

# Records in social media: a new (old) understanding of records management

Babatunde Kazeem Oladejo

*Department of Computer Science and Information Systems,  
Sarajevo School of Science and Technology, Sarajevo, Bosnia and Herzegovina, and*

Darra Hofman

*School of Information, San Jose State University, San Jose, California, USA*

## Abstract

**Purpose** – Social media posts have been an integral part of our society’s communication and serve purposes from the personal to the national, from the mundane to the silly to the momentous. This study aims to examine social media posts as records, discussing how social media technology serves, perhaps unexpectedly, to reinforce traditional archival understandings of issues such as provenance, custody, access, disposition and preservation.

**Design/methodology/approach** – This study follows a four-step methodology. First, this study analyzes literature for a matching definition of the social media record. In the second step, we appraise three social media postings previously curated and cited in news articles by journalists to determine their characteristics – Are these social media posts “records?” Third, this study evaluates the sample records against two dominant theoretical record models, the life cycle and the continuum and attempt to apply the model specifications to the data samples. Finally, this study proposes appropriate records management solutions to address governance issues from the study findings in the conclusion section.

**Findings** – This study shows that, even by the most traditional of definitions, social media posts are records. The paper also demonstrates that platform mediation transforms simple narrative documents into records whose provenance, custody and control are dictated by platform logics and governance, outside of the control of their creators. Through appraisal of a small sample of “important” social media posts, this study illustrates that, rather than obsolete, traditional records management concepts and approaches are necessary to ensuring the ongoing accessibility, usability and evidentiary character of social media posts in the broader “platformized” context.

**Research limitations/implications** – This is exploratory, theoretical work. In future works, this study plans to expand and validate aspects of this study.

**Originality/value** – This paper tests existing theoretical frameworks, namely, the Records Life cycle and the Records Continuum for applicability to the social media record. The paper also offers a view of the potential for traditional archival and records management concepts in service of a just and inclusive recordkeeping, because such concepts allow us to demonstrate the centralized, elite-serving, bureaucratic structures which underpin social media records are obscured by the seemingly decentralized, participatory nature of social media.

**Keywords** Social media, Records management, Social media record, Record, Non-record, Platform governance

**Paper type** Research paper



## Introduction

The records challenges associated with social media are well-known; in the US National Archives and Records Administration (NARA)’s social media guidance from a decade ago,

“noteworthy challenges” include: “recordkeeping in a collaborative environment; content located in multiple places; ownership and control of data that resides with a third party; [ . . . ] development and implemented of records schedules” (2014, 2), as well as the ubiquitous digital challenges of trustworthiness, privacy and volume. Despite the well-established challenges inherent in social media records, which should make the area interesting to researchers, a review of the electronic records management literature found that “social media is underdeveloped and without a strong records management presence” (Oladejo and Hadžidedić, 2021, 74). Even though social media platforms have become important repositories for personal, corporate and governmental information, our understanding of the impact of these platforms *vis-à-vis* records and recordkeeping is still developing. Social media mediates creation, custody, control and ownership in ways that feel unprecedented, challenging our understanding of fundamental issues as provenance, appraisal and access. As Jessica Bushey (2014, 34) wrote in her paper on digital photographs in social media:

[ . . . ] social networking platforms as repositories for digital photographs and social memory should be examined from the archival perspective in which consideration of ownership, copyright and privacy must be weighed along with ongoing accessibility and long-term preservation.

Social media have only grown more prominent in the near decade since Bushey wrote these words. Archival scholars continue to do important work on issues ranging from detecting “fake video” (Hamouda *et al.*, 2019), to long-term preservation and the role of platform application programming interfaces as “technologies of custody” (Acker and Kreisberg, 2020) and the gap between record *making* and record *keeping* (Sheffield, 2018). But the larger question of social media posts as records, and of the utility of records management approaches to social media posts, remains open. Diffuse digital records – what Anne Gilliland (2014) persuasively describes as “networked records” – have raised so many questions that one finds regular assertions in the literature that fundamentals of archival practice and theory are simply no longer relevant. As an example, Gilliland, in her article argued at the time that “[t]raditional records management activities such as records retention scheduling and appraisal [ . . . ] are struggling in the network society, and indeed will likely become obsolete” (p. 29). Lynch (2017, 1) goes further, arguing that:

[t]hinking rooted in traditional archival methodology – focusing on the preservation of physical and digital objects, and perhaps the accompanying preservation of their environment to permit subsequent interpretation or performance of the objects - has been a total failure for many reasons.

And yet, records retention schedules, appraisal and even provenance remain intact, even as digital content and approaches to its use, preservation and access proliferate. Indeed, the Library of Congress abandoned its attempt to preserve the whole of Twitter, moving instead to a strategy of appraisal and selection (Kim *et al.*, 2013). Why have digital technologies – and especially social media, which allow billions of people to participate in documented public conversations in unprecedented ways – not overturned these traditional concepts and approaches? Is it simply a matter of inertia, a failure of imagination or an unwillingness to embrace change?

There is a common belief that technology, by giving users greater access to information and to the production of information, will be democratizing and liberatory, challenging extant power structures and inequalities that underlie society. In such a state of liberation, new models of records and recordkeeping become necessary. For example, Gilliland notes that some scholars challenge “the dominant model of provenance” as “perpetuat[ing] existing bureaucratic power structures and elites” (2014, 24). Networked technologies – social media, cryptocurrency and blockchain, the internet itself – have, each in their turn,

been looked upon as solutions that would change these power structures. For archives, networked technologies, including social media, have been touted as a solution to a limited, exclusionary picture of society perpetuated by archives whose mandates and limited resources ensured that a very privileged sliver of society's documentary heritage survived (Levi, 2013, 35).

It is obvious that our burgeoning networked information and communications technologies (ICTs) have led to an explosion of both documents and records, with people and organizations contributing to the creation in ways that might initially seem trivial, but later evoke serious consequences. Lauren Goode (2021) planned her wedding on social media, then called it off at the last moment but continues to get anniversary greetings and family themed shopping offers, three years later, at an untold psychological cost. Tragically, social media content has become a major health risk to children, not just for risky behaviors and addictions but also for suicides (Klepper, 2021). The various controversies around Facebook, including the Cambridge Analytic scandal, performing unconsented research on users that influenced the US elections in 2016 (Rehman, 2019), not only threaten the largest democracy in the world, but served as a model for copycat follow-ups in in other countries (Cesarino, 2020). And the many controversies around Twitter, since its acquisition by Elon Musk (Barrie, 2022), highlight the fact that social media platforms are, at the end of day, owned and governed by private actors whose primary interest is profit. So where is liberation?

It is evident that these technologies remain centralized, bureaucratic and elite-controlled. No matter how many of us may *use* Twitter, Facebook or TikTok, users have limited control over their posts and especially over the ongoing trustworthiness and accessibility of social media posts. Gayo-Avello puts it bluntly: "social media is the product of communicative capitalism, and the goal is not to boost political action but to commoditize and monetize individual communication" (Gayo-Avello, 2015, 10). However problematic their origins, traditional archival concepts such as provenance and the evidentiary character of records remain relevant because the bureaucratic, legally defended and dictated power structures of society remain intact. Networked technologies may obscure those structures, but they have not eliminated them. In reimagining provenance, for example, do we risk obscuring power dynamics and structures that continue to exist and are even reified through these ICT infrastructures?

Ultimately, there is, these authors argue, a fundamental question at play in our understanding of social media posts as records: Should our approach to records reflect what is, or what should be? In other words, if we accept *a priori* that archives heretofore have both reflected and participated in perpetuating deep inequalities and injustices and have an obligation to pursue justice moving forward, must we also create new understandings of records and archives? Let us start by addressing the abuse of platform powers by social media moguls that exercise *permanent* control over public records for profit. Would a traditional application of records management suffice in taming that power by ensuring that records are kept only for a specified purpose (based on classification), kept only for a specified period (retention) and removed upon term-completion (disposition)? In this context, can a traditional understanding of records and archives support a more just future for the public user?

### Methodology

This study follows a four-step methodology. First, we analyze literature for a matching definition of the social media record. While we acknowledge that social media platforms use numerous backend systems in which a diversity of records are created, for the purpose of this paper, "social media records" refers specifically to user-created postings, such as tweets on Twitter. In the second step, we appraise three social media postings previously curated and cited in news articles by journalists to determine their characteristics – Are these social

---

media posts “records?” Third, we evaluate the sample records against two dominant theoretical record models, the life cycle and the continuum and attempt to apply the model specifications to the samples. A good fit would further assert that the samples are records and gaps would be considered opportunities for further research. Finally, we propose appropriate records management solutions to address governance issues from the study findings in the conclusion section.

## Findings

### *Social media posts as records*

There are few questions which bring out more passion from record professionals than “What is a record?” Yet, identifying where – and indeed if – there are records created in new ICTs, such as social media, is key to managing the data and information created, disseminated and stored through such technologies. While records professionals may debate the concepts, definitions and even archival relevance, the broader world largely defaults to treating social media postings as records: in courts of law (Faklaris and Hook, 2016), government operations (NARA, 2014), commerce (Tuten and Mintu-Wimsatt, 2018) and societal memory (Johnston, 2016). However, the distinction between records, information and data is critical in determining how to appropriately treat social media postings; as Yeo (2018, 141) reminds us, “Records resound with a complexity of meaning and performativity, which the simple concept of records as information is not rich enough to encompass.”

Although there is no shortage of discussions of the definition of “records” in English, it is nonetheless necessary to begin this discussion by anchoring what we mean when we define social media postings as “records.” Let us begin with Duranti’s (2009, 44) definition of a record as “a document made or received in the course of a practice activity as its instrument or by-product and set aside for action or reference.” Despite the varying models of how to capture and preserve records (recordkeeping models), or even competing fundamental definitions from scholars, the core definition of record as “an important document set aside for action or reference” is common to all (Bearman and Trant, 1998, 14). The recordkeeping mechanism is where the various record defining models differ (Lappin *et al.*, 2021), and at this point in the definition, recordkeeping is secondary. We must first ascertain and answer the question – *What is a social media record?* before we attempt to manage it.

The traditional record definition connotes that not all documents are records, essentially designating the remaining documents as non-records. While the record undergoes an elaborate recordkeeping scheme, the non-record is treated as part of an aggregation that is minimally managed and scheduled for early disposition (Read and Ginn, 2015, 7). It is, however, essential to recognize that selecting data as non-record does not make it ineligible for becoming a record. This fact was best asserted by de Perio Wittman (2021, 71):

Non-records are all records exempted from the [Presidential Records Act] PRA and the [Federal Records Acts] FRA. Federal records and nonrecords may be considered agency records under the Freedom of Information Act (FOIA).

In essence, non-records are records, but records that are deemed to have non-substantial information value. Dionne and Carboni (2009, 50) in a E-record management case review found that while the corporate filing system contained 84% records and 16% non-records, the less formal email system had 4% records and 96% non-records. The study was only able to provide these statistics because they captured and managed non-records. Patricia Franks (2016, 49) in an advice to government agencies, cautions that the failure to manage social media non-records will cause the agencies difficulty in information retrieval, inefficient resource utilization and increased e-discovery costs in event of FOIA requests and lawsuits.

The same guidance should apply to public social media, where the flood of data is currently unrelenting, poorly managed and without archival controls.

From the perspective of the authors, if both records and non-records are records, then all social media postings are records in the most traditional sense: a corporation using Twitter to put forth a press release or a local government posting an update about garbage collection on Facebook are creating records to effectuate a practice activity. Mosweu (2022, 47) shares the example of the Public Records Office Victoria, which:

[...] stat[e]s that social media posts created or received by a public officer in the course of their duties, are evidence of government business, as they document the actions taken by public officers and should be retained for reasons of accountability and transparency.

However, for every local government sharing infrastructure updates on Facebook or Wendy's roasting their competition on Twitter, there are legions of people posting about their children, their favorite sports team or their lunch:

Social media is an online environment where content is created, consumed, promoted, distributed, discovered or shared for purposes that are primarily related to communities and social activities rather than to functional, task-oriented objectives (Mosweu, 2019, 51).

Indeed, our understanding of social media records is complicated by the fact that personal social media posts which, from the perspective of the authors, are not instrumental to any practical activity are nonetheless records. A casual "hello" social media posting between two individuals would nominally be a non-record, without an evidential value. However, should the two people be involved in a criminal investigation and claim unacquaintance, suddenly the otherwise trivial "hello" transaction becomes evidential in a court of law. Consider a hypothetical applicant for lawful permanent residence in the USA based on marriage to a US citizen. If the applicant wrote a private letter to her sister, calling her spouse any number of names and saying how much she regretted marrying them, the letter would be sad, but likely, juridically irrelevant. If, however, she did so in a public Facebook post, the United States Department of Homeland Security, which monitors "publicly available" social media accounts of immigration applicants, could take her post as evidence of fraud, leading to the denial or rescinding of her permanent residence, a significant juridical consequence (DHS, 2019). Even banal posts – such as a picture of a meal with a "check in" tag to a restaurant – are now transactions. By receiving the record into its *fonds*, the social media platform seeks to obtain the effects (data aggregation and monetization) guaranteed in contracts, such as the terms of service, and broader commercial law. Outside of social media, such social interactions would not have juridical consequence, nor would they be preserved for their evidentiary capacity, unless otherwise documented.

In surveillance capitalism, it is the presumed evidentiary nature, the transactional value of the user-generated records which motivates platforms to provide their services for free to the user. The records – including both the posts and their metadata – are valuable because they are presumed capable, through sophisticated data analytics, to serve as the *factum probans* for any number of *factum probandum*. It is precisely their mediation by social media platforms which makes social media posts into records, as they are instrumental to the practical activity of the platforms, which function as part of the platform's *fonds*:

The *fonds* is thus the conceptual 'whole' that reflects an organic process in which a records creator produces or accumulates series of records which themselves exhibit a natural unity based on shared function, activity, form or use (Cook, 1993, 33).

In other words, social media platforms have made themselves not mere repositories, but creators of a *fonds* of immense size and diversity, receiving records from almost countless

---

authors. To advance this argument, let's evaluate a sample of social media data of diverse composition and evaluate them against the two most referenced record models: Records Lifecycle and Records Continuum.

### *Sample records appraisal*

Appraisal in the context of social media is somewhat complicated. A decision was required on what should constitute a meaningful sample. Prudence of judgment led us to elect a curated record source, social media postings cited by journalists in articles in reputable news publications, as they are often regarded as credible sources of information (Deacon, 2007). Journalists are record professionals in their field, competent in the selection of social media postings as “evidence of an activity or event” (matching the definition of a record). Although it is possible to select a larger pool of social media records, we limited this appraisal exercise to just three samples to devote sufficient attention to the appraisal detail without overwhelming the reader.

The next challenge was deciding on an appraisal method. In practice, application of the record appraisal method can be variable and largely subjective with the final decision of “record or not record” often reliant on specialist knowledge and disputable human judgment (Caron and Brown, 2013). It was therefore essential to apply a template to the appraisal process. In this study, we used Boles and Young's (1985) three appraisal elements of value of information, cost of retention and implications of appraisal recommendation. We evaluated each of the three sample social media citations on these elements to arrive at record decisions. Table 1 presents the records appraisal report for each of the social media data samples.

### *Records Lifecycle analysis of the appraised records*

The Records Lifecycle theory posits that a record's life starts at its creation and ends at disposition. There are various adaptations of the life cycle model varying from Roper's three stages to Goodman's ten stages (Yusof and Chell, 2000, 136–37). For this paper however, we adopted Tayfun and Gibson's (1996) three stages of records life cycle as creation, maintenance and use and disposition, which corresponds to what Pearce-Moses (2005, 232) describes as common to all the life cycle models. Figure 1 presents the diagram of the records life cycle inclusive of the sub-components used in this analysis. For organizations that use social media for official purposes, the life cycle as-is works relatively well for social media records management, *insofar* as those records are received into the organization's *fonds*. For social media at the platform level, the life cycle is more complex and tests the model in ways that the corporate electronic record had not previously done.

## **Creation**

Traditionally, record creation, not to be confused with document creation, is the selection of a finalized business document as a record via an appraisal process. Upon record creation, file classifications and retention schedules are applied, and the record is available for business use, in the second stage. The record is the official copy, while other non-record copies might promulgate (Tayfun and Gibson, 1996, 2).

For our social media record samples, a similar analogy can be made. Although Jack Dorsey's “just setting up my twttr” tweet was posted in 2006, presumably as a non-record tweet, it achieved record status by 2021 after its crowd-sourced appraisal of 122,000 retweets, 19,500 quotes and 180,100 likes culminated into a sale worth \$2.9m. Darnella Frazier's George Floyd video on the contrary went instantly viral, although there was still a pre-appraisal moment, even if only a few minutes or hours before it became worthy of the *viral* public appraisal status on Facebook. Elon Musk's Taking Tesla Private in violation of

**Table 1.**  
Social media data  
sample appraisal  
report

	Record appraisal 1	Record appraisal 2	Record appraisal 3
Article title	Jack Dorsey sells his first tweet ever as an Non-Fungible Token (NFT) for over \$2.9 million (short title: Jack Dorsey's First Tweet)	Darmella Frazier captured George Floyd's death on her cellphone. The teenager's video shaped the Chauvin trial (short title: Darmella Frazier's George Floyd Video)	Elon Musk's tweet about taking Tesla private has triggered a federal lawsuit (short title: Elon Musk Taking Tesla Private)
Article date	March 24, 2021	April 20, 2021	September 28, 2018
Article source	Consumer News and Business Channel (CNBC)	New York Times (NYT)	Vox News
Article link	<a href="http://www.cnbc.com/2021/03/22/jack-dorsey-sells-his-first-tweet-ever-as-an-nft-for-over-2point9-million.html">www.cnbc.com/2021/03/22/jack-dorsey-sells-his-first-tweet-ever-as-an-nft-for-over-2point9-million.html</a>	<a href="http://www.nytimes.com/2021/04/20/us/darmella-frazier-video.html">www.nytimes.com/2021/04/20/us/darmella-frazier-video.html</a>	<a href="http://www.vox.com/2018/9/27/17911826/elon-musk-tesla-sec-twitter-lawsuit">www.vox.com/2018/9/27/17911826/elon-musk-tesla-sec-twitter-lawsuit</a>
Cited source	Twitter	Facebook	Twitter
Cited Post ID	20	1670313089836457	1026872652290379776
Cited Post	@jack "just setting up my twitr" 9:50 PM · Mar 21, 2006 122.1K Retweets 19.5K Quote Tweets 180.1K Likes	Darmella Frazier March 11, 2021 "I still can't get over how quick the news tried to cover up George Floyd's death. Just makes me think what else got covered up if it was no evidence to see what really happened. This world we live in is sick and things need to change! If you think Derek Chauvin was "just doing his job" YOU'RE APART OF THE PROBLEM. George Floyd was already cuffed on the ground, a knee to the neck when you're already restrained is absolutely unnecessary. Despite this man's past, nobody deserves to die because they have a past. That man was begging for his life and Chauvin did not care. He deserves to go down. Anyone who thinks differently, you're apart of the problem." #justiceforGeorgefloyd"	@elonmusk "Am considering taking Tesla private at \$420. Funding secured." 6:48 PM · Aug 7, 2018 14.3K Retweets 7,243 Quote Tweets 83.5K Likes

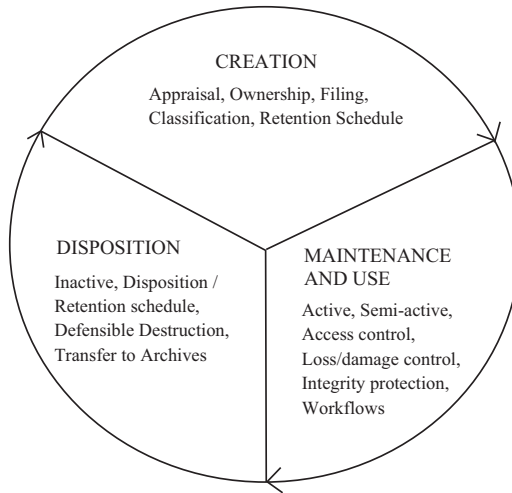
(continued)

	Record appraisal 1	Record appraisal 2	Record appraisal 3
Value of information	This tweet, the first ever posted on Twitter, has become a societal memory item. A search on Google Scholar for the exact string “just setting up my twtr” produced 375 results in Dec 2022. The purchaser of the tweet’s nonfungible token, Sina Estavi, stated: “This is not just a tweet! I think years later people will realize the true value of this tweet, like the Mona Lisa painting”	An ordinary citizen captured a video of police brutality and posted it on social media. The video went viral and became the central evidence in a major court case.	The United States Securities Exchange Commission (SEC) considered the tweet misleading information aimed at Tesla stockholders and inflating the stock price. This a serious violation of SEC regulations
Cost of retention	Publicly available free of charge as of this writing. Service provider cost undisclosed	Publicly available free of charge as of this writing. Service provider cost undisclosed	Publicly available free of charge as of this writing. Service provider cost undisclosed
Implications of the appraisal recommendations	By converting this tweet into a Non-Fungible Token (NFT), with the tamper-resistance provided by blockchain technology (Locke 2021), the platform provider is capable of creating social media records compliant with the traditional ISO-15489 record specifications of authenticity, reliability, usability and integrity (Hamidovic, 2010)	The video was authenticated and admitted as evidence at trial, with several news media replaying it from social media. The police officer was found guilty and sentenced to jail	The tweet led to civil and criminal investigation of Elon Musk, with civil fines levied by the SEC.
Appraisal decision	Select as a record	Select as a record	Select as record

Source: Table by authors

Table 1.





Source: Figure by authors

Figure 1.  
The Records  
Lifecycle model

an SEC rule indicates a regulatory appraised selection process but the tweet was only consequential because of the public access and reaction to the “misleading information aimed at Tesla stockholders.”

The main challenge of social media record creation is the influence of crowd-sourced appraisal, which, irrespective of professional appraisal, empowers the public user to decide what is record worthy or not. It is also important to note that social media record ownership is unlike the traditional model because it is shared beyond the original creator, and even the platform’s right to the record, to the public. For example, in the case of Jack Dorsey’s first tweet the *de facto* owner is now Sina Estavi, but it is still shared amongst the stakeholders. Although filing is proved by the platform storage location of the record, there was no apparent filing or retention classification information found on the study data samples.

Finally, the life cycle model depicted in Figure 1 shows a flow from disposition back to (re)creation. We found this flow correct for the traditional record, where it can be superseded via a disposal and recreation, for example, it is common practice to supersede work-practice records in this manner (Bradshaw and Rickards, 2018, 7). For social media, the flow from disposition to (re)creation is however, variable from platform to platform, for example, Facebook allows users to edit posted text but not photos or videos, whereas Twitter only allows a short-term window for text edits (McCluskey, 2022).

### Maintenance and use

After creation, the life cycle record enters an active stage, when it is highly used. Tayfun and Gibson reported that for the traditional record common estimate is that “as much as 80% of the activity on a record occurs in the first 20% of its life” (Tayfun and Gibson, 1996, 7). After the initial utility, the record undergoes a semi-active stage when it might be stored at an off-site location, but restored to active stage if demand increases. Irrespective of the active or semi-active status, access control, loss or damage control and integrity protection mechanisms are enforced during maintenance and use.

Interestingly, the social media record also undergoes *trending* and *dormant* stages analogous to the traditional active and inactive stages (Jansen *et al.*, 2021). Platform storage location is undisclosed and irrelevant to the user who can access the record through the same URL, however, IT infrastructure that offload dormant data to secondary, cheaper storage is commonly used in the industry (Pu *et al.*, 2019). While under maintenance and use, social media data enjoys all the access control and integrity protection available to the traditional digital data because it is also digital.

### Disposition

Disposition, not to be confused with (*ad-hoc*) deletion which is part of maintenance and use, occurs after the expiration of the assigned retention period. There are two primary disposition outlets in the life cycle: destruction or transfer to the archives. The life cycle destruction is considered defensible disposition because it is a pre-negotiated disposal of the record based on record schedules, supported by policy and applicable laws (Lemieux *et al.*, 2019, 49).

Given the lack of retention classification and schedule information on our social media record samples, disposition is presumed missing. As Barnard *et al.* (2019, 117) note, “records cannot be disposed of unless the reasons for creating them in the first place are understood,” in essence without applying filing and retention classifications at creation, disposition is disenfranchised. It would therefore appear from the social media platform perspective, that the appraisal decision was made permanent retention for everything not deleted by a user.

### *Records Continuum analysis of the appraised records*

The continuum model prescribes a multi-faceted, continuously and recursively applied set of processes (Frings-Hessami, 2022), in contrast to the life cycle model which sets the processes in linearly progressive phases. There are no shortages of theoretical discussions about the Records Continuum model, with praises and criticism. Clear and concise application examples are however rare. For this study, we adopted Karabinos’s (2018) experiments with the Foreign and Commonwealth Office “Migrated Archives” data set. The report clearly demonstrated the power of the continuum model by applying the dimensional constructs of Create, Capture, Organize and Pluralize to a set of records over a time (period) and space (events). For example:

In the early 1980s, at the Hayes repository, records were reviewed and inventoried by FCO staff, while some were most likely destroyed. I would consider this action both re-creation (1D) and re-capture (2D). Once the public became aware of the Migrated Archives during the Mau Mau court case and their movement to The National Archives a further re-creation occurred (1D). Here, they were also captured, organized, and pluralized by The National Archives (2D, 3D, and 4D).

[...]

Furthermore, a new process begins at pluralization (4D), as the Migrated Archives then became an impetus for an untold number of records to be created, such as court records, parliamentary inquiries, and internal memorandums (1D). These include the Cary Report, but also all the records that went into the creation of the Cary Report that were then captured and organized (2D, 3D). Cary, while writing his report, would have re-created records that helped him in his research (1D). The Cary report has also been pluralized and made accessible (4D) (p. 218).

With this example, one can begin to see the dynamic nature of the continuum model as the experiment navigates the dimensions. However, this example still does not apply the full spectrum of the Record Continuum model’s capability. Application of the axis elements were not discussed, perhaps in a simplification attempt, or were the dimensions only applied to the recordkeeping axis? Irrespective of this concern, Karabinos provided the best example

we found in our literature search and served as the guide for our experiment. We tabulated and numbered the continuum dimension and axis processes to produce intersectional elements, which were also numbered for reference ease (Table 2). We then applied the records continuum actions to the three social record samples.

*Jack Dorsey’s first tweet.* On March 21, 2006, Jack Dorsey (1D-1A: Actor) posted a message, “just setting up my twttr” (1D-4A: Document) visible to the public users on Twitter. This seemingly ordinary tweet was sold to Sina Estavi (1D-1A: Actor) for \$2.9m (1D-3A: Transaction) as a non-fungible token record (2D-4A: Record) after having acquired over 122,000 retweets, 19,500 quotes and 180,100 likes social media valuation (2D-3A: Activity), as reported by CNBC (4D-2A: Collective Memory) on the March 24, 2021.

A simplified version of the continuum application above would be Jack Dorsey posted a non-record tweet (1D), which underwent public user records appraisal (2D) leading to a sale (3D) to Sina Estavi (1D) as announced by major news channels (4D). There are indications of missing events in this continuum application. What happened between the initial document creation (1D) and the record purchase transaction (3D)? The tweet must have attained record status before the purchase transaction, which would indicate an unknown capture (2D) with at least a distribution (4D). Is it possible to find the missing “shadow” continuum events?

*Darnella Frazier’s George Floyd video.* On April 20, 2021, *New York Times* (4D-1A: Institution) reported the role Darnella Frazier’s George Floyd video (2D-4A: Record) played in the murder trial (4D-3A: Purpose) of former police officer Derek Chauvin (1D-1A: Actor). The newspaper report cited a Facebook message (2D-4A: Record) posted on March 11, 2021 by Darnella Frazier (1D-1A: Actor), recalling the events surrounding George Floyd’s (1D-1A: Actor) death and the public reaction (4D-2A: Collective Memory).

A simplified version of the continuum application to this use case would be *New York Times* reported (4D) the role of Darnella Frazier’s George Floyd video (2D) in the murder trial (4D) of Derek Chauvin (1D) and cited a Facebook message (2D,1D,1D,4D). This use case also invites the analyst to troubleshoot and find missing records. Where is the original video posted by Ms Frazier? When was that recorded? Why is that record no longer publicly available?

*Elon Musk taking Tesla private.* On 28-Sep-2018, Vox News (4D-1A: Institution) reported that the SEC (4D-1A: Institution) has filed a lawsuit (4D-3A: Purpose) against Elon Musk (1D-1A: Actor) for violating the agency’s rule on information disclosure to shareholders. The tweet “Am considering taking Tesla private at \$420. Funding secured.” (2D-4A: Record) led to fines (1D-3A: Transaction) being levied (4D-2A: Collective Memory) against Mr Musk.

A simplified version of the continuum analysis would be: Vox news (4D) reported SEC’s lawsuit (4D) against Elon Musk (1D) for a social media posting that violated agency rules. The tweet (2D) led to levied fines (4D). Here again, with three pluralization events on a single record, the continuum invites the analyst to investigate shadow events with probes, such as,

	1A: Identity	2A: Evidentiality	3A: Transactionality	4A: Recordkeeping
1D: Create	1D-1A: Actor(s)	1D-2A: Trace	1D-3A: Transaction	1D-4A: Document
2D: Capture	2D-1A: Unit(s)	2D-2A: Evidence	2D-3A: Activity	2D-4A: Record
3D: Organize	3D-1A: Organisation	3D-2A: Corporate/ Individual Memory	3D-3A: Function	3D-4A: Archive
4D: Pluralize	4D-1A: Institution	4D-2A: Collective Memory	4D-3A: Purpose	4D-4A: Archives

**Table 2.**  
Tabulated and  
numbered records  
continuum processes

Source: Table by authors

---

was there a complaint about Mr Musk's tweet? Were there other related records created before the Vox news report?

## Discussion

### *Social media record and platform governance*

The now iconic Jack Dorsey tweet “just setting up my twttr” is remarkably plain, perhaps even trite. What gave the tweet its record worthiness is its social network characteristics: the comments, the likes and retweets. When combined with the buyer, Sina Estavi's tweet, “*This is not just a tweet! I think years later people will realize the true value of this tweet, like the Mona Lisa painting,*” the record value becomes better established. And it is a good example of Gilliland's (2014) created by the crowd, combined and stitched-together description of the networked record and a proof that hugely diverse content can become a record. But it also points to the much older concept of the archival bond. It has always been true that identical documents can be very different records, depending upon their archival bond, and that identical documents can be records in many different *fonds*.

Social media demonstrates that the legal, and especially the evidentiary, nature of records remains critical and contingent. It also demonstrates that archives both participate in, and are a product of the broader culture, including legal and regulatory limitations. For example, critics of provenance argue that, by failing to acknowledge:

[...] the multi-provenance bureaucratic record and the record created by the crowd [...] provenance] renders others who participate in the production of the record as mere subjects rather than co-creators with rights in those records (Gilliland, 2014, 24).

And indeed, it would seem that social media platform and the law both acknowledge that users are creators; Twitter's Terms of Service famously says, “What's yours is yours — you own your Content (and your incorporated audio, photos and videos are considered part of the Content),” and copyright law in many countries, including the USA and Canada, acknowledges copyright as belonging to the creator at the moment of creation. However, a regulatory regime that largely leaves large technology companies to self-govern (Barret, 2020) and allows them to enforce standardized “click-wrap” agreements in which the technology companies could hardly be said to be arms-length, betrays the power imbalances at play in the system. Ultimately, the power of tech companies – and specifically social media companies – is immense and regulation instruments extraordinarily light. As the *Washington Post* reported:

[t]he Jan. 6 committee spent months gathering stunning new details on how social media companies failed to address the online extremism and calls for violence that preceded the Capitol riot. The evidence they collected was written up in a 122-page memo that was circulated among the committee, according to a draft viewed by The Washington Post. But in the end, committee leaders declined to delve into those topics in detail in their final report [...] concerned about the risks of a public battle with powerful tech companies (2023).

Cheney-Lippold (2017, 254) argues that, among other algorithm-driven technologies, social media *datafy* everything in a way that short-circuits traditional relationships with records and evidence:

I really mean *everything*. Love, friendship, criminality, citizenship, and even celebrity have all been *datafied* by algorithms we will rarely know about. These proprietary ideas about the world are not open for debate, made social and available to the public. They are assigned from behind a private enclosure, a discursive trebuchet that assails us with meaning outside the castle walls.

There is little to no engagement with the record *in its context*, acknowledgement of the evidentiary character of records, the possibility of multiple perceptions nor values beyond

that which generates profit. While it is arguably true that archivists have always assigned meaning to records through such inescapably human – and therefore limiting – processes as arrangement and description, there is a qualitative difference in the *datafied*, algorithmic process and archival record awareness. However, by engaging with the monetized, *datafied* nature of social media records through archival and diplomatic lenses, the archivist can participate in and contribute to the debate regarding improvements to the regulatory instruments that will shape the future of social media platform governance.

#### *Records management model*

The social media record embodies the records continuum ideal, one that lives in a continuous utility of space and time, changing with every new comment, like, forward, reaction and as McKemmish (1994) says, “is always in a process of becoming.” The Records Continuum model encourages the records analyst to follow evidential traces, and discover related records in the shadow of the known, unlike the life cycle model that simply organizes the known record, optimized for disposition. However, despite the acceptance of the social media record’s “always becoming” characteristic and the push to “keep everything” because storage is (purported to be) cheap, or to avoid blame for unfair destruction, “secure, compliant information disposition has its place!” (Franks, 2017). Gable (2015) emphasizes that mature information governance must realize the true benefit of legally defensible destruction, implemented on the principles of scheduled retention, is significant, not just for cost savings but also for risk reduction. We argue that for social media being an informal information system, analogous to email and indeed more vernacular, non-records should be destroyed early for a healthy management of the remaining records. “Defensible destruction” as a disposition option is, however, conspicuously missing in the continuum model specification. Of missing terms in the Records Continuum model, Upward (2000) wrote:

In selecting terms I tried to choose ones that have a reliable dimensional locus and general significance to practice in archives and records management, but there are many unexpressed points of practical significance. Words like file and series are examples. Locating them on the continuum, however, is not something that can be done with any certainty. Files and series exist somewhere in the recordkeeping containers continuum, usually between the record and the archive.

File classification and record series are essential components of the record schedule, which would explain why the continuum model might be weak at defensible destruction, a proven strength of the life cycle model. Although a creative RM program might be able to effectively use the continuum model for defensible destruction, an argument can be made that this feature should be more prominent in the model specification.

#### *Record retention/disposition as a regulatory instrument*

If any *unsubstantial* non-record can become a *substantial* record, does it then make sense to keep all data perpetually for the chance that they might become records? In 2010, Twitter and the USA Library of Congress started a project to preserve all tweets ever created permanently, since Twitter’s inception in 2006, as an archival endowment, societal memory and research resource. In the first 2 years alone, approximately 170 billion tweets were captured (Kim *et al.*, 2013). The volume grew from 50 million tweets per day in 2010 to over 500 million per day by 2017 when the project was stopped (Fondren and Menard McCune, 2018). The noted challenges to the project were not only technological, but also of record-value. What is the value of keeping everything? Can someone actually read through or extract real value from the overwhelming volume of data? Although the archival debates of *keep-everything* versus *scheduled destruction* continues, we side with the opinion that it

---

essential to select records from a pool of content, remove the non-records early and manage the selected records through a well-defined process that includes *defensible destruction*.

In the authors' world-view of a regulated social media, the service providers would be mandated to disclose any retention/disposition schedule their internal algorithms assign to social media content and allow the public "crowd" users to update, add, remove these metadata on the items/containers to a point of aggregation. The proposed scheme is not simple, but neither are the coauthoring, emotional gratification and other features of social media. We acknowledge that such a scheme would take a while to perfect, but a start point is needed.

## Conclusion

Larsen and McGraw (2014, 267) reminds us of the timeless empirical analysis paradox:

The theory 'All ravens are black' rules out the existence of white ravens; and observation of a white raven refutes the theory.

Despite social media's volume of billions of postings per day, we answer the study's primary research question "does social media data contain records?" by appraising only three sample postings for their record worthiness. We could possibly manually appraise a hundred more, and it would remain a "drop in the sea" of social media data. We, however, believe these three "white ravens" are sufficient. To further assert that social media records exist, we applied the two most cited records management models, life cycle and continuum to the three sample data. Our analysis reveals that although at the fundamental level, the models fit the social media record, there are rooms for improvement, or perhaps synergetic cooperation between the two models. Records Continuum captured the essence of the social media record but is weak at defensible destruction, a key competence of the life cycle model.

The study's second goal was to apply traditional archival optics to the imperative complexities of social media and see if age-tested records management principles hold true for the social media record. Our investigation reveals that at the very fundamental levels of record vs non-record, social media is no different than its traditional equivalent. What diverges vastly, however, is the dynamic composition of the social media record. If the paper recordkeeping system was rendered obsolete by the manifold complexities of the electronic record, the social media record, being indeed, a subtype of the superlative network and algorithmic record had changed that dynamics by exponents.

A single click at the social media platform's terms-of-service splits record ownership unfairly between the technology company and actual record creator. While the naive public user enjoys gratifications of social engagements, unaware of the record value of every "like," "reaction," and *comment*, the well-funded platform owners amass *record-quality* data on them for corporate profit. Without archival guidance, and for the sake of technological progress, a weak self-regulatory regime in the USA spreads the pervasive effect across the world with the resultant chaos leaving social, cultural and democratic ideals in peril. The circles of regulatory fines and penalties for power abuse and violations are trivial if the companies earn multiple folds back, neither has the insincere executive apologies nor their promises to solve the problems. Without control of the social media content through an archival guidance that mandates the fundamental records management elements inclusive of scheduled disposition, governmental efforts in regulating social media will remain weak, *ad-hoc* and ineffectual. In the meantime, real people will continue to suffer the consequences.

## *Limitations and future work*

While this initial paper began with an exploratory "close-up" analysis, a three-record sample dataset is extremely small for social media, which boasts of big data in billions. In a future

paper, we plan to experiment with a bigger sample dataset and apply big data computing analytic methods to derive further insights into the characteristics of the social media record. And such future work might also propose a record model that combines both the life cycle and continuum models as a better fit model for the social media record.

### References

- Acker, A. and Kreisberg, A. (2020), "Social media data archives in an API-Driven world", *Archival Science*, Vol. 20 No. 2, pp. 105-123.
- Barnard, A., Bonilla, E., Franks, P.C., Rocha, C.C.M.L., Schenkolewski-Kroll, S. and Tractinsky, A. (2019), "Retention and disposition", *Trusting Records in the Cloud*, Vol. 117.
- Barret, P.M. (2020), "Regulating SocialMedia", *Center for Business*, New York University Stern, New York, NY.
- Barrie, C. (2022), "Did the Musk takeover boost contentious actors on Twitter?", ArXiv Preprint ArXiv:2212.10646.
- Bearman, D. and Trant, J. (1998), "Electronic records research working meeting, may 28-30, 1997: a report from the archives community", *Bulletin of the American Society for Information Science and Technology*, Vol. 24 No. 3, pp. 13-17.
- Boles, F. and Young, J. (1985), "Exploring the black box: the appraisal of university administrative records", *The American Archivist*, Vol. 48 No. 2, pp. 121-140.
- Bradshaw, E. and Rickards, L. (2018), "South Atlantic tide gauge data management plan".
- Bushey, J. (2014), "Convergence, connectivity, ephemeral and performed: new characteristics of digital photographs", *Archives and Manuscripts*, Vol. 42 No. 1, pp. 33-47.
- Caron, D. and Brown, R. (2013), "Appraising content for value in the new world: establishing expedient documentary presence", *The American Archivist*, Vol. 76 No. 1, pp. 135-173, doi: [10.17723/aarc.76.1.g5x055x8228xx1mu](https://doi.org/10.17723/aarc.76.1.g5x055x8228xx1mu).
- Cesarino, L. (2020), "how social media affords populist politics: remarks on liminality based on the Brazilian case", *Trabalhos Em Linguística Aplicada*, Vol. 59 No. 1, pp. 404-427, doi: [10.1590/01031813686191620200410](https://doi.org/10.1590/01031813686191620200410).
- Cheney-Lippold, J. (2017), "We are data", *We Are Data*, New York, NY University Press.
- Cook, T. (1993), "The concept of the archival fonds in the Post-Custodial era: theory, problems and solutions", *Archivaria*, pp. 24-37.
- de Perio Wittman, J. (2021), "A trend you can't ignore: social media as government records and its impact on the interpretation of the law", *Albany Law Journal of Science and Technology*, Vol. 31, p. 53.
- Deacon, D. (2007), "Yesterday's papers and today's technology: digital newspaper archives and 'push button' content analysis", *European Journal of Communication*, Vol. 22 No. 1, doi: [10.1177/0267323107073743](https://doi.org/10.1177/0267323107073743).
- DHS (2019), "Agency information collection activities: generic clearance for the collection of social media information on immigration and foreign travel forms", available at: [www.federalregister.gov/documents/2019/09/04/2019-19021/agency-information-collection-activities-generic-clearance-for-the-collection-of-social-media](https://www.federalregister.gov/documents/2019/09/04/2019-19021/agency-information-collection-activities-generic-clearance-for-the-collection-of-social-media)
- Dionne, M. and Carboni, A. (2009), "How to successfully implement an E-Records management program", *Information Management Journal*, Vol. 43 No. 2, pp. 49-54.
- Duranti, L. (2009), "From digital diplomatics to digital records forensics", *Archivaria*, pp. 39-66.
- Faklaris, C. and Hook, S.A. (2016), "Oh, snap! the state of electronic discovery amid the rise of Snapchat, WhatsApp, Kik, and other mobile messaging apps".

- Fondren, E. and Menard McCune, M. (2018), "Archiving and preserving social media at the Library of Congress: institutional and cultural challenges to build a Twitter archive", *Preservation, Digital Technology and Culture*, Vol. 47 No. 2, pp. 33-44.
- Franks, P.C. (2016), "Applying records management principles to managing public government social media records", *SocialMedia for Government: Theory and Practice*, Routledge, New York, NY, p. 15.
- Franks, P.C. (2017), "Even in a 'never delete anything' world, compliant information disposition has its place", *Information Management*, Vol. 51 No. 6, available at: [www.proquest.com/openview/e4234d255f44f610d302307fc7824313/1?pq-origsite=gscholar&cbl=47365](http://www.proquest.com/openview/e4234d255f44f610d302307fc7824313/1?pq-origsite=gscholar&cbl=47365).
- Frings-Hessami, V. (2022), "Continuum, continuity, continuum actions: reflection on the meaning of a continuum perspective and on its compatibility with a life cycle framework", *Archival Science*, Vol. 22 No. 1, pp. 113-128, doi: [10.1007/s10502-021-09371-2](https://doi.org/10.1007/s10502-021-09371-2).
- Gable, J. (2015), "Making a business case for the principles", *Information Management Journal*, Vol. 49 No. 5, pp. 34-38.
- Gayo-Avello, D. (2015), "Social media, democracy, and democratization", *IEEE Multimedia*, Vol. 22 No. 2, pp. 10-16.
- Gilliland, A.J. (2014), "Reconceptualizing records, the archive and archival roles and requirements in a networked society", *Knygotyra*, Vol. 63.
- Goode, L. (2021), "I called off My wedding. The internet will never forget", *Wired*, 2021, available at: [www.wired.com/story/weddings-social-media-apps-photos-memories-miscarriage-problem/](http://www.wired.com/story/weddings-social-media-apps-photos-memories-miscarriage-problem/)
- Hamouda, H., Bushey, J., Lemieux, V., Stewart, J., Rogers, C., Cameron, J., Thibodeau, K. and Feng, C. (2019), "Extending the scope of computational archival science: a case study on leveraging archival and engineering approaches to develop a framework to detect and prevent 'fake video'", *2019 IEEE International Conference on Big Data (Big Data)*, IEEE, 3087-97.
- Haris Hamidovic, C.I.A. (2010), "An Introduction to Digital Records Management", *ISACA Journal*, Vol. 6, pp. 1-6.
- Jansen, N., Hinz, O., Deusser, C. and Strufe, T. (2021), "Is the buzz on?—a buzz detection system for viral posts in social media", *Journal of Interactive Marketing*, Vol. 56 No. 1, pp. 1-17.
- Johnston, L. (2016), "Social news= journalism evolution? How the integration of UGC into newswork helps and hinders the role of the journalist", *Digital Journalism*, Vol. 4 No. 7, pp. 899-909.
- Karabinos, M. (2018), "In the shadows of the continuum: testing the records continuum model through the foreign and commonwealth office 'migrated archives'", *Archival Science*, Vol. 18 No. 3, pp. 207-224.
- Kim, A.E., Hansen, H.M., Murphy, J., Richards, A.K., Duke, J. and Allen, J.A. (2013), "Methodological considerations in analyzing twitter data", *JNCI Monographs*, Vol. 2013 No. 47, pp. 140-146.
- Klepper, D. (2021), "TikTok boosts posts about eating disorders, suicide", *AP News*, 2021, available at: <https://apnews.com/article/technology-health-eating-disorders-center-government-and-politics-0c8ae73f44926fa3daf66bd7caf3ad43>
- Lappin, J., Jackson, T., Matthews, G. and Ravenwood, C. (2021), "Rival records management models in an era of partial automation", *Archival Science*, Vol. 21 No. 3, pp. 243-266, doi: [10.1007/s10502-020-09354-9](https://doi.org/10.1007/s10502-020-09354-9).
- Larsen, J.T. and McGraw, A.P. (2014), "The case for mixed emotions: mixed emotions", *Social and Personality Psychology Compass*, Vol. 8 No. 6, pp. 263-274, doi: [10.1111/spc3.12108](https://doi.org/10.1111/spc3.12108).
- Lemieux, V., Hofman, D., Batista, D. and Joo, A. (2019), "Blockchain Technology and Recordkeeping", ARMA International Educational Foundation, ARMA Canada.
- Levi, A.S. (2013), "Humanities 'big data': myths, challenges, and lessons", *2013 IEEE International Conference on Big Data*, IEEE, 33-36.
- Lynch, C. (2017), "'Stewardship in the' age of algorithms", *First Monday*.



- McCluskey, M. (2022), "The promise—and possible perils—of editing what We say online", *Time Magazine*, September 23, 2022, available at: <https://time.com/6215340/edit-tweets-imessages-consequences/>
- McKemmish, S. (1994), "Are records ever actual", *The Records Continuum: Ian Maclean and Australian Archives First Fifty Years*, Ancora Press, Clayton.
- Mosweu, T. (2019), "The good, the bad and the ugly: social media prospects and perils for records management", *ESARBICA Journal*, Vol. 38 No. 1.
- Mosweu, T.L. (2022), "A review of the legislative framework for social media records in Botswana", *Records Management Journal*, Vol. 32 No. 1, pp. 62-74.
- NARA (2014), "Guidance on managing social media records" *National Archives and Records Administration*, available at: [www.archives.gov/records-mgmt/bulletins/2014/2014-02.html](http://www.archives.gov/records-mgmt/bulletins/2014/2014-02.html)
- Oladejo, B. and Hadžidedić, S. (2021), "Electronic records management—a state of the art review", *Records Management Journal*, Vol. 31 No. 1, pp. 74-88.
- Pearce-Moses, R. (2005), "A glossary of archival and records terminology", *Archival Fundamentals Series*, Society of American Archivists, Chicago.
- Pu, Q., Venkataraman, S. and Stoica, I. (2019), "Shuffling, fast and slow: scalable analytics on serverless infrastructure", in *NSDI*, Vol. 19, pp. 193-206.
- Read, J. and Ginn, M.L. (2015), "Alphabetic Records Management", *Equipments and procedures in records management*, Cengage Learning, Boston, 10th, Ed.
- Rehman, I.U. (2019), *Facebook-Cambridge Analytica Data Harvesting: What You Need to Know, Library Philosophy and Practice*, pp. 1-11.
- Sheffield, R.T. (2018), "Facebook live as a recordmaking technology", *Archivaria*, Vol. 85, pp. 96-121.
- Tayfun, A.C. and Gibson, S. (1996), *A Model for Life Cycle Records Management*, TRW Environmental Safety Systems, Inc, Vienna, VA (United States).
- Tuten, T. and 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*, Vol. 26 Nos 1/2.
- Upward, F. (2000), "Modelling the continuum as paradigm shift in recordkeeping and archiving processes, and beyond: a personal reflection", *Records Management Journal*, Vol. 10 No. 3, pp. 115-139.
- Yeo, G. (2018), "Records, Information and Data: Exploring the Role of Record Keeping in an Information Culture", Facet Publishing, London.
- Yusof, Z.M. and Chell, R.W. (2000), "The records life cycle: an inadequate concept for Technology-Generated records", *Information Development*, Vol. 16 No. 3, pp. 135-141, doi: [10.1177/0266666004240413](https://doi.org/10.1177/0266666004240413).

### Further reading

- Zakrzewski, C., Lima, C. and Harwell, D. (2023), "What the Jan. 6 probe found out about social media, but didn't report", *Washington Post*, January 18, 2023, available at: [www.washingtonpost.com/technology/2023/01/17/jan6-committee-report-social-media/](https://www.washingtonpost.com/technology/2023/01/17/jan6-committee-report-social-media/)

### Corresponding author

Babatunde Kazeem Oladejo can be contacted at: [kazeem.oladejo@ssst.edu.ba](mailto:kazeem.oladejo@ssst.edu.ba)