhimu (DLDM)*). By 2016, there were over 9,000 shops across the country licensed as ADDOs by the TFDA (Embrey et al., 2016).

AJIC 25 - 5-July-20 - 20h14.indd 2-3 7/5/2020 8:22:48 PM

Given the importance of ADDOs in Tanzania, the lack of reliable records on medicines and on the conditions for which the medicines were purchased became a concern to the Ministry (Rutta, 2014). It was difficult for supervisors and inspectors from the Ministry and the TFDA to assess ADDO dispensaries' performance and the shops' compliance with regulations, as most of the records were stored manually.

In 2018, the Ministry, through the TFDA, introduced the mHealth system under the Febrile Illness Management Project, which was established by the Ministry in conjunction with the National Malaria Control Programme and the Maternal, Newborn and Child Health Unit. The project aims to improve the diagnostic services and treatment for three febrile illnesses: malaria, diarrhea, and pneumonia. The mHealth system was established to monitor the project through the collection of information from the ADDOs. The project was launched in the Morogoro Region and is now also being implemented in the Mbeya, Iringa and Njombe Regions.

The system transmits data to a District Health Information Software version 2 (DHIS 2) platform for subsequent analysis and reporting. This system was acquired from the Integrated Disease Surveillance and Response (IDSR) system employed by the President's Malaria Initiative (PMI). This USAID-funded project aims to reduce malaria-related mortalities (PMI, 2017).

Nearly 1,000 ADDO shops have been registered for, and are using, the mHealth system in the Iringa, Mbeya, and Njombe regions. Since the system's introduction, these dispensaries have been recording information and sending it to the Ministry and the TFDA via their mobile phones. The system received positive reception during adoption, with the dispensaries initially using the system intensively, but there is anecdotal evidence to suggest that use has gradually diminished due to challenges faced by the ADDOs.

mHealth information systems use

The adoption and use of mHealth systems is becoming common in developing countries due to the decreasing cost and increasing network coverage of mobile services (Chib et al., 2014). However, studies have indicated that these systems often do not bring about their expected benefits (Aranda-Jan et al., 2014; Chang et al., 2011; Holden & Karsh, 2010) as users tend to stop using them or to use relatively few features, beyond the period of intensive use after initial adoption (Kim, Lee, Hwang, & Yoo, 2016). For instance, users started underreporting after some years of operation of the mHealth system implemented in Kenya (Githinji et al., 2013), and Uganda (Chang et al., 2011). This underutilisation of mHealth systems has been described as a partial failure (DeLone & McLean, 2016; Heeks, 2006).

While initial acceptance is an important step towards realising mHealth system success, long-term benefits depend on continued use (Bhattacherjee, 2001). Some studies have found that users tend to stop using the system or use relatively few features after a few years of use (Kim et al., 2016). As a result, achieving the intended long-term benefits of such systems is unlikely (Vaghefi & Tulu, 2019). The success of information systems depends on users' continued use of the systems beyond the adoption stages (Karahanna, Straub, & Chervany, 1999).

There has not been extensive research into how to ensure long-term benefits from mHealth system implementation (with Leon et al. (2012) being a notable exception). Aranda-Jan et al. (2014) argue that existing studies of mHealth systems in developing countries tend to focus on initial adoption and implementation, with little attention being paid to an evaluation of the long-term use of systems.

It is important to ensure that these systems are used beyond the adoption stages to reduce the chances of systems' failures (DeLone & McLean, 2016). In addition, users must be able to use the majority of a system's features (Burton-Jones & Volkoff, 2017), and it should be used habitually (Limayem, Hirt, & Cheung, 2007). Currently, studies tend to focus on specific aspects of a system, project, or organisation during an evaluation of information system success, with little attention being paid to how the system is being used (DeLone & McLean, 2016). As a result, the long-term benefits of mHealth systems adopted in developing countries remain uncertain (Aranda-Jan et al., 2014).

Conceptual models for IS success and technology acceptance

The two models deployed in the research were:

- the updated D&M IS success model (DeLone & McLean, 2013); and
- the extended unified theory of acceptance and use of technology (UTAUT2) (Venkatesh, Thong, & Xu, 2012).

The updated D&M IS success model was extended from the original D&M model (DeLone & McLean, 1992). The original D&M model consists of six factors: system quality, information quality, system use, user satisfaction, individual impact, and organisational impact. The updated D&M model replaced individual impact and organisational impact with the net benefits, while service quality was added as a new factor (DeLone & McLean, 2003).

The extended unified theory of acceptance and use of technology (UTAUT2) extends the original UTAUT order to study acceptance and use of technology in a consumer context (Venkatesh et al., 2012). The original UTAUT model has four key factors: performance expectancy, effort expectancy, social influence, and facilitating conditions that influence behavioural intention to use a technology and/or technology use. The UTAUT2 incorporates three factors into UTAUT: hedonic motivation, price value, and habit. Individual differences—namely, age, gender, and experience—are hypothesised to moderate the effects of these constructs on behavioural intention

AJIC 25 - 5-July-20 - 20h14.indd 4-5 7/5/2020 8:22:48 PM

and technology use. Therefore, the UTAUT2 consists of seven factors, namely, effort expectancy, performance expectancy, facilitation conditions, social influence, price value, hedonic motivation, and habit.

These two models—the updated D&M IS success model and UTAUT2—have been widely used to understand the reasons for low or non-use of information systems in various contexts.

3. Research design

Research model

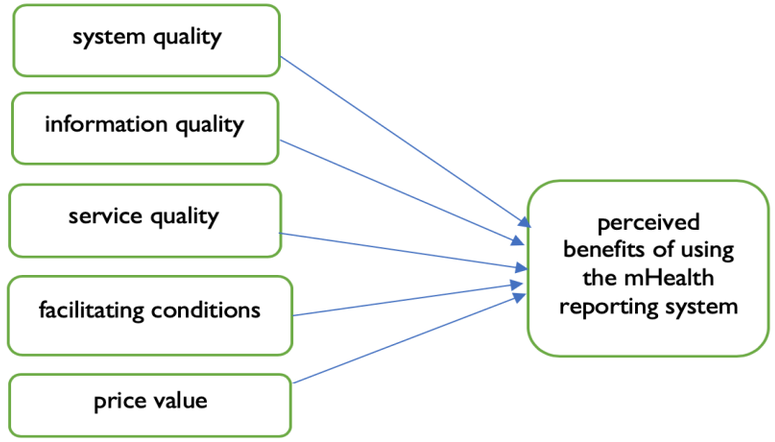
To explore the degree to which certain use factors may influence dispensaries' perceptions of the potential benefits of using the ADDO programme's mHealth system, relevant factors were extracted from the updated D&M IS success model (DeLone & McLean, 2013), and the extended unified theory of acceptance and use of technology (UTAUT2) model (Venkatesh et al., 2012) was adapted. Table 1 shows the main factors from two selected research models, and the five use factors that were found to be most relevant to the research's focus on perceived benefits of use (and thus adopted for the research model).

Table 1: Perceived-benefits-of-use factors adopted for the research model

2 models from the literature	Main factors identified in the 2 models in the literature	5 perceived-benefits-of-use factors adopted for the research model
updated D&M IS success model (DeLone & McLean, 2013)	• system quality • information quality • service quality	1. system quality 2. information quality 3. service quality
extended unified theory of acceptance and use of technology (UTAUT2) model (Venkatesh et al., 2012)	• performance expectancy • effort expectancy • social influence • facilitating conditions • hedonic motivation • price value • habit • behavioural intention • use	4. facilitating conditions 5. price value

Figure 1 provides a visualisation of the research model, with the five use factors chosen for the model—system quality, information quality, service quality, facilitating conditions, and price value—all contributing to perceived benefits of the use of an mHealth system.

Figure 1: Research model for investigating impacts on dispensaries' perceptions of potential benefits of using ADDO mHealth reporting system



Research hypotheses

Hypotheses were developed for each of the five perceived-benefits-of-use factors in the research model.

System quality

System quality is concerned with whether or not there are "bugs" in the system, the consistency of the user interface, and the ease of use of the system (Seddon & Kiew, 1995). The quality of an information system has been shown to have an impact on users' perceptions of it, as well as on users' satisfaction (DeLone & McLean, 2013). Therefore, it was important to include system quality as a factor that will influence the ADDO programme dispensaries' perceptions of the potential benefits of using the mHealth system. The hypothesis for this perceived-benefits-of-use factor was:

> **H1:** System quality has an impact on ADDO programme dispensaries' perception of the potential benefits of using the mHealth information system.

AJIC 25 - 5-July-20 - 20h14.indd 6-7 7/5/2020 8:22:48 PM

Information quality

Moreover, the primary motivation of information systems is to provide users with accurate, timely, and relevant information (DeLone & McLean, 2016). Therefore, information quality was considered an important factor in users' perception of the benefits of using the ADDO mHealth system. Information quality is the information system's output, which is determined by the quality of data and reports generated by the system (DeLone & McLean, 2016). Dispensaries are likely to favourably perceive using the mHealth system if they find that the system generates quality data and reports that meet their needs. The hypothesis for this perceived-benefits-of-use factor was:

> **H2:** Information quality has an impact on ADDO programme dispensaries' perception of the potential benefits of using the mHealth information system.

Service quality

According to DeLone and Mclean (2003), system support service is the measure of the quality of support services that users receive from the IS department and IT support personnel. It is expected that users with good system support are likely to continue using the system (DeLone & McLean, 2016). To ensure smooth use of the system, the Ministry of Health has a dedicated IT Unit that provides support to users of the mHealth system. It was important to measure the quality of support services offered by this IT Unit. The hypothesis for this perceived-benefits-of-use factor was:

> **H3:** Service quality has an impact on ADDO programme dispensaries' perception of the potential benefits of using the mHealth information system.

Facilitating conditions

In the ADDO mHealth context, key facilitating conditions include access to mobile telephony, internet access, and related resources. Facilitating conditions were found to be key determinants of users' continued use of various systems such as those researched by Ng et al. (2015) and Smith and Motley (2010). The hypothesis for this perceived-benefits-of-use factor was:

> **H4:** Facilitating conditions have an impact on ADDO programme dispensaries' perception of the potential benefits of using the mHealth information system.

Price value

The price value is considered positive when using the system results in net cost benefits (Venkatesh et al., 2012). When the system was introduced, dispensaries were initially provided with free internet data bundles to facilitate training in the system and to get the project started. Since that initial period, there has been an expectation that dispensaries will use their resources to pay for the internet data needed to continue using the system, i.e., that dispensaries will be willing to bear the cost of using this system because they perceive that the system has direct benefits in their day-to-day activities. The price value was found to be a predictor in the use of the aforementioned SMS for Life reporting app in Tanzania (Mikkelsen-Lopez et al., 2014). Therefore, it was important to include price value as an important factor for this study. The hypothesis for this perceived-benefits-of-use factor was:

> **H5**: Price value has an impact on ADDO programme dispensaries' perception of the potential benefits of using the mHealth information system.

Data collection

The data collection took place in November 2019, via (1) a self-administered questionnaire composed of closed-ended questions; and (2) focus group discussions. For both the questionnaire and the focus groups, the participants were representatives of ADDO programme dispensaries using the mHealth reporting system in the Iringa, Mbeya, and Njombe Regions.

Questionnaire

Questionnaire respondents were asked to respond to statements using a five-point Likert scale (1=strongly disagree; 2=disagree; 3=neutral; 4=agree; 5= strongly agree). The statements, provided in Table 2, were grounded in the updated D&M IS success model (DeLone & McLean, 2013) and the extended unified theory of acceptance and use of technology (UTAUT2) model (Venkatesh et al., 2003), with modifications necessary to fit the context of this particular study. Table 2 shows the non-demographic elements of the questionnaire.

AJIC 25 - 5-July-20 - 20h14.indd 8-9 7/5/2020 8:22:48 PM

Table 2: Contents of the questionnaire

Factor	Code	Item
system quality	SQ1	The mHealth system is easy to learn.
	SQ2	The mHealth system is easy to use.
	SQ3	The mHealth system requires only the minimum number of fields and screens to enter data and generate the required reports.
	SQ4	The mHealth system includes the necessary features and functions for performing my day-to-day activities.
info. quality	IQ1	The data, information, and reports that I receive from the mHealth system are accurate.
	IQ2	The mHealth system provides an output that is exactly what is needed.
	IQ3	The information that I receive from the mHealth system is complete.
	IQ4	Information and reports from the mHealth system are readable, clear, and well-formatted.
	IQ5	Information needed from the mHealth system is always available.
	IQ6	The information that I receive from the mHealth system is up to date.
service quality	SEQ1	IT Unit staff provide prompt support through various communication means such as email, telephone, chat, etc.
	SEQ2	The training provided by the IT Unit staff has enhanced my ability to use the mHealth system.
	SEQ3	The IT Unit staff are competent in the use of the mHealth system.
	SEQ4	The IT Unit staff have adequate knowledge to help me when I experience any problems with the mHealth system.
facilitating conditions	FC1	I have the resources (e.g., mobile phone, Internet connection, etc.) necessary to use the mHealth system.
	FC2	I have the knowledge necessary to use the mHealth system.
	FC3	The mHealth system is similar to other systems I use for my day-to-day activities.
	FC4	Help is available when I get a problem with the mHealth system.
price value	PV1	The mHealth system is a good value for the money.
	PV2	At the current cost of using the mHealth system, the system provides good value.
	PV3	The cost I incur in using the mHealth system is reasonable.
perceived benefits of using the mHealth reporting system	PB1	Using the mHealth system has improved information quality in the Ministry.
	PB2	Using the mHealth system has improved daily information management processes.
	PB3	Using the mHealth system has reduced information errors.

The ADDO programme dispensaries targeted for the survey questionnaire were selected from all the wards in each region, to avoid sampling bias by including a disproportionately high number of stores from a few wards. To determine the minimum acceptable sample size for the questionnaire, the formula proposed by Green (1991) was adopted. That formula states that the minimum sample size should be N > 50+8m, where m is the number of independent variables. The study had five factors. Therefore, the minimum sample size required for this research was 50 + (8*5) = 90. Of the 500 potential respondents who were given the questionnaire in Iringa, Mbeya, and Njombe, 318 returned usable questionnaires. Of these 318 respondents, 64% were female, and 34% were male. Most (87.1%) had attained secondary education, while 11.9% had primary education. A small fraction of the respondents had a tertiary education qualification (1%).

Focus group discussions

Seven focus groups were convened—three in Iringa, two in Mbeya, and two in Njombe—and each group consisted of between five and eight participants, with a total of 38 respondents. The respondents who participated in the focus group discussions were selected from those who completed the questionnaire. The discussions aimed to check the validity of data obtained from the questionnaire and to get a broad range of viewpoints about the issues raised in the questionnaires. The focus groups focused on factors that were found to be significant and helped to identify additional factors that were influencing dispensaries' perceived benefits of using the ADDO mHealth reporting system.

Data analysis

The collected quantitative data was analysed to determine the relationships between the *independent variables* (system quality, information quality, service quality, facilitating conditions, and price value) and the *dependent variable* (perceived benefits of using the mHealth reporting system). The data obtained from focus group discussions were analysed by reviewing each focus group recording to identify commonalities across the responses.

4.Questionnaire findings

Reliability testing

The Cronbach's Alpha was used to test the reliability of the instrument. The Cronbach's Alpha coefficient of the 25 items was 0.917, indicating that the instrument was reliable as the value of Cronbach's Alpha was above 0.5. Furthermore, the result of Cronbach's Alpha for all six variables was above 0.5, as shown in Table 3.

AJIC 25 - 5-July-20 - 20h14.indd 10-11 7/5/2020 8:22:48 PM

Table 3: Cronbach's Alpha coefficients for construct reliability measurement

Factor	Final number of items	Lowest loading	Cronbach's Alpha
system quality (SYSQ3 dropped)	3	0.820	0.891
information quality	4	0.687	0.847
service quality	4	0.649	0.871
facilitating conditions	3	0.665	0.572
price value	3	0.743	0.775
perceived benefits of using the mHealth reporting system	3	0.652	0.870

Correlation results

Correlation between perceived benefits and the 5 dependent variables shows four of five are significant and positive (See Table 4). Only the facilitating conditions factor is not significant. Looking across the correlation matrix there are no correlations among the independent variables greater than 0.700, therefore collinearity is not of great concern and all will proceed to be entered in multiple regression.

Table 4: Correlation matrix

		perceived benefits	system quality	info. quality	service quality	facilitating conditions	price value
perceived benefits	Pearson Correlation	1	.488**	.600**	.614**	-.095	.492**
	Sig. (2-tailed)		.000	.000	.000	.092	.000
	N	318	318	318	318	317	318
system quality	Pearson Correlation	.488**	1	.514**	.561**	-.035	.297**
	Sig. (2-tailed)	.000		.000	.000	.530	.000
	N	318	318	318	318	317	318
information quality	Pearson Correlation	.600**	.514**	1	.689**	.027	.341**
	Sig. (2-tailed)	.000	.000		.000	.637	.000
	N	318	318	318	318	317	318
service quality	Pearson Correlation	.614**	.561**	.689**	1	.045	.412**
	Sig. (2-tailed)	.000	.000	.000		.428	.000
	N	318	318	318	318	317	318
facilitating conditions	Pearson Correlation	-.095	-.035	.027	.045	1	-.182**
	Sig. (2-tailed)	.092	.530	.637	.428		.001
	N	317	317	317	317	317	317
price value	Pearson Correlation	.492**	.297**	.341**	.412**	-.182**	1
	Sig. (2-tailed)	.000	.000	.000	.000	.001	
	N	318	318	318	318	317	318

** Correlation is significant at the 0.01 level (2-tailed).

Identifying the factor structure

Factor analysis was performed using the principal component analysis extraction method on 25 items, with the use of varimax rotation with Kaiser Normalization. Items with loadings greater than 0.3 on any of the factors (i.e., SQ3, IQ1, IQ6) were removed, and the analysis was re-run, as suggested by Samuels (2016). The loadings for each item are shown in Table 5.

Table 5: Factor loadings

Factor	Items in varimax rotation	Loadings
system quality	SQ1	0.863
	SQ2	0.827
	SQ4	0.820
information quality	IQ2	0.687
	IQ3	0.783
	IQ4	0.707
	IQ5	0.704
service quality	SQ1	0.675
	SQ2	0.821
	SQ3	0.814
	SQ4	0.649
facilitating conditions	FC1	0.665
	FC2	0.828
	FC3	0.683
price value	PV1	0.787
	PV2	0.806
	PV3	0.743
perceived benefits of using the mHealth reporting system	IU1	0.652
	IU2	0.835
	IU3	0.810

AJIC 25 - 5-July-20 - 20h14.indd 12-13 7/5/2020 8:22:49 PM

Hypothesis-testing

The five hypotheses derived from the five factors in the research model—system quality, information quality, service quality, facilitating conditions, and price value—were subjected to linear regression analysis using the SPSS software to predict the causal relationship between these factors and the dispensaries' perceptions of the potential benefits of using the mHealth system.

Using the enter method, a significant model emerged: $F(5, 316) = 63.915$, $p < .001$. The model explains 49.9% of the variance (adjusted $R^2 = 0.499$) in dispensaries' perceptions of the potential benefits of use of the mHealth system, as shown in Table 6.

Table 6: Summary of the research model extracted from SPSS

Model	R	R^2	Adjusted R^2
1.	0.712a	0.507	0.499

This model shows that 50.1% of the factors impacting the ADDO dispensaries' perceptions of the potential benefits of using the mHealth reporting system cannot be explained by factors proposed in this model. This is to say, the five factors—system quality, information quality, service quality, facilitating conditions, and price value—can explain only 49.9% of factors impacting the dispensaries' perceptions of the mHealth reporting system in the Iringa, Mbeya and Njombe Regions.

In testing the hypotheses, the beta values indicated the contribution by each factor in the research model, after performing regression analysis using SPSS. The finding in respect of each hypothesis (i.e., the finding as to whether the factor in question had a positive or negative impact) was determined to be significant when the finding had a p-value (calculated probability) of less than 0.05. The results from this study showed that, among the five factors presented in this study, four were found to be having a significant positive impact (i.e., their beta values were positive, at p-values less than 0.05). The *facilitating conditions* factor was not found to be having a significant effect on dispensaries' perceptions of the potential benefits of using the mHealth reporting system, as its p-value was not less than 0.05. Table 7 provides the analysis of variance (ANOVA) values, and Table 8 provides a summary of beta values and p-values for the five factors, as obtained from a linear regression analysis.

Table 7: ANOVA values

Model		Sum of squares	Df	Mean square	F	Sig.
1	Regression	82.693	5	16.539	63.915	.000[b]
	Residual	80.474	311	.259		
	Total	163.167	316			

Table 8: Standardised regression coefficients

Construct	Standard error	Beta	P-value	Implication of p-value
system quality	0.046	0.128	0.010	$p < 0.05$
information quality	0.059	0.278	0.000	$p < 0.05$
service quality	0.058	0.253	0.000	$p < 0.05$
facilitating conditions	0.040	-0.065	0.115	$p > 0.05$
price value	0.039	0.242	0.000	$p < 0.05$

In Table 8, system quality's significant positive impact is indicated by its positive beta and its p-value being less than 0.05 (beta = 0.128, $p < 0.05$). This finding implies that the surveyed dispensaries felt the mHealth system was of high quality and therefore it will bring higher benefits. Information quality's significant positive impact is shown by its positive beta and its p-value of less than 0.05 (beta = 0.278, $p < 0.05$). This finding suggests dispensaries perceived that the mHealth system generated quality reports for record-keeping as well as for sending the data to the higher authorities. Service quality's significant positive impact is indicated by its beta being positive and its p-value being less than 0.05 (beta = 0.253, $p < 0.05$). The lack of significant impact generated by the facilitating conditions factor is indicated by its p-value being greater than 0.05 (beta = -0.065, $p > 0.05$). Price value's significant positive impact is indicated by its positive beta and its p-value being less than 0.05 (beta = 0.242, $p < 0.05$). This finding suggests dispensaries perceived that the cost associated with using the mHealth system was not a limiting factor. The five hypotheses and their beta values and p-values are shown in Table 9.

AJIC 25 - 5-July-20 - 20h14.indd 14-15 7/5/2020 8:22:49 PM

Table 9: Hypothesis-testing findings

Hypothesis	Results
System quality has an impact on ADDO programme dispensaries' perception of the potential benefits of using the mHealth reporting system.	significant with positive effect (beta = 0.128, $p < 0.05$)
Information quality has an impact on ADDO programme dispensaries' perception of the potential benefits of using the mHealth reporting system.	significant with positive effect (beta = 0.278, $p < 0.05$)
Service quality has an impact on ADDO programme dispensaries' perception of the potential benefits of using the mHealth reporting system.	significant with positive effect (beta = 0.253, $p < 0.05$)
Facilitating conditions have an impact on ADDO programme dispensaries' perception of the potential benefits of using the mHealth reporting system.	not significant (beta = -0.065, $p > 0.05$)
Price value has an impact on ADDO programme dispensaries' perception of the potential benefits of using the mHealth reporting system.	significant with positive effect (beta = 0.242, $p < 0.05$)

5. Focus group findings

As a supplement to testing the hypotheses via the survey, the focus group discussions were conducted to gain more insight into the factors influencing dispensaries' perceptions of the potential benefits of using the mHealth information system. The focus groups focused on the four factors that were found via the questionnaire to have a statistically significant impact on the dispensaries' perceptions of the potential benefits: system quality, information quality, service quality, and price value.

System quality

In line with the survey finding that system quality had a significant positive impact on dispensaries' perceptions of the potential benefits of using the mHealth reporting system, there was consensus among the focus group participants that the mHealth system was more convenient to use than the system of filling in paper documents to send to the Ministry. In the words of one participant:

> It is very convenient to use the system for preparing and sending the reports compared to the time when we were doing this manually. We were required to fill in the paper documents and send them to the Ministry. It was taking a lot of time, and the chance of making errors was huge.

Another participant said:

> I have to make a choice, to use the system for managing my office and sending the report electronically or fill in the paperwork. If you compare the benefits, I should learn and use this system.

Nonetheless, some dispensary personnel in the focus groups indicated that the system was difficult to use—because, among other things, it was slow, and it was difficult to see on a mobile device, given the size of the mobile screen. Some focus group participants indicated that they could not use most of the features of the system. Two of the focus group inputs were as follows:

> The system is slow. It takes a bit of time to load the data once you click [the] send button.

> The system is okay but look at my phone and the size of the screen. I have to take [out] my glasses every time I have to use this system. Maybe if they reduce the number of screens, it might be better.

Some dispensary personnel said they preferred using tablets rather than mobile phones, due to small screen sizes making data entry difficult.

Information quality

In keeping with the survey finding that information quality had a significant positive impact on dispensaries' perceptions of the potential benefits of using the mHealth reporting system, there was a consensus in the focus groups that the current system provided better quality information when compared to a paper-based system. However, focus group participants also pointed to potential information-quality improvements that could be made. According to one participant:

> The system should be able to send feedback on our phones when we send monthly reports. This is missing in the current system.

In the words of another focus group participant:

> It would be nice if the system can help me to see the items that are out of stock and are missed from the ordered quantity.

It was also found that some dispensaries wanted to be able to generate information reports for their own use, in addition to generating the reports in the formats required

AJIC 25 - 5-July-20 - 20h14.indd 16-17 7/5/2020 8:22:49 PM

by the Ministry. For instance, dispensaries wanted to be able to keep records of their sales and stocks of the goods for their own use.

Some focus group participants said they sometimes had difficulty meeting the monthly reporting deadlines. It was suggested that notifications could be sent via SMS to remind dispensaries of the deadlines for submitting reports.

Service quality

Although the survey found that service quality had a significant positive impact on dispensaries' perceptions of the potential benefits of using the mHealth system, many participants in the focus groups were critical of what they saw as a lack of training support from the Ministry's IT Unit staff. It seems that the training that was conducted at the beginning of the mHealth project was not sufficient to enable dispensaries to use the system with full effectiveness. As a result, the majority of the focus group participants indicated that they needed more training. For instance, in the words of one participant:

> My opinion is that this system is good, but I ask for the experts to give us more training, as the system has many features which we cannot use […] properly. In the beginning, I had to spend time asking our friends in other shops [as to] how to send monthly reports, but later I mastered it. But the problem has remained in the stock module, in which I think I need more support.

Another participant said:

> The IT staff should visit us more often to check if we have problems. Since they conducted this training [in 2017], it [has been] almost two years now, and I have not seen them. Look, now I have a lot to ask, but I cannot do it. The system is not very difficult to use, but I have so many ideas on how to improve it.

Price value

While the survey results pointed to the price value factor as having a significant positive impact on dispensaries' perceptions of the potential benefits of using the mHealth system, the sentiments in the focus groups were found to be mixed on the matter of cost. Many participants felt that the mHealth system had positive cost implications—a sentiment voiced by this participant:

> I can save time and transportation costs [through] avoiding unnecessary trips to the central office to submit the printed forms. So, it is [really] important to use this system.

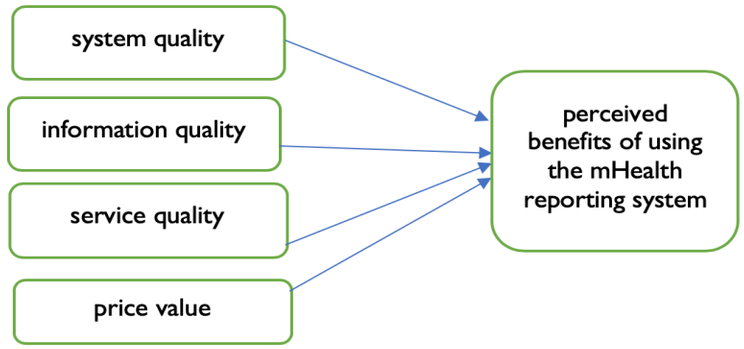
However, at the same time, this same participant criticised the fact that although the system required internet access when sending reports to the Ministry, the cost of the internet was not covered by the Ministry. Other participants were also critical of the internet costs created by the system, indicating that the cost was a barrier to their dispensaries. In the words of one such participant:

> On average, I can spend almost TSH2,000 [approximately USD1) for sending monthly reports. This is a lot of money for a small shop like this.

6. Analysis and conclusions

Based on the results of this study, we propose, in Figure 2, a four-factor model for understanding and addressing the factors currently impacting dispensaries' perceptions of the potential benefits of using the mHealth reporting system in Tanzania.

Figure 2: Factors impacting dispensaries' perceptions of potential benefits of using ADDO mHealth reporting system



Our study provides clear evidence that four of the five factors, according to our survey results, have a significantly positive impact on dispensaries' perceptions of the potential benefits of using the mHealth reporting system. Despite these findings, the focus group discussions made it clear that there is still much room for improvements. In respect of system quality, the core focus group critiques, that the ADDO mHealth reporting system is too slow and that it is difficult to use on a mobile handset, certainly need to be addressed by the Ministry. In respect of price value, the chief difficulty is the lack of free internet provision to the dispensaries for use in their reporting. The implementers of mHealth projects need to find ways to provide subsidised or free internet use for the system's users. The ADDO programme's managers should negotiate with mobile firms for the provision of discounted or free internet services to the participating dispensaries. There are already precedents in Tanzania for such arrangements. In Zanzibar, for instance, mobile operator Airtel provided free internet services for users of an mHealth system aimed at reducing maternal mortality in rural Tanzania (Mbaruku et al., 2018).

AJIC 25 - 5-July-20 - 20h14.indd 18-19 7/5/2020 8:22:49 PM

In respect of service quality, the focus group discussions indicated that the Ministry's training provision for dispensaries has not been sufficiently intensive. Building the perceived benefits of the system in the eyes of the dispensaries is, therefore, contingent on improved and ongoing training support from the Ministry's IT Unit. In respect of information quality, a shortcoming cited in the focus group discussions is the lack of provision for ADDO dispensaries to use the mHealth system to generate reports, not only for the Ministry, but also for their own use. To increase perceived benefit, the Ministry needs to build in the functionality necessary for dispensaries to customise reports to meet their own needs at the facility level.

References

Aranda-Jan, C. B., Mohutsiwa-Dibe, N., & Loukanova, S. (2014). Systematic review on what works, what does not work and why of implementation of mobile health (mHealth) projects in Africa. *BMC Public Health*, *14*(188). https://doi.org/10.1186/1471-2458-14-188

Barclay, V. C., Smith, R. A., & Findeis, J. L. (2012). Surveillance considerations for malaria elimination. *Malaria Journal*, *11*, 2–5. https://doi.org/10.1186/1475-2875-11-304

Bhattacherjee, A. (2001). Understanding information systems continuance: An expectation-confirmation model. *MIS Quarterly*, *25*(3), 351–370. https://doi.org/10.2307/3250921

Bogan, M., Mushi, C., Esch, J. Van, Wakabi, T., Lesh, N., Derenzi, B., & Mitchell, M. (2009). Improving standards of care with mobile applications in Tanzania. In *Presented at the W3C Workshop on the Role of Mobile Technologies in Fostering Social and Economic Development in Africa*.

Burton-Jones, A., & Volkoff, O. (2017). How can we develop contextualized theories of effective use? A demonstration in the context of community-care electronic health records. *Information Systems Research*, *28*(3), 468–489. https://doi.org/10.1287/isre.2017.0702

Chalker, J. C., Vialle-Valentin, C., Liana, J., Mbwasi, R., Semali, I. A., Kihiyo, B., … Ross-Degnan, D. (2015). What roles do Accredited Drug Dispensing Outlets in Tanzania play in facilitating access to antimicrobials? Results of a multi-method analysis. *Antimicrobial Resistance and Infection Control*, *4*(1), 1–11. https://doi.org/10.1186/s13756-015-0075-2

Chang, L. W., Kagaayi, J., Arem, H., Nakigozi, G., Ssempijja, V., Serwadda, D., … Reynolds, S. J. (2011). Impact of a mHealth intervention for peer health workers on AIDS care in rural Uganda: A mixed methods evaluation of a cluster-randomized trial. *AIDS and Behavior*, *15*, 1776–1784. https://doi.org/10.1007/s10461-011-9995-x

Chib, A., Velthoven, M. H. Van, & Car, J. (2014). mHealth adoption in low-resource environments : A review of the use of mobile healthcare in developing countries. *Journal of Health Communication: International Perspectives*, *20*(1), 37–41. https://doi.org/10.1080/10810730.2013.864735

DeLone, W. H., & McLean, E. R. (1992). Information systems success - the quest for a dependent variable. *Information Systems Research*, *3*(1), 60–95. https://doi.org/10.1287/isre.3.1.60

DeLone, W. H., & McLean, E. R. (2003). The DeLone and McLean model of information systems success: A ten-year update. *Journal of Management Information Systems / Spring*, *19*(4), 9–30. https://doi.org/10.1080/07421222.2003.11045748

DeLone, W. H., & McLean, E. R. (2013). The DeLone and McLean model of information systems success : A ten-year update. *Journal of Management Information Systems*, *8*(4), 9–30. https://doi.org/10.1080/07421222.2003.11045748

DeLone, W. H., & McLean, E. R. (2016). Information systems success measurement. *Foundations and Trends in Information Systems*, *2*(1), 1–116. https://doi.org/10.1561/2900000005

DeRenzi, B., Birnbaum, B., Findlater, L., Mangilima, J., Payne, J., Parikh, T., … Lesh, N. (2012). Improving community health worker performance through automated SMS. In *Proceedings of the Fifth International Conference on Information and Communication Technologies and Development* (pp. 25–34). Atlanta, GA, USA. https://doi.org/10.1145/2160673.2160677

Dillip, A., Kimatta, S., Embrey, M., Chalker, J. C., Valimba, R., Malliwah, M., … Johnson, K. (2017). Can formalizing links among community health workers, Accredited Drug Dispensing Outlet dispensers, and health facility staff increase their collaboration to improve prompt access to maternal and child care? A qualitative study in Tanzania. *BMC Health Services Research*, *17*(1), 1–11. https://doi.org/10.1186/s12913-017-2382-1

Embrey, M., Vialle-Valentin, C., Dillip, A., Kihiyo, B., Mbwasi, R., Semali, I. A., … Ross-Degnan, D. (2016). Understanding the role of accredited drug dispensing outlets in Tanzania's health system. *PLoS ONE*, *11*(11), 1–16. https://doi.org/10.1371/journal.pone.0164332

Francis, F., Ishengoma, D. S., Mmbando, B. P., Rutta, A. S. M., Malecela, M. N., Mayala, B., … Michael, E. (2017). Deployment and use of mobile phone technology for real-time reporting of fever cases and malaria treatment failure in areas of declining malaria transmission in Muheza district north-eastern Tanzania. *Malaria Journal*, *16*(1), 1–14. https://doi.org/10.1186/s12936-017-1956-z

Githinji, S., Kigen, S., Memusi, D., Nyandigisi, A., Mbithi, A. M., Wamari, A., … Zurovac, D. (2013). Reducing stock-outs of life saving malaria commodities using mobile phone text-messaging: SMS for Life study in Kenya. *PLoS ONE*, *8*(1), 1–8. https://doi.org/10.1371/journal.pone.0054066

Green, S. . (1991). How many subjects does it take to do a regression analysis. *Multivariate Behavioral Research*, *26*(3), 499–510. https://doi.org/10.1207/s15327906mbr2603_7

Heeks, R. (2006). Health information systems: Failure, success and improvisation. *International Journal of Medical Informatics*, *75*(2), 125–137. https://doi.org/10.1016/j.ijmedinf.2005.07.024

Holden, R. J., & Karsh, B. (2010). The Technology Acceptance Model : Its past and its future in health care. *Journal of Biomedical Informatics*, *43*(1), 159–172. https://doi.org/10.1016/j.jbi.2009.07.002

Karahanna, E., Straub, D. W., & Chervany, N. L. (1999). Information technology adoption across time: A cross-sectional comparison of pre-adoption and post-adoption beliefs. *MIS Quarterly: Management Information Systems*, *23*(2), 183–213. https://doi.org/10.2307/249751

Kim, S., Lee, K., Hwang, H., & Yoo, S. (2016). Analysis of the factors influencing healthcare professionals' adoption of mobile electronic medical record (EMR) using the unified theory of acceptance and use of technology (UTAUT) in a tertiary hospital. *BMC Medical Informatics and Decision Making*, *16*(12), 1–12. https://doi.org/10.1186/s12911-016-0249-8

AJIC 25 - 5-July-20 - 20h14.indd 20-21 7/5/2020 8:22:49 PM

Leon, N., Schneider, H., & Daviaud, E. (2012). Applying a framework for assessing the health system challenges to scaling up mHealth in South Africa. *BMC Medical Informatics and Decision Making*, *12*(1). https://doi.org/10.1186/1472-6947-12-123

Limayem, M., Hirt, S. G., & Cheung, C. M. K. (2007). How habit limits the predictive power of intention: The case of information systems continuance. *MIS Quarterly: Management Information Systems*, *31*(4), 705–737. https://doi.org/10.2307/25148817

Mbaruku, G., Therrien, M. S., Tillya, R., Mbuyita, S., Mtema, Z., Kinyonge, I., … Miller, S. (2018). Implementation project of the non-pneumatic anti-shock garment and m-communication to enhance maternal health care in rural Tanzania. *Reproductive Health*, *15*(1), 177. https://doi.org/10.1186/s12978-018-0613-5

Mikkelsen-Lopez, I., Shango, W., Barrington, J., Ziegler, R., Smith, T., & DeSavigny, D. (2014). The challenge to avoid anti-malarial medicine stock-outs in an era of funding partners: The case of Tanzania. *Malaria Journal*, *13*(1), 1–9. https://doi.org/10.1186/1475-2875-13-181

Ng, S. N., Matanjun, D., D'Souza, U. J. A., & Alfred, R. (2015). Understanding pharmacists' intention to use medical apps. *Electronic Journal of Health Informatics*, *9*(1), 1–17.

Rutta, E. (2014). *Medicines in Health Systems : Advancing access, affordability, and appropriate use*. *Alliance for Health Policy and Systems Research Flagship Report 2014*. https://doi.org/ISBN 978 92 4 150762 2

Rutta, E., Liana, J., Embrey, M., Johnson, K., Kimatta, S., Valimba, R., … Sillo, H. (2015). Accrediting retail drug shops to strengthen Tanzania's public health system: An ADDO case study. *Journal of Pharmaceutical Policy and Practice*, *8*(1), 1–15. https://doi.org/10.1186/s40545-015-0044-4

Rutta, E., Senauer, K., Johnson, K., Adeya, G., Mbwasi, R., Liana, J., … Alphonce, E. (2009). Creating a new class of pharmaceutical services provider for underserved areas: The Tanzania Accredited Drug Dispensing Outlet experience. *Progress in Community Health Partnerships : Research, Education, and Action*, *3*(2), 145–153. https://doi.org/10.1353/cpr.0.0063

Samuels, P. (2016). *Advice on Exploratory Factor Analysis*. https://doi.org/10.13140/RG.2.1.5013.9766

Seddon, P. B., & Kiew, M. (1995). A partial test and development of Delone and Mclean's model of IS success. *Australasian Journal of Information Systems*, *4*(1), 90–109. https://doi.org/http://dx.doi.org/10.3127/ajis.v4i1.379

Smith, A. D., & Motley, D. (2010). Operational and customer relationship management considerations of electronic prescribing among pharmacists. *International Journal of Electronic Healthcare*, *5*(3), 245. https://doi.org/10.1504/IJEH.2010.034175

US President's Malaria Initiative. (2017). *Malaria Operational Plan FY 2017*.

Vaghefi, I., & Tulu, B. (2019). The continued use of mobile health apps: Insights from a longitudinal study. *JMIR MHealth and UHealth*, *7*(8), 1–11. https://doi.org/10.2196/12983

Venkatesh, V., Morris, M. G., Hall, M., Davis, G. B., Davis, F. D., & Walton, S. M. (2003). User acceptance of information technology : Toward a unified view. *MIS Quarterly*, *27*(3), 425–478. https://doi.org/10.2307/30036540

Venkatesh, V., Thong, J. Y. L., & Xu, X. (2012). Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. *MIS Quarterly*, *36*(1), 157–178. https://doi.org/10.2307/41410412

Pilot Testing of an Information Extraction (IE) Prototype for Legal Research

Brenda Scholtz
Associate Professor, Faculty of Science, Nelson Mandela University, Port Elizabeth, South Africa
https://orcid.org/0000-0002-0844-1383

Thashen Padayachy
Master's Graduate, Faculty of Science, Nelson Mandela University, Port Elizabeth, South Africa
https://orcid.org/0000-0001-9319-7356

Oluwande Adewoyin
PhD Candidate, Faculty of Science, Nelson Mandela University, Port Elizabeth, South Africa
https://orcid.org/0000-0001-6097-0795

Abstract
This article presents findings from pilot testing of elements of an information extraction (IE) prototype designed to assist legal researchers in engaging with case law databases. The prototype that was piloted seeks to extract, from legal case documents, relevant and accurate information on cases referred to (CRTs) in the source cases. Testing of CRT extraction from 50 source cases resulted in only 38% (n = 19) of the extractions providing an accurate number of CRTs. In respect of the prototype's extraction of CRT attributes (case title, date, journal, and action), none of the 50 extractions produced fully accurate attribute information. The article outlines the prototype, the pilot testing process, and the test findings, and then concludes with a discussion of where the prototype needs to be improved.

Keywords
information retrieval (IR), information extraction (IE), natural language processing (NLP), legal cases, document databases, source cases, cases referred to (CRTs)

Acknowledgement
Financial assistance towards this research from the South African National Research Foundation (NRF) is hereby acknowledged. Opinions expressed and conclusions arrived at are those of the authors and are not to be attributed to the NRF.

DOI: https://doi.org/10.23962/10539/29192

AJIC 25 - 5-July-20 - 20h14.indd 22-1 7/5/2020 8:22:49 PM

Recommended citation
Scholtz, B., Padayachy, T., & Adewoyin, O. (2020). Pilot testing of an information extraction (IE) prototype for legal research. *The African Journal of Information and Communication (AJIC), 25*, 1-20. https://doi.org/10.23962/10539/29192

This article is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) licence: https://creativecommons.org/licenses/by/4.0

1. Introduction

The research we describe in this article—the pilot testing of an information extraction (IE) prototype designed to assist legal researchers—is the continuation of a research project that was outlined in an earlier publication. As outlined in that Padayachy, Scholtz and Wesson (2018) paper, the ultimate aim of the research project is to develop an information extraction (IE) model that, when applied to a database of legal cases, can (1) determine the case most applicable to a point of law; and (2) store the findings in a database.

The Padayachy et al. (2018) paper described the process of designing the IE model. Following a design science research (DSR) framework, the model was based on a literature review, a review of existing systems at a sample firm, and interviews with experts. The 2018 paper also described results from the initial testing of the prototype with small amounts of data and using basic queries in a graph database. In this article, we describe and discuss the results of the next phase of testing of the IE prototype, using a larger pilot sample (50 legal case documents) and testing two of the prototype's four processes: information extraction (IE) and information storage. The legal documents used were provided to us by an organisation in the South African legal domain (which we call LegalCo, for anonymity purposes).

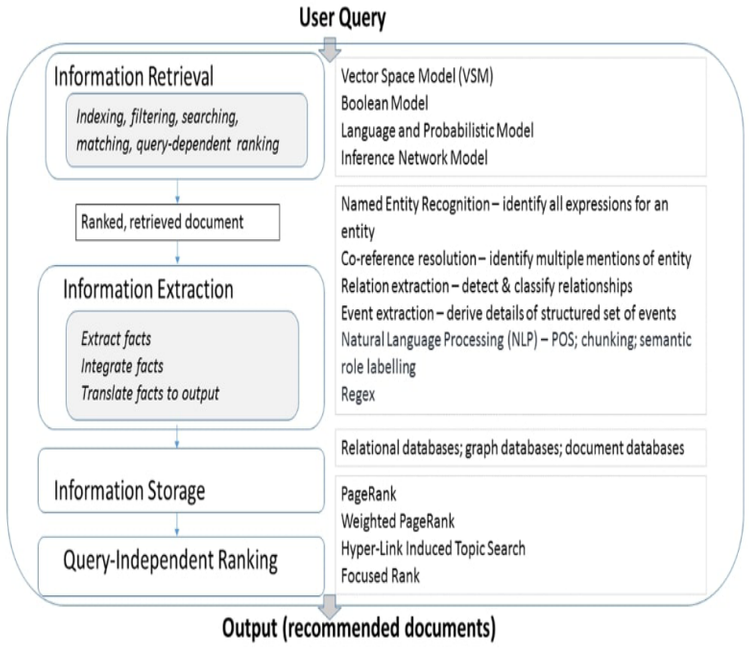
2. The IE prototype

Our model, which we call the IE Model for Legal Cases, consists of the following four main processes:

- information retrieval (IR), with indexing;
- information extraction (IE);
- information storage; and
- query-independent ranking of the most applicable case (MAC).

Figure 1 provides detail on the elements assembled in order to actualise the four processes, and to integrate the processes in a manner designed to produce relevant case information for a legal researcher. The model also incorporates the options of using a document database, a graph database, and/or a relational database for document storage.

Figure 1: IE Model for Legal Cases



Information retrieval (IR)

The first set of processes that needs to take place in our model, based on a query from the user, is IR, which is the process that deals with the representation, storage, and search of a collection of data in response to a request (Roshdi & Roohparvar, 2015). The data could be in the form of video, audio, or text. For our model, text is the focus, with the assumption that information compilation for legal cases is currently, for the most part, in textual form. An IR system does its job by recovering relevant documents from a compilation of resources made by autonomous modules or a database (Kumar & Sharma, 2018).

As illustrated in Figure 1, *indexing* is the first process in IR. Indexing is the process of representing a document's content by creating a logical view of the document in a collection by means of keywords or terms (Kumar & Sharma, 2018). An index is constructed from a case's keywords in order to act as a pointer to a stored document for rapid, accurate retrieval and storage. *Filtering*, also represented in Figure 1, is done once a query has been entered by a user. Filtering removes all stop words from the user's query. (A stop word is a word that is very commonly used and that is programmatically ignored by a search system during indexing, search, and retrieval of information (Schofield, Magnusson, & Mimno, 2017).) *Searching* is then done on the indexed documents. *Matching* compares the two representations (i.e., the

AJIC 25 - 5-July-20 - 20h14.indd 2-3 7/5/2020 8:22:49 PM

indexed documents and the user's information need) so that *query-dependent ranking* and *ranked retrieval* of documents can take place. The output of this process is a set of ranked, retrieved documents. The user can then provide feedback, if different information is needed, by altering the query.

For text retrieval, four matching techniques are deployed, using the following four models:

- the vector space model (VSM);
- the Boolean model;
- the language and probabilistic model; and
- the inference network model.

The VSM recovers text automatically, through representation of documents and queries as weighted vectors. The VSM was successfully used in the study by Firdhous (2010) to retrieve legal documents based on user queries. The VSM was also used, by Aritomo and Watanabe (2019), to generate a searchable encryption technique that enabled a keyword search for documents through encryption. (However, the technique was applied to request-for-comment documents and not legal case documents.) Irrespective of the domain, the VSM is the most commonly used IR model (Al-Anzi & AbuZeina, 2018). Frequency–inverse document frequency (tf–idf or TFIDF for short) can be used in VSM for weight estimation; it is a numerical statistic that is used to reflect the importance or weighting of a word to a document in a collection. Retrieval is based on the degree of similarity between the term vector and the query vector, while recovered results are ranked using cosine similarity (a measure of similarity between two non-zero vectors of an inner product space that measures the cosine of the angle between them), which is a major strength of the VSM. However, the VSM suffers from the inability of the vector space to deal with polysemy and synonymy. Also, index terms are assumed to be mutually independent (Kumar & Sharma, 2018).

The Boolean model is the simplest matching technique for text retrieval. The model uses Boolean algebra for exact matching, and represents documents and queries as sets of terms (Liu, 2011). Results of the Boolean model can be true or false, with queries indicated using OR, XOR and AND relators. Application of the Boolean model is characterised by a large assembly of terms, and information is outputted if all the Boolean terms are present in the resource—but the Boolean model is limited by its requirement for strict matching that generates "nothing or too many" problems (Pandey, Mathur, & Joshi, 2019). Also, the model does not deal with frequency and term weights, leading to unranked results. Therefore, users must have a strong knowledge of query-making.

The language and probabilistic model recovers documents or text by placing emphasis on the probabilities of different factors in decision-making, such as documents' relevance, for ranking (Losee, Bookstein, & Yu, 1986; Losee, 2015). However, the model is affected by the difficulties of combining different ranking functions into a single function, ranking speed, the non-convex nature of ranking algorithms, and retrieval of irrelevant information (Pandey et al., 2019).

The inference network model consists of a directed, acyclic graph that contains nodes. The nodes represent events with possible outcomes whilst the arcs of the network represent probabilistic dependencies between the events (Croft et al., 2015). In the context of IR, nodes represent the observation of a document or document features. The events in an inference network model are binary, meaning that true and false are the only two possible outcomes.

Information extraction (IE)

The output of the IR processes (the ranked documents) can then be used as inputs to the IE process. In the IE process, additional filtering is applied to the ranked data and the results are saved to a data repository. From the data repository, the saved results can then be parsed using algorithms such as those that perform clustering or query-independent ranking (e.g., Google PageRank). The main techniques used in IE are:

- named entity recognition (NER) (Abdelmagid et al., 2015; Piskorski & Yangarber, 2013);
- co-reference resolution (Iida, Inui, & Matsumoto, 2003; Piskorski & Yangarber, 2013);
- relation extraction (RE) (Piskorski & Yangarber, 2013);
- event extraction (EE) (Piskorski & Yangarber, 2013);
- natural language processing (NLP) (Piskorski & Yangarber, 2013); and
- regular expressions (regex) usage (Goyvaerts & Levithan, 2009).

NER is an IE technique that processes extracted information from unstructured and structured text (Abdelmagid et al., 2015). When the technique is applied, all the expressions related to an entity are identified. In addition, NER can involve extracting descriptive information from text about an entity, and completing a template based on the extracted information. Two main tasks are involved in NER: identification and classification of predefined entities. Piskorski and Yangarber (2013) identify organisations, persons, temporal expressions, and numerical expressions as examples of predefined entities.

Co-reference resolution requires the identification of multiple mentions of the same entity (Piskorski & Yangarber, 2013). These mentions can be named, pronominal, nominal, or implicit. A named mention refers to an entity by name, e.g., "General Electric", while a pronominal mention refers to an entity by use of a pronoun, e.g., "*he* forgot to buy food". A nominal mention refers to an entity by a noun phrase, e.g., "*The company* unveiled future plans". Implicit mention uses zero-anaphora to refer

AJIC 25 - 5-July-20 - 20h14.indd 4-5 7/5/2020 8:22:49 PM

to an entity. Zero-anaphora is a gap in a sentence that has an anaphoric function and is often used to refer to an expression that provides necessary information to understand the sentence (Iida et al., 2003). An example of an implicit mention that uses zero-anaphora is seen in "There are two roads to eternity, *a straight and narrow,* and *a broad and crooked.*" In this example, the gaps of the sentence are "a straight and narrow" and "a broad and crooked".

RE is a technique for the detection and classification of predefined relationships between entities identified in a body of text (Piskorski & Yangarber, 2013), e.g., *PersonA is an employee of PersonB*.

EE looks for events in the text and derives detailed and structured sets of information about the events. It is said to be one of the most difficult IE tasks, as it needs to extract the information necessary to answer the questions "who did what to whom, when, where, through what methods?"

NLP can be used to analyse and produce meaning from text that has been extracted from sources such as documents or websites (Singh, 2018). NLP is divided into two categories, namely language processing and language generation. Language processing refers to the analysis of language to produce meaningful representations, whilst language generation refers to producing language from a representation (Liddy, 2001). NLP can be applied to various activities such as speech understanding, IE, and knowledge acquisition (Chowdhary, 2012). In the context of IE, NLP can be applied during the fact extraction process. In the context of our interest in ultimately developing a prototype that can rank and recommending the MAC, NLP can be applied to text that has been extracted from legal cases. Common NLP techniques are part-of-speech (POS) tagging, stop-word removal, parsing, chunking, NER, and semantic role labelling (Chopra, Prashar, & Sain, 2013; Collobert et al., 2011; Vijayarani et al., 2015).

With POS tagging, each word in a set of text is labelled with a unique tag to indicate the word's syntactic role. Words are labelled based on English parts of speech such as nouns, verbs, and adjectives (Collobert et al., 2011). POS tagging is a simplified form of morphological analysis, as words are only tagged, not analysed to find internal structure (Indurkya & Damerau, 2010). Stop-word removal involves removing commonly-used words that are usually articles, prepositions, or pronouns. Parsing refers to determining the grammatical structure of phrases or sentences. Chunking, also known as shallow parsing, labels segments of a sentence with syntactic constituents such as noun or verb phrases (Collobert et al., 2011). In the context of NLP, NER involves labelling elements in a sentence according to different categories, e.g., "person", "location". Semantic role labelling assigns semantic roles to syntactic constituents of a sentence (Collobert et al., 2011).

In regex usage, specific text patterns are used for searching bodies of text, replacing text, segregating text into smaller bodies, and rearranging pieces of text (Goyvaerts & Levithan, 2009). Regex, if correctly used, can simplify programs and text processing tasks by minimising the amount of code needed for processing. Regex usage differs from NLP as none of the NLP phases need to be applied when using regex with bodies of text, i.e., regex can be used directly on an unprocessed body of text.

Information storage

The information storage element is a crucial part of IE, and the data repository selected impacts the efficiency and performance of the model. For text that is structured or semi-structured, Mooney and Bunescu (2005) recommend that IE is performed first on the text, and then the extracted text is transformed to a relational database. However, relational databases tend to decrease in efficiency and performance over time, especially when the data stored increases. An alternative, to overcome the inefficiencies of relational databases, is to use a graph database (Batra & Tyagi, 2012) or a document database (Roy-Hubara & Sturm, 2020). A graph database uses graphs to store information in nodes and allows for the creation of relationships between nodes using edges (Batra & Tyagi, 2012). Among the benefits of graph databases are performance and flexibility. A document database is a non-relational database that stores data in the form of documents that can be grouped together to form collections (Moniruzzaman & Hossain, 2013). Documents can be viewed as objects that contain typed values such as strings, binary values, or arrays. Unlike relational databases that store data across multiple tables and columns, document databases store data in a single document. This helps to eliminate the need for JOIN operators. Data can be stored in three types of structures, namely XML, JavaScript Option Notation (JSON), or Binary JSON (BSON).

Query-independent ranking of selected cases

The information storage process needs to apply techniques and algorithms to save the extracted facts into a data repository so that the last process—query-independent ranking of the selected cases—can be performed. In our model, the query-independent ranking process will return a recommendation of the MAC to the user. For this process, the user of our model will either use an adapted version of Google PageRank or will execute a query on the nodes within the database.

3. Testing of the IE prototype

For our pilot testing of two of our model's components—IE and information storage—the company that we refer to as LegalCo provided the sample data, which consisted of legal case files from the *All South African Law Reports* (LexisNexis South Africa, n.d.).

AJIC 25 - 5-July-20 - 20h14.indd 6-7 7/5/2020 8:22:49 PM

Legal cases can be separated into two parts:

- *the source case* and its primary attributes (case name, case division, case date); and
- *cases referred to*—which we have given the acronym CRTs—with each CRT having four attributes: title, journal, action, and date.

Each source case can have many CRTs.

The main problem addressed by our pilot testing was how to accurately retrieve CRT information from source case files. The prototype we tested in this pilot study did not include the initial step of IR or the final step of query-independent ranking, i.e., determining the MAC. The requirements for the prototype for this study were therefore only those relating to the IE and information storage steps, as summarised in Table 1. The Natural Language Toolkit (NLTK) was used to implement NLP for IE, since it supports most NLP tasks.

Table 1: Techniques and software used for the prototype's IE and information storage components

Processes	Functionality and techniques used	Software
information extraction (IE)	Extract the XML contents of a .docx formatted legal case	Zipfile, Python library
	Parse legal cases in .docx (MS Word) format (NLP, regular expressions, tokenisation, and stop-word removal)	NLTK RE, LXML, Python library
IE and information storage	Develop the prototype	Pycharm IDE
information storage	Interact with the Neo4j graph database	Neo4j, Python library
	Set up document database	MongoDB
	Manage document database	MongoDB Compass desktop application

Table 2 provides an overview of the iterations/equations and experiments we conducted using the prototype.

Table 2: Iterations and experiments

Iteration	Measures	Experiment Set	Number of Source Cases
1: Difference ratio for a source case	Basic functionality, error checking and effectiveness	1	10
2: Total difference ratio for CRTs	Accuracy and execution time	2	10
		3	50, 102

Iteration 1: Extraction of facts from a legal source case

The aim of iteration 1 (not discussed in detail in this article) was to analyse a typical legal case document from the *All South African Law Reports*, determine what information would be important, and test the accuracy of the processes and techniques applied (as specified in Table 2). In this iteration, the CRTs were not considered, since this was the more complex requirement. In the set of experiments conducted in this iteration, 10 legal cases from the LegalCo database—all cases from the *All South African Law Reports*—were used as a test set (coded as cases U1 to U10). In this iteration the experiments used a graph database first and then a document database. Use of the graph database allowed for the extracted facts to be stored in nodes that were connected to each other by specific relationships. Use of the document database allowed for the extracted facts to be stored as documents with multiple types of data embedded into a single document.

Iteration 2: CRT extraction difference ratio

The aim of iteration 2 (the focus of this article) was to test the accuracy of the prototype's extraction of CRTs from source cases. The equation used to determine the *CRT extraction difference ratio for a single source case* was as follows:

Equation 1

$$Xi = \frac{Ai}{Bi}$$

Where:
Xi = the CRT extraction difference ratio for source case *i*;
Ai = the CRT output frequency count for source case *i* that differs from the expected CRT output for source case *i*; and
Bi = the expected CRT output frequency count for source case *i*.

AJIC 25 - 5-July-20 - 20h14.indd 8-9 7/5/2020 8:22:49 PM

The equation used to determine the *total CRT extraction difference ratio for multiple source cases* was as follows:

Equation 2

$$Y = \frac{\sum_{i=1}^{n} A_i}{\sum_{i=1}^{n} Bi}$$

Where:
Y = total CRT extraction difference ratio for multiple source cases;
Ai = the CRT output frequency count for source case *i* that differs from the expected CRT output for source case i; and
Bi = the expected CRT output frequency count for source case *i*.

An ideal value for the difference ratio is 0, since this indicates that there is no difference between the actual output and the expected output (Conroy, 2016). In our experiments, as described in the next section of this article, an error margin of 10% was applied, implying that a difference ratio of 0.1 or less was considered acceptable.

4. Findings

Iteration 1: Experiment set 1

Both a graph database and a document database were investigated and tested during experiment set 1. The graph database implementation used the Python library provided by Neo4j, and the Neo4j desktop application, to locally create and connect to a graph database. During the first phase, dummy data was created and inserted into the graph database as nodes. The dummy data consisted of recipes and drinks associated with each recipe. The insertion of the data was successful. The nodes were the recipes and the relationships were the associated relationships with one or more drinks. However, errors were encountered when trying to create the nodes with relationships. Multiple attempts were made to resolve this error, but no solution could be found. This led to iteration 2, experiment set 2, where a document database was used successfully, as described in the next section.

Iteration 2: Experiment set 2

Experiment set 2 used the document database MongoDB (MongoDB, n.d.) and 10 source cases (coded as cases U1 to U10). In a document database, the data is stored as key-value pairs where both the keys and values are searchable. For experiment set 2, the extracted data was converted to a Python dictionary, whereby data values could be associated with keys, thus allowing for the keys and values to be searchable.

It was found that the primary attributes from the documents were correctly extracted after some cleaning, but the CRTs were not always extracted correctly. U1 to U10 had a total CRT extraction difference ratio of 0.006, implying that 0.6% of the

number of observed CRTs in the source cases were different from what was expected. Since a difference ratio of 0.1 or less was considered acceptable, the result of 0.006 was therefore an acceptable value. (The sample size of 10 was small, and thus only appropriate for a pilot study of the kind we were conducting.)

Some of the CRTs that had no actions were in different formats, which could account for why they were not extracted. The use of .docx-formatted documents was found to be better than .pdf-formatted documents, since the paragraph tags that store text in XML could, in the .docx-formatted (MS Word) documents, be accessed and processed. All subsequent testing was thus conducted using MS Word documents, and this was one limitation of this study.

Iteration 2: Experiment set 3

The first part of experiment set 3 used 50 source cases (F1 to F50) from South African case law during the period 1996 to 2018. The aim of this first part was to determine the prototype's extraction accuracy for extracting the number of CRTs. The total CRT extraction difference ratio, determined through the application of equation 2, was used as the measure of accuracy. Table 3 provides the frequency count results for each of the 50 source cases.

Table 3: CRTs extracted

Case	Actual no. of CRTs in the source case	No. of CRTs extracted from the source case	Absolute difference (excess or shortfall in no. of CRTs extracted)	Difference ratio for CRTs (difference between extracted no. and actual no.)
F1	45	46	1	0.02
F2	5	5	0	0.00
F3	42	43	1	0.02
F4	5	4	1	0.20
F5	6	6	0	0.00
F6	2	2	0	0.00
F7	15	11	4	0.27
F8	7	8	1	0.14
F9	5	5	0	0.00
F10	1	2	1	1.00
F11	7	8	1	0.14
F12	3	3	0	0.00
F13	15	18	3	0.20
F14	9	10	1	0.11

AJIC 25 - 5-July-20 - 20h14.indd 10-11 7/5/2020 8:22:50 PM

F15	33	33	0	0.00
F16	5	4	1	0.20
F17	6	8	2	0.33
F18	23	23	0	0.00
F19	8	4	4	0.50
F20	10	25	15	1.50
F21	10	10	0	0.00
F22	11	27	16	1.45
F23	23	24	1	0.04
F24	10	19	9	0.90
F25	23	19	4	0.17
F26	26	25	1	0.04
F27	12	13	1	0.08
F28	6	6	0	0.00
F29	15	15	0	0.00
F30	15	19	4	0.27
F31	18	19	1	0.06
F32	11	12	1	0.09
F33	6	6	0	0.00
F34	7	9	2	0.29
F35	10	11	1	0.10
F36	17	17	0	0.00
F37	27	21	6	0.22
F38	10	10	0	0.00
F39	75	60	15	0.20
F40	5	5	0	0.00
F41	6	6	0	0.00
F42	8	8	0	0.00
F43	13	13	0	0.00
F44	13	13	0	0.00
F45	12	11	1	0.08
F46	6	5	1	0.17
F47	15	17	2	0.13
F48	19	23	4	0.21
F49	13	17	4	0.31
F50	3	3	0	0.00
Totals	697	731	110	9.46
Average				0.19

Table 4 provides a summary of the accuracy results for the 50 cases according to the frequency, and frequency percentage per range of difference ratios, calculated for the CRT extraction numbers. As can be seen in the table, of the 50 extractions (from the 50 source cases used in the test), 38% (n = 19) had a difference ratio of 0, i.e., in only 38% of the extractions, the number of CRTs extracted was the same as the actual number of CRTs in the source case.

As also shown in Table 4, an overall difference ratio of 0.1578 was observed across the 50 source cases, i.e., the number of CRTs extracted was, on average, 16% different from the actual number in the source case. This difference ratio was higher than the 10% margin of error specified (i.e., in the first three rows in Table 4). Based on the 10% margin of error, 54% (n = 27) of the CRT extraction numbers fell within the margin of error, and 44% (n = 22) of the CRT extraction numbers fell outside the margin of error, thus highlighting the complexities involved in information extraction of this nature.

Table 4: Difference ratio ranges for CRT extractions

Difference ratio range	Frequency (n = 50)	Frequency %		
0	19	38		
1	1	2		
0.01 to 0.09	7	14		
0.1 to 0.5	19	38		
0.6 to 0.9	1	2		
1 to 1.5	3	6		
Totals	**50**	**100**	**Total difference ratio**	**0.157819**

Extraction accuracy for CRT attributes

To analyse the accuracy of the extraction of CRT attributes (title, date, journal, and action), we assigned the following categorisations:

- perfect;
- partial; and
- not (extracted).

For example, if a source case had five CRTs and all five had perfectly extracted attributes, then the extraction would have been classified as "perfect". If only some of the five CRTs were perfectly extracted, and others were only partially extracted, the extraction was categorised as "partial". If none of the five CRTs was extracted,

AJIC 25 - 5-July-20 - 20h14.indd 12-13 7/5/2020 8:22:50 PM

the extraction was categorised as "not" extracted. Table 5 provides a summary of the count of cases that were categorised as being perfect, partial, not, and noisy. Among the extractions of CRT attributes from the 50 source cases, none were found to be "perfect"; 96% (n = 48) were categorised as "partial"; and 4% (n = 2) were categorised as "noisy". We regarded the "partial" extractions as satisfactory for the purposes of this pilot phase, but with further study needed into the causes of incomplete extractions.

Table 5: Extraction accuracy for CRT attributes

Categories	Frequency (n = 50)
perfect	0
partial	48
not	0
noisy	2

For the 48 partial extractions and 2 noisy extractions, further investigation was conducted to determine the cause. Three main causal categories were identified:

- incorrect extraction of extra lines;
- incorrect splitting of CRTs; and
- noisy CRT data.

"Extra lines" errors occurred when non-CRT lines of text, which were similarly formatted to CRT text, were incorrectly detected and extracted as part of a CRT. "Splits" errors occurred when a CRT was incorrectly split into two parts, resulting in a false additional CRT being extracted. "Noise" errors occurred when CRTs had noisy data. (Noisy data is data that negatively affects data processing techniques (Quinlan, 1986).)

Table 6 provides the extraction accuracy results for the CRTs' attributes (title, date, journal, and action). As shown in the table 27% (n = 196) of CRT titles were perfectly extracted, while 48% (n = 353) were partially extracted, and 14% (n = 99) were not extracted. In respect of CRT dates, 83% (n = 604) were perfectly extracted, none was partially extracted, and 6% (n = 44) were not extracted. In respect of CRT journals, 72% (n = 531) were perfectly extracted, fewer than 1% (n = 4) were partially extracted, and 15% (n = 113) were not extracted. In respect of CRT actions, 40% (n = 292) were perfectly extracted, fewer than 1% (n = 3) were partially extracted, and 48% (n = 353) were not extracted.

Table 6: Extraction accuracy for CRT attributes

Attribute	Category	Frequency (n = 731)
CRT title	Perfectly extracted CRT titles	196
	Partially extracted CRT titles	353
	Titles not extracted	99
	Extra line instances	25
	Split instances	19
	Noisy instances	39
CRT date	Perfectly extracted CRT dates	604
	Partially extracted CRT dates	0
	CRT dates not extracted	44
	Extra line instances	25
	Split instances	19
	Noisy instances	39
CRT journal	Perfectly extracted CRT journals	531
	Partially extracted CRT journals	4
	CRT journals not extracted	113
	Extra line instances	25
	Split instances	19
	Noisy instances	39
CRT action	Perfectly extracted CRT actions	292
	Partially extracted CRT actions	3
	CRT actions not extracted	353
	Extra line instances	25
	Split instances	19
	Noise instances	39

None of the 50 extractions could earn an overall "perfect" categorisation (see Table 5 above) because each extraction had at least one CRT attribute error, i.e., an error in extraction of one or more of the CRT's title, date, journal, or action.

Table 7 summarises the difference ratios for the perfectly extracted attributes. For CRT titles, a difference ratio of 0.73 was observed, indicating that 73% of the extracted CRT titles were different from the actual titles. For CRT dates, 17% of

AJIC 25 - 5-July-20 - 20h14.indd 14-15 7/5/2020 8:22:50 PM

the extractions were different from the actual dates; for CRT journals, 27% of the extractions were different from the actual journals; and for CRT actions, 60% were different from the actual actions.

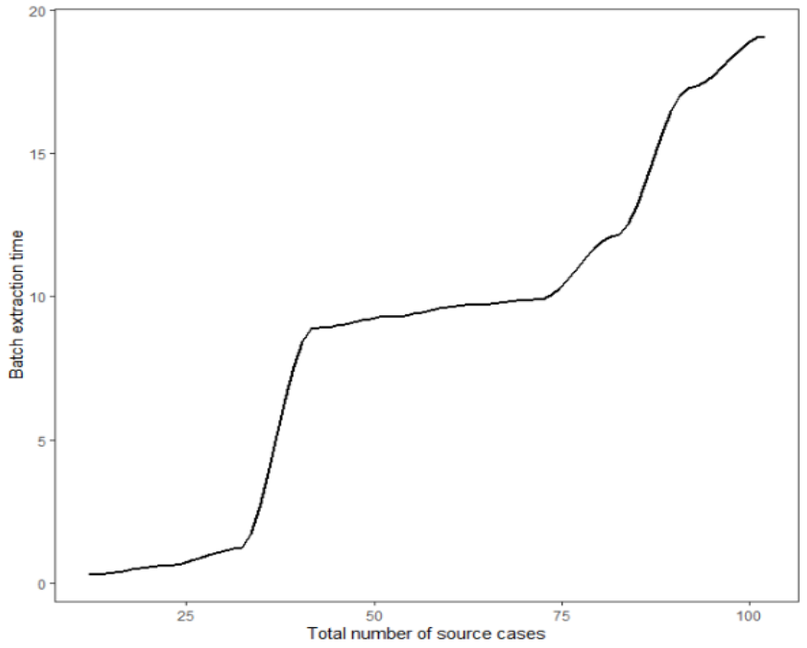
Table 7: Difference ratios for perfectly extracted CRT attributes

Attribute	Frequency	Difference ratio
Titles	196	0.73
Number of CRT dates	604	0.17
Number of CRT journals	531	0.27
Number of CRT actions	292	0.60

Execution times

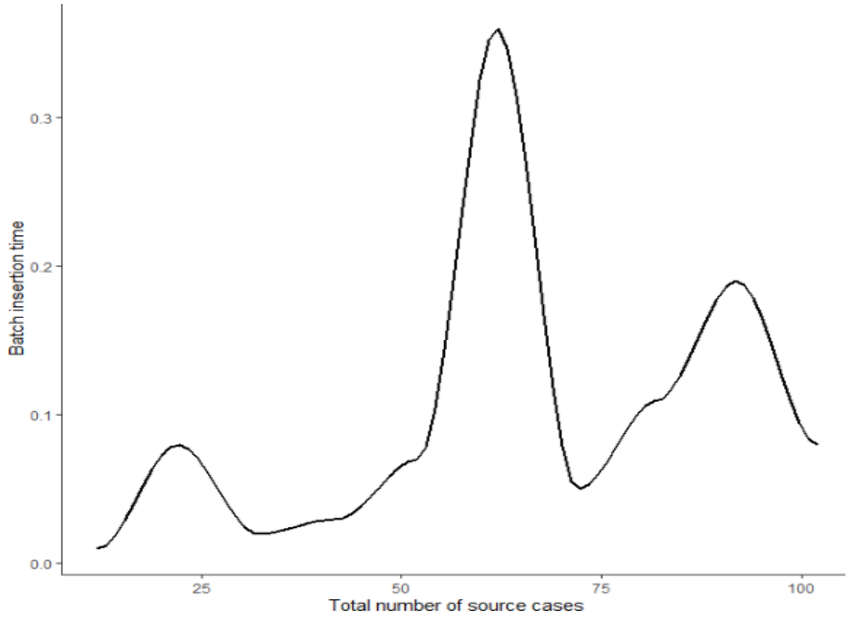
In the second part of experiment set 3, aimed at evaluating the prototype's information storage performance, we evaluated the prototype's execution times. Two metrics for execution time were used, namely extraction time and insertion time (the time taken to insert the source cases in the batch into the database as objects). Ten batches of tests were run, and the extraction and insertion results are illustrated in Figure 2 and Figure 3, respectively. (The extraction process we timed was for the extraction of source cases, not for the extraction of CRTs.) The first test batch had 12 case documents. In each successive test, an additional 10 cases were added, until reaching the last batch of 102 case documents.

Figure 2: Time taken to extract source cases



For the 10 batches, an average time of 8.8 seconds per batch was taken to extract the source cases. The batch extraction time was found to be satisfactory, at 1.2 seconds, for the batch of 32 cases, but the batch extraction time jumped to 8.9 seconds for the 42-case batch. A likely reason for this marked increase in extraction processing time was the increased word counts of the case documents.

Figure 3: Time taken to insert case objects



The insertion time was calculated in terms of the time it took to insert the case objects, as key-value pairs, into the document database. The insertions were found to be quick, taking an average of 0.10 seconds per source case object (n = 102). But no pattern could be found for the insertion times, as the times were not found to increase accumulatively as the size of the batch increased. The reasons for this were not clear, and could be investigated in future research. The longest time taken to insert case objects was for the batch of 62 case objects, which took 0.36 seconds per case. The shortest time taken was 0.01 seconds per case (n = 12).

5. Conclusions

The pilot testing described in this article, of a prototype for two components—the IE process and the information storage process—in our proposed IE Model for Legal Cases, highlights the challenges that need to be addressed in order to accurately and efficiently extract and store CRTs. Testing of our prototype's CRT extraction from 50 source cases resulted in only 38% (n = 19) of the extractions providing an accurate

AJIC 25 - 5-July-20 - 20h14.indd 16-17 7/5/2020 8:22:50 PM

number of CRTs. None of the 50 extractions resulted in fully accurate extractions of the CRT title, date, journal, and action attributes. It is thus clear that the IE prototype needs to be improved. A key area requiring improvement is the prototype's ability to extract CRT information from a wider variety of case formats, in order to reduce data noise and, in turn, the number of errors caused by the splitting of a single CRT's information into more than one CRT extraction, and the number of errors caused by the addition of non-CRT-related text into a CRT's extraction.

References

Abdelmagid, M., Ahmed, A., & Himmat, M. (2015). Information extraction methods and extraction techniques in the chemical document's contents: Survey. *ARPN Journal of Engineering and Applied Sciences, 10*(3), 1068–1073.

Al-Anzi, F. S., & AbuZeina, D. (2018). Beyond vector space model for hierarchical Arabic text classification: A Markov chain approach. *Information Processing & Management, 54*(1), 105–115. https://doi.org/10.1016/j.ipm.2017.10.003

Aritomo, D., & Watanabe, C. (2019). Achieving efficient similar document search over encrypted data on the cloud. In *2019 IEEE International Conference on Smart Computing (SMARTCOMP)* (pp. 1–6). https://doi.org/10.1109/smartcomp.2019.00020

Batra, S., & Tyagi, C. (2012). Comparative analysis of relational and graph databases. *International Journal of Soft Computing and Engineering (IJSCE)*, *2*(2), 509–512.

Chopra, A., Prashar, A., & Sain, C. (2013). Natural language processing. *International Journal of Technology Enhancements and Emerging Engineering Research*, *1*(4), 131–134.

Chowdhary, K. (2012). *Natural language processing*. Jodhpur, India: MBM Engineering College. Retrieved from http://www.krchowdhary.com/me-nlp12/nlp-01.pdf

Conroy, R. (2016). *Sample size: A rough guide.* Dublin: Royal College of Surgeons in Ireland. http://doi.org/10.1080/08897077.2011.640215

Collobert, R., Weston, J., Bottou, L., Karlen, M., Kavukcuoglu, K., & Kuksa, P. (2011). Natural language processing (almost) from scratch. *Journal of Machine Learning Research, 12,* 2493–2537. https://doi.org/10.1.1.231.4614

Croft, W. B., Metzler, D., & Strohman, T. (2015). *Information retrieval in practice*. New York: Pearson.

Firdhous, M. (2010). Automating legal research through data mining. *International Journal of Advanced Computer Science and Applications (IJACSA)*, *1*(6), 9–16.

Goyvaerts, J., & Levithan, S. (2009). *Regular expressions cookbook*. Boston: O'Reilly Media.

Iida, R., Inui, K., Takamura, H., & Matsumoto, Y. (2003). Incorporating contextual cues in trainable models for coreference resolution. In *Proceedings of the 2003 EACL Workshop on the Computational Treatment of Anaphora* (pp. 23–30).

Indurkhya, N., & Damerau, F. J. (Eds.).(2010). *Handbook of natural language processing (Vol. 2)*. Boca Raton, FL: CRC Press.

Kumar, R., & Sharma, S. C. (2018). Information retrieval system: An overview, issues, and challenges. *International Journal of Technology Diffusion (IJTD)*, *9*(1), 1–10.

LexisNexis South Africa. (n.d.). *All South African law reports*. Retrieved from https://store.lexisnexis.co.za/products/all-south-african-law-reports-2020-skuZASKUPG1994

Liddy, E. D. (2001). Natural language processing. In M. A. Drake (Ed.), *Encyclopedia of library and information science* (2nd ed.). New York: Marcel Dekker. https://doi.org/10.1017/S0267190500001446

Liu, B. (2011). *Web data mining: Exploring hyperlinks, contents, and usage data*. New York: Springer-Verlag.

Losee, R. M. (2015). Validating a model predicting retrieval ordering performance with statistically dependent binary features. *International Journal of Information Retrieval Research (IJIRR)*, *5*(1), 1–18. https://doi.org/10.4018/ijirr.2015010101

Losee, R. M., Bookstein, A., & Yu, C. T. (1986). Probabilistic models for document retrieval: A comparison of performance on experimental and synthetic databases. In *SIGIR '86: Proceedings of the 9th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 258–264). https://doi.org/10.1145/253168.253222

MongoDB. (n.d.). Introduction to MongoDB: Documents. Retrieved from https://docs.mongodb.com/manual/core/document/

Moniruzzaman, A., & Hossain, S. (2013). NoSQL database: New era of databases for big data analytics – Classification, characteristics and comparison. *International Journal of Database Theory and Application*, *6*(4), 43–45. https://doi.org/10.1016/S0262-4079(12)63205-9

Mooney, R. J., & Bunescu, R. (2005). Mining knowledge from text using information extraction. *ACM SIGKDD Explorations Newsletter*, *7*(1), 3–10. https://doi.org/10.1145/1089815.1089817

Padayachy, T., Scholtz, B., & Wesson, J. (2018). An information extraction model using a graph database to recommend the most applied case. In *Proceedings of the 2018 International Conference on Computing, Electronics and Communications Engineering (ICCECE)* (pp. 89–94). doi: 10.1109/iCCECOME.2018.8658659

Pandey, S., Mathur, I., & Joshi, N. (2019). Information retrieval ranking using machine learning techniques. In *2019 Amity International Conference on Artificial Intelligence (AICAI)* (pp. 86–92). https://doi: 10.1109/AICAI.2019.8701391

Piskorski, J., & Yangarber, R. (2013). Information extraction: Past, present and future. In T. Poibeau, H. Saggion, J. Piskorski, & R. Yangarber (Eds.), *Multi-source, multilingual information extraction and summarization*. Berlin: Springer. https://doi.org/10.1007/978-3-642-28569-1_2

Quinlan, J. R. (1986). The effect of noise on concept learning. In R. S. I. Michalski, J. G. Carboneel, & T. M. Mitchell (Eds.), *Machine learning*. Burlington, MA: Morgan Kaufmann Publishers.

Roshdi, A., & Roohparvar, A. (2015). Review: Information retrieval techniques and applications. *International Journal of Computer Networks and Communications Security*, *3*(9), 373–377.

Roy-Hubara, N., & Sturm, A. (2020). Design methods for the new database era: A systematic literature review. *Software & Systems Modeling*, *19*, 297–312. https://doi.org/10.1007/s10270-019-00739-8

Schofield, A., Magnusson, M., & Mimno, D. (2017). Pulling out the stops: Rethinking stopword removal for topic models. In *15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers* (pp. 432–436). https://doi.org/10.18653/v1/e17-2069

AJIC 25 - 5-July-20 - 20h14.indd 18-19 7/5/2020 8:22:50 PM

Singh, S. (2018). Natural language processing for information extraction. arXiv preprint arXiv:1807.02383.

Vijayarani, S., Ilamathi, M. J., & Nithya, M. (2015). Preprocessing techniques for text mining – An overview. *International Journal of Computer Science & Communication Networks*, *5*(1), 7–16.

Cyber-Threat Information-Sharing Standards: A Review of Evaluation Literature

Nenekazi N. P. Mkuzangwe
Researcher, Defence and Security, Council for Scientific and Industrial Research (CSIR), Pretoria
https://orcid.org/0000-0002-2073-4082

Zubeida C. Khan
Senior Researcher, Defence and Security, Council for Scientific and Industrial Research (CSIR), Pretoria
https://orcid.org/0000-0002-1081-9322

Abstract
Cyber-threat information-sharing tools, through which cybersecurity teams share threat information, are essential to combatting today's increasingly frequent and sophisticated cyber-attacks. Several cyber-threat information-sharing standards exist, but there is at present no single standard or set of standards widely adopted by organisations and by computer security incident response teams (CSIRTs) operating at organisational, sectoral, national, and international levels. This introduces an interoperability problem in respect of communication across the various organisations and CSIRTs. Harmonised adoption of threat information-sharing standards would be of great benefit to cybersecurity efforts. In an effort to support harmonised use of cyber-threat information-sharing standards, this article provides findings from a review of the extant literature on such standards.

Keywords
cybersecurity, cyber-threats, information-sharing, standards, protocols, harmonisation, computer security incident response teams (CSIRTs), ontologies, use cases, semantic elements, syntactic elements, privacy, information security

DOI: https://doi.org/10.23962/10539/29191

Recommended citation
Mkuzangwe, N. N. P., & Khan, Z. C. (2020). Cyber-threat information-sharing standards: A review of evaluation literature. *The African Journal of Information and Communication (AJIC), 25*, 1-12. https://doi.org/10.23962/10539/29191

This article is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) licence: https://creativecommons.org/licenses/by/4.0

AJIC 25 - 5-July-20 - 20h14.indd 20-1 7/5/2020 8:22:50 PM

1. Introduction

Cyber-attacks are increasing in both frequency and sophistication, and no organisation is immune from attack. It thus becomes imperative for organisations to have mechanisms that will help improve their security and ability to defend against cybercrime—and, in turn, decrease their risks of suffering financial and/or reputational damage. Information-sharing about the various cyber-threats, vulnerabilities, and other malicious cyber-artefacts is one of the mechanisms used to help fight against the ever-growing and increasingly sophisticated cyber-attacks.

Key information-sharing entities at organisational, sector, national, and international levels are computer security incident response teams (CSIRTs), which are staffed by professionals performing both reactive and proactive services. These services include information-sharing, threat-sharing, incident-handling, and proactive threat intelligence. An example of an international CSIRT is the European Union Agency for Cybersecurity (known by the acronym ENISA, based on the agency's original name). CSIRTs typically have cyber-threat information-sharing standards and protocols in place, such as structured threat information expression (STIX), trusted automated exchange of intelligence information (TAXII), and cyber observable expression (CybOX). However, there are numerous available standards and protocols, and CSIRTs and organisations use varying standards, depending on their particular preferences. There is, thus, currently no widely adopted international set of standards for security teams' sharing of cyber-threat information. This lack of adoption of common standards can serve to undermine effective communication regarding cyber threats between organisations, and between organisations and sector, national and international CSIRTs.

According to Johnson et al. (2016), standardised data formats and transport protocols are important building blocks for interoperability, as they enable automation and allow information-sharing amongst organisations to occur at machine speed. Rantos et al. (2020) state that interoperability issues need to be addressed before any sharing of cyber-threat information and intelligence may occur, and they delineate the issues into the following categories: legal; policy and procedural; technical; and semantic and syntactic. Legal interoperability ensures the alignment of legal frameworks under which organisations operate and provide services, and it also caters to matters of data privacy. The policies and procedures for interoperability consist of formal statements that reflect organisations' objectives and detailed instructions to achieve these objectives. Technical interoperability relates to the implementation of tools that support the automated exchange of information (including delivery and consumption) and the underlying communication protocols used for the transport of the information. Semantic and syntactic interoperability involves conveying the necessary meaning via syntactically correct messages.

Several data formats (standards) are currently used for the exchange of cyber-threat information between entities, which is a situation that poses interoperability problems. Harmonised adoption of threat information-sharing standards is necessary for optimal cybersecurity. In an effort to support dialogue on harmonisation, this article provides findings from our review of literature evaluating existing cyber-threat information-sharing standards. The literature review was primarily conducted by means of keyword searches, using the following search string: "cyber threat information sharing standard", "cyber threat intelligence", and "cyber threat intelligence ontology". Initially Google Scholar was searched. Thereafter, academic databases such as Science Direct and IEEE Xplore Digital Library were searched. Citations and websites were gleaned from the collected literature and visited in order to augment the collection. It was found that the literature evaluating standards could be divided into four categories:

- ontologies;
- use cases;
- semantic and syntactic elements; and
- privacy and information security implications.

2. Ontologies

A proposed taxonomy for threat-sharing technologies and ontologies

Burger et al. (2014) propose a five-layer taxonomy for classifying threat-sharing technologies and for classifying ontologies of such technologies. The five layers proposed are: *transport*, *session*, *indicators*, *intelligence,* and *5Ws* (who, what, when, where, and why). The transport and session layers often ride over the hypertext transport protocol/transport layer security (HTTP/TLS). On the transport layer, the TLS is responsible for the encryption of the byte stream, which can be synchronous or asynchronous in order to ensure the confidentiality and integrity of the raw data (payload) that is being transported. The session layer is responsible for authentication and authorisation of users by defining the way in which users are authenticated to the system and what threat data they can access. The indicators layer represents cyber-intelligence payload and indicators. The intelligence layer specifies action and includes queries that are formulated from information gathered from the indicator layer about a target or targets. The 5Ws layer is used to gather information using "who", "what", "when", "where" and "why" questions, e.g., who is interested in the user or organisation?

Burger et al. (2014) use their model to evaluate two transport protocols—TAXII, and real-time internet-work defence (RID)—and two data representation formats: STIX, and versions of incident object description exchange format (IODEF). TAXII

AJIC 25 - 5-July-20 - 20h14.indd 2-3 7/5/2020 8:22:50 PM

and RID are transport protocols for STIX and IODEF, respectively. Burger et al. (2014) found, via their evaluation, that:

- With respect to the transport protocol standards, TAXII and RID are in the *transport* and *session* layers since they provide secure transportation of cyber-threat intelligence. RID ensures that the client and the server are authenticated to each other, that the payload (the actual cyber-threat intelligence being transported) is encrypted (RID is primarily used over HTTP/TLS network protocol), and that privacy between partners is enforced. TAXII also encrypts the cyber-threat intelligence that is being transported, and authenticates users through network protocols like HTTP/S.
- With respect to the two data representation standards, STIX falls in the *indicators*, *intelligence* and *5Ws* layers, due to its wide range of objects. STIX's objects give it the ability to represent indicators, specific actions to be carried out, and the 5Ws. IODEF falls in the *indicators* layer since it consists of a data model that provides an XML representation of threat information shared amongst CSIRTs in relation to computer security incidents and events. IODEF for structured cybersecurity information (IODEF-SCI) falls in the *indicators* and *intelligence* layers since it extends IODEF to carry intelligence information. When deployed in conjunction with RID, IODEF can fall in the *intelligence* layer since it can be used as a query language.

Ontologies for semantic reasoning services

Asgarli and Burger (2016) analyse STIX and IODEF in order to map them to RDF/XML and to propose ontologies for semantic reasoning services. Semantic reasoning is the process of inferring new knowledge from an existing knowledge base using logical rules. The benefit in having an ontology is that, in addition to using it for threat intelligence, it can go a step further towards a more strategic approach that enables the system to make inferences about potential cyber-attacks in an effort towards automated response. Ontologies are created by entities (classes, object properties, and data properties) that are used to represent a domain. The Asgarli and Burger (2016) mapping process for STIX and IODEF results in an ontology for STIX containing 153 classes, 237 object properties, and 49 data properties. The resulting IODEF ontology contains 39 classes, 45 object properties, and 54 data properties. The STIX ontology is considerably larger because STIX contains definitions from CybOX, from common attack pattern enumeration and classification (CAPEC), and from malware attribute enumeration and characterisation (MAEC) standards.

Towards a comprehensive threat intelligence ontology

Mavroeidis and Bromander (2017) present a cyber-threat intelligence (CTI) model to characterise threat intelligence in terms of various dimensions. This model can be used for potential attack attribution. The model is characterised by *detective capabilities* and *preventive capabilities*. The authors use the CTI model to compare a set of 27 cyber-threat standards, taxonomies, and ontologies. They find that only two

standards and two ontologies contain comprehensive threat information, according to their CTI model's characteristics, namely: the STIX1 and STIX2 standards, and the two unified cybersecurity ontology (UCO) standards. The authors find that other existing ontologies are not sufficiently comprehensive for use in representing information about cyber-threat intelligence, i.e., they lack formal constraints, which are used in ontologies to provide more specialised information about concepts such as cardinality (e.g., specifying that a certain threat has exactly one actor). Mavroeidis and Bromander (2017) also find that the ontologies target specific sub-domains of threat intelligence, and thus cannot be used for a wide range of cyber-threats.

In respect of the taxonomies they examine, Mavroeidis and Bromander (2017) find that relationships are not sufficiently established, in the taxonomies, between the motivations, goals, and strategies of the attackers— meaning the taxonomies are not sufficient for use in sharing information about cyber-threats. Mavroeidis and Bromander (2017) propose that the way forward is to develop a heavyweight ontology (one enriched with logical axioms describing concepts in details) for cyber-threats, so that information is represented in a uniform and logical format with the high degree of expressivity necessary for complex cyber-threat information.

3. Use cases

Overview of 22 standards' use cases

Kampanakis (2014) examines 22 standards in terms of the following: *each standard's purposes*, *other similar standards*, *where the standard is used*, and *its adoption level.*

Table: Standards examined in Kampanakis (2014)

Language standards	Transport standards	Scoring systems standards	Enumeration standards	Other
STIX	TAXII	common vulnerability scoring system (CVSS)	common platform enumeration (CPE)	software identification (SWID)
CybOX	RID	common configuration scoring system (CCSS)	common vulnerability enumeration (CVE)	
MAEC		common weakness scoring system (CWSS)	common configuration enumeration (CCE)	

AJIC 25 - 5-July-20 - 20h14.indd 4-5 7/5/2020 8:22:50 PM

open vulnerability and assessment language (OVAL)			common weakness enumeration (CWE)	
extensible configuration checklist description format (XCCDF)				
open checklist interactive language (OCIL)				
IODEF				
malware metadata exchange format (MMDEF)				
common vulnerability report format (CVRF)				
open indicators of compromise (OpenIOC)				
vocabulary for event recording and incident scoring (VERIS)				

Kampanakis (2014) finds that many of the standards overlap, and the choice of which standard to use depends on the context and use case. Accordingly, Kampanakis (2014) recommends that the first step should be to identify the use case of the security information to be represented and exchanged, followed by selection of the standard that covers that specific use case.

Incident reporting formats' strengths, weaknesses, use cases

Menges and Pernul (2018) propose a three-pronged model for evaluating incident reporting formats, based on the *structural*, *general*, and *additional* evaluation criteria. The *structural* evaluation criteria are based on a model that the authors call a universal pattern for structured incident exchange (UPSIDE). These criteria evaluate

the incident reporting formats in terms of indicator, attacker, attack, defender, and contentual coverage. The *general* evaluation criteria, based on those proposed by Steinberger et al. (2015), evaluate the formats according to machine-readability, human-readability, unambiguousness of semantics, interoperability, extensibility, aggregability, practical application, and external dependencies. The *additional* evaluation criteria are licensing terms, maintenance effort, and documentation. Menges and Pernul (2018) apply this framework to four incident reporting formats:

- STIX versions 1 and 2;
- IODEF and IODEF version 2;
- VERIS; and
- extended abuse reporting format (X-ARF).

They find that, in terms of the structural/UPSIDE evaluation criteria, STIX and STIX 2 are able to represent the indicator, attacker, attack, defender, and contentual coverage specified in UPSIDE, while IODEF does not provide indicator, attacker, or defender coverage. IODEF also does not provide sufficient contentual coverage. IODEF 2 extends the first version of IODEF in being able to represent the attacker and defender entities of UPSIDE, and in having increased contentual coverage. VERIS represents the attacker, defender and attack entities of UPSIDE, with less contextual coverage than the two STIX and two IODEF versions. X-ARF provides attacker and attack coverage but no indicator or defender coverage, and the lowest contentual coverage among the reporting formats considered.

In terms of the general evaluation criteria, Menges and Pernul (2018) find that both versions of STIX meet most of the general evaluation criteria, except for (in the case of STIX 1) human readability and extensibility, and (in the case of STIX 2) human readability and practical application. IODEF is found to have high interoperability, extensibility, human readability, aggregability and practical application, but low machine readability, ambiguity problems, and no external dependencies. IODEF 2 is found to have improved (over the first version of IODEF) via better machine readability, changes to prevent ambiguity, and use of external references, but with poor human readability. The interoperability and extensibility of IODEF 2 are similar to that of the original IODEF. VERIS has adequate machine readability and human readability, and good interoperability and extensibility, but no aggregability, external dependencies, or wide practical application. Meanwhile, X-ARF is found to score well in human readability, poorly in machine readability and ambiguity, and to lack aggregability or external dependencies.

In terms of the additional evaluation criteria, Menges and Pernul (2018) find that all the examined standards are licensed, maintained, and documented. However, for STIX 1 and IODEF, maintenance effort has fallen away with the introduction of the new versions.

AJIC 25 - 5-July-20 - 20h14.indd 6-7 7/5/2020 8:22:50 PM

4. Semantic and syntactic elements

Fenz et al. (2008) provide a framework for evaluating the semantic elements of security advisory standards in terms of their *semantic usability*, *information complexity*, and *distribution*. In respect of semantic usability, Fenz et al. (2008) analyse the degree to which a standard uses a common language to ensure machine readability, and the degree to which it provides clear and unambiguous semantics to ensure machine recognition. In terms of information complexity, the authors analyse the extent to which the standard provides necessary elements for describing information technology (IT) security incidents. In respect of distribution, they analyse the degree to which the standard is used by major CSIRTs, whether it is still supported, and the last time the standard has been updated.

Fenz et al. (2008) use their framework to evaluate the following six standards:

- advisory and notification markup language (ANML);
- European information security promotion programme (EISPP);
- common announcement interchange format (CAIF);
- IODEF;
- common alerting protocol (CAP); and
- OVAL.

Fenz et al. (2008) find that, in terms of *semantic usability*, OVAL is the only standard of the six that met the elements of this criterion fully; in terms of *information complexity*, four of the six standards—ANML, EISPP, CAIF and OVAL—satisfy the criterion; and in terms of *distribution*, four of the six standards—EISPP, IODEF, CAP, and OVAL—are satisfactory. Fenz et al. (2008) thus deduce that OVAL is the most suitable of the six standards for automatic or semi-automatic interpretation of security threats, though they do at the same time find that OVAL falls short on some requirements, e.g., patch information such as download locations or required reboots.

Steinberger et al. (2015) evaluate standards in terms of their *interoperability*; *extensibility; scalability*; *aggregability*; *protocol independency*; *human readability*; *machine readability*; *confidentiality, integrity and authenticity*; and *practical application* and *reliable message transport* for exchange protocols. They apply their evaluation framework to the following exchange formats:

- IODEF;
- CAIF;
- ARF;
- X-ARF versions 1 and 2;
- common event expression (CEE); and
- Syslog RFC 5424.

Steinberger et al. (2015) also evaluate two exchange (transport) protocols—RID and common intrusion detection framework (CIDF)—and the following extensible messaging and presence protocols:

- XMPP extension protocol;
- incident handling protocol XEP-0268;
- intrusion detection exchange protocol (IDXP);
- simple mail transfer protocol (SMTP);
- common event expression (CEE) log transport (CLT); and
- Syslog RFC 3164 and RFC 5425.

For the exchange formats they have considered, the authors find that *extensibility* is high for all the exchange formats, while the *confidentiality, integrity and authenticity* and *practical application* criteria are not well-satisfied by any of the exchange formats.

ARF, CEE and both versions of X_ARF are found to be high on *interoperability*. All the exchange formats they have considered were low on *scalability*, while *aggregability* is found to be high in CAIF and X_ARF v0.2. *Protocol independency* is found to be high in CAIF, ARF and both versions of X_ARF. ARF, CEE and both versions of X_ARF are found to be high in *human readability*, while *computer readability* is high in all the exchange formats considered. The authors did not report the evaluation of Syslog RFC 5424 exchange format.

For the exchange protocols and the extensible messaging and presence protocols they have considered, Steinberger et al. (2015) found that *reliable message transport* and *scalability* are high for all the protocols, except for Syslog RFC 3164 and RID. *Confidentiality, integrity and authenticity* are high for all the considered protocols, except for SMTP and RFC 3164. *Interoperability* is high for all the considered protocols, except for CIDF, XEP-0268 and CLT. The rest of the protocols are low (CIDF, XEP-0268 and CLT) or medium (RID and IDXP) gradings in terms of *practical application*.

5. Privacy and information security implications

Information-sharing may result in leaking of the private information of entities, or revealing sensitive information about the context (since attributing attacks and performing various security analyses require contextual information) (Kampanakis, 2014). Disclosure of sensitive information, and personally identifiable information (PII) can result in, inter alia, financial loss and loss of reputation (Johnson et al., 2016). Therefore, it is important to evaluate the sharing standards in terms of how much private information they leak about the sharing entities.

Kampanakis (2014), based on the review of 22 standards discussed in section 3 above, advises that cyber-threat information collection and sharing be done in a systematic manner in order to mitigate privacy risks, and points to NIST's Preliminary Cybersecurity Framework Appendix B as a methodology for the protection of privacy and civil liberties within a cybersecurity programme (NIST, 2013).

AJIC 25 - 5-July-20 - 20h14.indd 8-9 7/5/2020 8:22:50 PM

Mohaisen et al. (2017), in their exploration of the privacy risks associated with threat intelligence information-sharing, include an analysis of the private information-leaking risks posed by 14 widely-used information-sharing standards in three categories:

- enumeration standards: CCE, CWE, CVE, CPE and common attack pattern enumeration and classification (CAPEC);
- scoring systems standards: CVSS and CWSS; and
- language standards: CybOX, MAEC, OVAL, IODEF, XCCDF, STIX and CEE.

For the 14 standards, Mohaisen et al. (2017) apply the following information-leakage scoring system:

- 0 for non-leaked or public data;
- 1 for leaked inferential data;
- 2 for leaked sensitive data; and
- 4 for leaked PII data.

Mohaisen et al. (2017) find that the language standards have the highest overall scores, with CybOX having the highest score of 65 and STIX the second-highest score of 36. The language standards also leak the most PII data. Thus, adoption of CybOX and STIX require the deployment of supplementary privacy controls.

Albakri et al. (2018) provide an analysis of the information that is shared by STIX, determining which information is contained in the incident reports and the risks associated with leaking such information. For every STIX data field, the threat associated with the disclosure of the information that corresponds to the data field is identified and its severity evaluated. The authors also evaluate the extent to which the disclosure of the information that corresponds with the data field identifies an individual or an organisation. The study provides detailed understanding of which information in the cyber-incident reports needs to be protected against specific attacks, and of the potential severity of such attacks. The authors aim to derive a set of guidelines on how to use STIX in a disciplined way that reduces the information-security risks. Their analysis indicates that certain STIX data fields can leak PII, organisational information, financial information, or cybersecurity information—largely because STIX consists of many free text fields with unconstrained properties.

To avoid information leakage via these fields, the authors advise the use of templates and that organisations use customised versions of STIX that meet their specific risk profiles.

6. Conclusions

This study reviewed cyber-threat information standards to assist with addressing interoperability issues in cyber-threat information-sharing. From the reviewed literature, eight reporting formats—namely, STIX 1, STIX 2, IODEF, IODEF 2, VERIS, ARF, CEE and X_ARF—and one exchange protocol, RID, were identified as being able to facilitate interoperability. However, from the studies that examined the privacy implications of the standards, the language standards CybOX, MAEC, OVAL, IODEF, XCCDF, STIX and CEE were found to leak the most private information, followed by the enumeration standards CCE, CWE, CVE, CPE and CAPEC, while the scoring standards CVSS and CWSS were found to leak no private information.

As pointed out in the literature, the leaking of private information violates legal interoperability and needs to be addressed before any information-sharing can occur. The works reviewed also suggest that, before adopting a standard, the use cases applicable to the incidents to be reported must be determined, and the standard that is capable of handling such use cases can then be selected.

References

Albakri, A., Boiten, E., & De Lemos, R. (2018). Risks of sharing cyber incident information. In *Proceedings of the 13th International Conference on Availability, Reliability and Security* (pp. 1–10). Association for Computing Machinery (ACM). https://doi.org/10.1145/3230833.3233284

Asgarli, E., & Burger, E. (2016). Semantic ontologies for cyber threat sharing standards. In *2016 IEEE Symposium on Technologies for Homeland Security (HST)* (pp. 1–6). Institute of Electrical and Electronics Engineers (IEEE). https://doi.org/10.1109/ths.2016.7568896

Burger, E. W., Goodman, M. D., Kampanakis, P., & Zhu, K. A. (2014). Taxonomy model for cyber threat intelligence information exchange technologies. In *Proceedings of the 2014 ACM Workshop on Information Sharing & Collaborative Security* (pp. 51–60). Association for Computing Machinery (ACM). https://doi.org/10.1145/2663876.2663883

Fenz, S., Ekelhart, A., & Weippl, E. (2008). Semantic potential of existing security advisory standards. In *Proceedings of the FIRST 2008 Conference-Forum of Incident Response and Security Teams*. https://doi.org/10.1109/aina.2008.69

Johnson, C., Badger, L., Waltermire, D., Snyder, J., & Skorupka, C. (2016). *Guide to cyber threat information sharing*. NIST Special Publication 800-150. https://doi.org/10.6028/nist.sp.800-150

Kampanakis, P. (2014). Security automation and threat information-sharing options. *IEEE Security & Privacy, 12*(5), 42–51. https://doi.org/10.1109/msp.2014.99

Mavroeidis, V., & Bromander, S. (2017). Cyber threat intelligence model: An evaluation of taxonomies, sharing standards, and ontologies within cyber threat intelligence. In *Intelligence and Security Informatics Conference (EISIC), 2017 European* (pp. 91–98). Institute of Electrical and Electronics Engineers (IEEE). https://doi.org/10.1109/eisic.2017.20

AJIC 25 - 5-July-20 - 20h14.indd 10-11 7/5/2020 8:22:50 PM

Menges, F., & Pernul, G. (2018). A comparative analysis of incident reporting formats. *Computers & Security*, *73*(March), 87–101. https://doi.org/10.1016/j.cose.2017.10.009

Mohaisen, A., Al-Ibrahim, O., Kamhoua, C., Kwiat, K., & Njilla, L. (2017, October). Rethinking information sharing for threat intelligence. In *Proceedings of the Fifth ACM/IEEE Workshop on Hot Topics in Web Systems and Technologies* (pp. 1–7). Association for Computing Machinery (ACM). https://doi.org/10.1145/3132465.3132468

NIST. (2013). Improving Critical Infrastructure Cybersecurity Executive Order 13636: Preliminary Cybersecurity Framework. https://doi.org/10.1002/9781119369141.app3

Rantos, K., Spyros, A., Papanikolaou, A., Kritsas, A., Ilioudis, C., & Katos, V. (2020). Interoperability challenges in the cybersecurity information sharing ecosystem. *Computers*, *9*(1), 18. https://doi.org/10.3390/computers9010018

Steinberger, J., Sperotto, A., Golling, M., & Baier, H. (2015). How to exchange security events? Overview and evaluation of formats and protocols. In *Integrated Network Management (IM), 2015 IFIP/IEEE International Symposium on Integrated Network Management* (pp. 261–269). Institute of Electrical and Electronics Engineers (IEEE). https://doi.org/10.1109/inm.2015.7140300

PUBLICATION REVIEW

AJIC 25 - 5-July-20 - 20h14.indd 12-1 7/5/2020 8:22:51 PM

Book Review: International Telecommunications Law and Policy

Uchenna Jerome Orji, International Telecommunications Law and Policy. *Newcastle upon Tyne, UK: Cambridge Scholars Publishing, 2018, 425 pages, £74.99 (hardcover). ISBN-13: 978-1-5275-0836-1; ISBN-10: 1-5275-0836-6*

Reviewer: Pontian N. Okoli
Lecturer, School of Law, University of Stirling, Scotland
https://orcid.org/0000-0003-2704-4161

Keywords
telecommunications, law, policy, regulations, instituitions, history, international, Africa

DOI: https://doi.org/10.23962/10539/29190

Recommended citation
Okoli, P. N. (2020). Book review: International telecommunications law and policy. *The African Journal of Information and Communication (AJIC), 25*, 1-5.
https://doi.org/10.23962/10539/29190

This article is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) licence:
https://creativecommons.org/licenses/by/4.0

This comprehensive volume by Nigerian legal academic Orji engages with the international, African continental, and African regional regimes that govern telecommunications regulation. The international organisations covered are the United Nations (UN), the International Telecommunication Union (ITU), and the World Trade Organisation (WTO). The African continental organisations dealt with are the African Union (AU), the African Telecommunications Union (ATU), the UN Economic Commission for Africa (UNECA), and the Regional African Satellite Communication Organisation (RASCOM). African regional entities considered are the Economic Community of West African States (ECOWAS), the Common Market for Eastern and Southern Africa (COMESA), the Economic Community of Central African States (ECCAS), the Economic and Monetary Community of Central Africa (CEMAC), the East African Community (EAC), and the Southern African Development Community (SADC).

The first chapter traces the historical origins of telecommunications, the evolution of telecommunications, the implications of various technologies for law and policy, and the sources of international telecommunications law. This chapter provides an

AJIC 25 - 5-July-20 - 20h14.indd 2-1 7/5/2020 8:22:51 PM

important introductory context that facilitates engagement with the contents of the 10 chapters that follow.

The second chapter provides a discussion of international telecommunication regimes within the UN framework, including conventions on the law of the sea—e.g., the UN Convention on the Law of the Sea (UNCLOS), the Convention for the Protection of Telegraph Cables, and the Convention on the Continental Shelf—and on international space law (e.g., the Outer Space Treaty). In the context of UNCLOS, Orji argues for expansion of the convention's limited application of universal jurisdiction—such that all states linked to submarine fibre optic telecommunications cables would be required to establish adequate measures, including penal sanctions, to deter intentional cable damage in their territorial waters (pp. 37–39). This critical treatment of the provisions of UNCLOS is a good precursor to the author's discussion of the Outer Space Treaty (p. 45), which he critiques in terms of, inter alia, definitional inadequacies and the consequences of "orbital debris".

Chapters 3 and 4 provide important institutional insights. Chapter 3 discusses the ITU and the obligations of its Member States, including obligations concerning installation and operation of telecommunication or radio services. There is an illuminating discussion of the no-harm obligation in international law (pp. 73–74). The author argues that the no-harm obligation arises from "the ITU's early recognition of the interconnectedness of national telecommunication networks, and the equal rights of States to establish and operate telecommunications installations without interference from either State or non-State actors", and further links the no-harm rule to international law principles on state responsibility as recognised in the *Trail Smelter*[1] and *Corfu Channel*[2] cases.

Chapter 4 arguably provides even better engagement with institutional frameworks, through a discussion of the ITU's International Telecommunication Regulations (ITRs). In examining the proposals of various state parties during the review of the ITRs at the ITU's 2012 World Conference on International Telecommunications (WCIT), the author provides insights into the socio-political underpinnings of certain national policy approaches to international telecommunications. Orji also identifies key factors that influenced ITU Member States' proposals for ITR treatment of concerns over the United States' perceived dominance, at the time, of the multi-stakeholder internet governance structure (the non-governmental Internet Corporation for Assigned Names and Numbers (ICANN); the perceived internet governance agendas of countries such as China, Russia and the Arab states; and questions on the contemporary relevance of the ITRs for global telecommunications governance. Such factors, Orij contends, were to a large extent responsible for

1 *Trail Smelter Arbitration (United States v Canada)*, 3 UNRIAA, p. 1905, 1952.
2 *Corfu Channel Case (United Kingdom v Albania); Merits*, International Court of Justice (ICJ), 9 April 1949.

impeding consensus on the ITRs at the 2012 WCIT. The author points out that high levels of contestation are nothing new in international telecommunications negotiations, and that, from the 2012 ITU WCIT onwards, "[i]nternet governance issues would continue to remain a source of contention amongst ITU Member States" (p. 127). These views provide a helpful foundation for the next chapter.

In chapter 5, while analysing several provisions of the 2012 ITRs, Orji revisits some Member State arguments made at the 2012 WCIT, including US arguments regarding obligations on Member States to ensure the security of international telecommunication networks,[3] and regarding the inclusion of spam control provisions.[4] The US, citing security concerns, contended that the ITU and the ITRs were not the appropriate venue or framework, respectively. Orji provides practical insights into the political and international security influences on the US position, presenting a balanced scholarly analysis with a presentation of the author's critique as well as critiques by other scholars (pp. 149–154). In respect of cybersecurity, Orji argues that "the ITU may provide a broader platform that will accommodate the perceived interests of developing countries that are opposed to the Council of Europe's Convention on Cybercrime" (p. 151), thus offering the real possibility of attaining a global cybersecurity regime. However, the author argues there is a need to first address the dualism of the ITR framework whereby countries are bound by either the 2012 ITRs or the 1988 ITRs (i.e., if they have not accepted the former). Orji argues that the fragmenting effects of such dualism could be addressed if Member States were to use the ITRs' special arrangements regime to negotiate mutual bilateral or multilateral telecommunication agreements, thus mitigating the perceived effects of dualism on international business conditions and practices in the telecommunications industry.

Chapter 6 continues the exploration of global efforts at harmonised telecommunications law and policy. Orji analyses the basic principles of the International Radio Regulations and discusses the challenge posed by the over-filing of requests for orbital slots for satellite systems ("paper satellites"), arguing that such challenges mean that the Radio Regulations remain relevant (p. 179). He traces the history of the Radio Regulations to the effects of Marconi's patent monopoly on wireless telegraphy and the tragedy of the sinking of the Titanic. Arguably, according to the author, it took the sinking of the Titanic in 1912 to force reconsideration of the inability to reach consensus, at the first International Radiotelegraph Convention in Berlin in 1906, with respect to intercommunication of radio systems (p. 173).

In chapter 7, the author discusses the WTO's telecommunications regulatory regime within the context of the General Agreement on Trade in Services (GATS). The GATS, inspired by the General Agreement on Tariffs and Trade (GATT), came into

3 Art. 5A of the ITRs of 2012.
4 Art. 5B of the ITRs of 2012.

AJIC 25 - 5-July-20 - 20h14.indd 2-3 7/5/2020 8:22:51 PM

force in 1995 and has promoted trade and development partly through a liberalisation of relevant national trade laws and policies of Member States. However, international regimes such as the GATS do not always guarantee an application on national levels, as countries can also regulate the supply of services. In this context, the author examines the domestic frameworks of countries including the United States, the United Kingdom and Nigeria, in order to explore how the GATS and relevant WTO agreements may be given effect either directly (p. 226) or indirectly (p. 229).

Orji introduces a robust African perspective in chapter 8. In examining telecommunications policies within relevant legal and institutional frameworks, including the AU, he considers the challenges that must be overcome if Africa is to leverage its unprecedented growth in the penetration of mobile telecommunications. The AU Commission's initiatives on telecommunications regulatory harmonisation include the AU Division of Information Society; the Reference Framework for the Harmonization of Telecommunication and ICT Policies and Regulation in Africa; the Comprehensive Continental ICT Strategy for Africa; and the New Partnership for Africa's Development's (NEPAD) Protocol on High Level Policy and Regulatory Framework Broadband ICT Infrastructure for Eastern and Southern Africa (Kigali Protocol). In identifying the impediments to harmonisation in Africa, Orji points out that the sub-regional intergovernmental bodies do not necessarily support harmonisation. The author correctly points to the need for integration of the efforts of sub-regional institutions (p. 276), but could have gone further to examine how such integration could be attained, given that the AU is apparently unable to provide effective coordination of sub-regional efforts—a task made complicated by, inter alia, "diverse legal traditions" (p. 275). This potential gap does not detract from the persuasive arguments presented in favour of harmonisation, which factor in the challenges that exist and contextualise the current potential for attainment.

In chapter 9, the author discusses the ATU, the AU, RASCOM, and African regional telecommunication harmonisation initiatives such as the African Information Society Initiative and the African Regional Action Plan on the Knowledge Economy. He identifies the challenges faced by the ATU in respect of, *inter alia*, poor funding and the ATU's under-utilisation in the development and implementation of harmonisation initiatives. Orji recommends that the AU increase utilisation of the ATU for harmonisation efforts, and that the ATU improve awareness of its programmes and their implications for development on the continent.

Chapter 10 focuses on the ECOWAS telecommunication regimes. Orji analyses several aspects of the ECOWAS Telecommunications Package, including legal frameworks on interconnection of ICT networks; services; service providers; and the management of radio frequency spectrum. The author also gives consideration to the West African Telecommunications Regulators Assembly (WATRA), and explores challenges to regional telecommunications regulation. This discussion resonates with

the analysis provided in Chapter 9 regarding challenges to harmonisation in Africa. Once more, the author highlights the potential for harmonisation, this time through the instrumentality of WATRA. This largely untapped potential is given further detailed consideration in the final chapter, Chapter 11, where Orji provides insights into other sub-regional efforts in Southern, East and Central Africa. In the Southern African context, for example, he argues that "SADC represents at least three main legal traditions, namely: the Common Law, Roman-Dutch Law and Civil Law" (p. 366). He further argues that these variations in legal systems within the region create impediments to the harmonisation of telecommunications law and policies. While there is scope for debate on this matter, the author has succeeded in articulating his views in a cogent manner, which is of great practical benefit to the body of academic literature in this area. Other aspects of Orji's analysis of sub-regional organisations include treatment of COMESA and the EAC. He concludes that, in respect of facilitating regional economic integration, "to a large extent, the […] challenges appear similar across all the sub-regions […]" (p. 389).

A major strength of this book lies in the scope of its coverage. In an area where there is a paucity of academic literature, especially literature providing in-depth consideration of African realities, this volume is a commendable effort that scholars will find to be a good basis for further research. The international comparative legal analysis makes the book potentially appealing to scholars in a variety of jurisdictions around the world. Beyond scholarly circles, the book should also be useful to lawyers, regulators, policymakers, and potential investors.

AJIC 25 - 5-July-20 - 20h14.indd 4-5 7/5/2020 8:22:51 PM

The African Journal of Information and Communication (AJIC)

Published by the LINK Centre
University of the Witwatersrand (Wits)
Johannesburg, South Africa
https://www.wits.ac.za/linkcentre

ISSN 2077-7213 (online version)
ISSN 2077-7205 (print version)

AJIC 25 - 5-July-20 - 20h14.indd 6 7/5/2020 8:22:51 PM