Data governance in healthcare information systems: A systematic literature review

Background: This study aimed to investigate data governance (DG) related to challenges associated with healthcare information systems (HIS), by reviewing guidelines emerging from academic sources as part of a consolidated systematic literature review (SLR). The research contributed theoretically towards the body of knowledge, by reviewing challenges and guidelines related to DG within the healthcare environment. It contributed practically to the body of knowledge through understanding the healthcare information’s systems status. The study also contributed methodologically and significantly to SLR strategies.

Objectives: The objective of this study was to understand the features of HIS; acquire information about DG success and understand the influence noted on DG.

Method: The study conducted an SLR over the period 2010–2020. Literature collection was not only restricted to South African publications but was extended to international sources. This study adapted a mono method.

Results: The study revealed that many organisations have realised that the only method to fix the data problem is the implementation of effective DG. With the increased adoption and rise of cloud computing, DG is gaining interest amongst specialists.

Conclusion: The shift from paper-based systems led organisations to seek organisational change through digital transformation. The proper collection and utilisation of electronic healthcare record is the foundation of the digital healthcare. Many organisations value DG as a promising method of maintaining data as a valuable asset.

Keywords: data governance, digitalisation, healthcare, healthcare information systems, eHealth, systematic literature review.

Introduction

Information systems (IS) play an important role in the digital era in many industries. Information systems are tools that help organisations to manage their data. Organisations invest in IS to drive their businesses to success in the changing of information technology (IT). The business components involve people, technology and processes. Without data, these business components are inactive, for data is treated as the lifeblood of the organisation (Russom 2012).

Organisations are facing a common goal that defines them as organisations. The types of organisations include higher business entities, education institutions and government departments, for example, healthcare. In the case of healthcare, the main goal is to deliver good health to people. The success of healthcare depends on data that flows within an organisation. Data are the most valuable asset in any organisation in this digital era (Lee et al. 2018). Many organisations realise that data governance (DG) as a promising method of keeping data as a valuable asset (Otto 2011a, 2011b).

Kitchenham and Charters (2007) suggest that a systematic literature review (SLR) is a strategy of assessing and interpreting all existing papers that are pertinent to the study. Siddaway (2014) explains an SLR as a method that addresses problems resulting from conflicting findings, produced by researchers. Piper (2013) points out that SLR permits complete, unbiased and literature-wide assessment of study results, design and quality. Okoli (2015) argues in detail that an SLR when properly done is valuable and turns into a highly cited part of the study that researchers pursue when undertaking a new investigation. Furthermore, such freestanding reviews summarise the evidence that is available to identify gaps in a research. The SLR method identifies, integrates and critically evaluates such findings.
Reasons for performing a systematic literature review

Various authors point out many reasons for undertaking an SLR:
• To synthesise the empirical confirmation of the limitations and benefits of a particular method;
• To recognise gaps in the existing research to provide directions for more investigations in these areas;
• To give a background or framework in order to correctly locate the activities of the current research (Kitchenham & Charters 2007).
• An SLR gives the opportunity to provide a structured and rigorous approach to conduct a standalone literature review (Okoli 2015); and
• The rigorous, systematic approach aims to minimise bias (Siddaway 2014).

The justifications listed above are relevant to the investigation undertaken in this study.

Advantages of a systematic literature review

An SLR adds rigour to the search strategy and minimises bias (Okoli 2015). Ryan (2010) identifies some advantages of an SLR that differ from a traditional review:
• A recognised methodology, which minimises bias in the outcome of the study, although the publication of bias in the literature does not protect it.
• Can give evidence about the impact of an occurrence over a wide range of empirical methods and settings. If studies provide trusted outcomes, SLRs give evidence that the phenomenon is strong and transferrable.
• The grouping of data using meta-analytic techniques is possible with quantitative studies, enhancing the possibility of noticing actual effects that minor studies are incapable of noticing.
• The research strategy report enables the readers to access the completeness, rigour and repeatability of the process.
• Explicit exclusion and inclusion criteria are required to evaluate each possible primary study.
• The information obtained from each primary study is specified and evaluated, using the quality criteria.

This study explores DG relative to challenges associated with healthcare information systems (HIS) via an SLR. To address the scope of this research study, the study posed one RQ.

How does HIS influence the possibility of DG success?

The RQ above concretise guidelines for the study and inform the research design and methods, namely an SLR.

Methodology

Although this study adopted an SLR, it is important to distinguish between the traditional literature review and the SLR in order to justify the chosen method. The researcher used an SLR to collect secondary data. In comparison to the traditional literature review, an SLR uses a properly defined approach to view the literature for a specific topic (Ryan 2010). Traditional reviews evaluate and summarise a body of literature and draw outcomes for the particular topic in question (Cronin, Ryan & Coughlan 2008). They collect information pertinent to what is known about the topic. Its vital purpose is to provide the reader with a complete familiarity in perceiving current knowledge and featuring the importance of new research. In comparison to a traditional literature review, an SLR uses a clear approach to review literature within a particular subject field.

Traditional reviews try to sum up a number of studies, whereas SLRs use a precise and clear approach to review literature in a particular subject field. Boell and Cecez-Kecmanovic (2010) point out that SLRs are of interests because of the significance they have in the literature search process. Furthermore, an SLR helps to analyse, assess and interpret research pertinent to a specific research topic (Kitchenham 2004). Cronin et al. (2008) argue that the aim of an SLR is to give as full as possible list of all published and unpublished studies on a specific subject field. Kitchenham and Charters (2007) reflect that the purpose of the SLR is to identify primary studies relevant to the RQ through an unbiased search strategy. Okoli (2015) argues that an SLR defines the content and quality of the knowledge of the previous studies that are available. Furthermore, the researchers added that the one factor that distinguishes an SLR from a traditional review is the rigour of the search process. These key guidelines scaffold the SLR process, namely structure (Boell & Cecez-Kecmanovic 2010), a systematically phased approach (Okoli 2015), inclusion and exclusion (Harpur 2018) and quality assessment criteria (QAC) (Inayat et al. 2014).

A four-phase strategy

This study aimed to explore DG relative to challenges associated with HIS, by reviewing guidelines emerging from academic
sources as part of a consolidated SLR. This study applied a four-phase strategy during the SLR: Phase 1 Planning, Phase 2 Selection, Phase 3 Extraction and Phase 4 Execution (Okoli 2015).

**Phase 1 Planning: Identifies the purpose and drafts the protocol**

The first phase consists of two steps, namely to identify the purpose of the SLR and to draft the protocol. The intention of the research is to answer the question posed from the perspective of previously published data on the topic. The RQ determined the focus of the planning stage. The study used a selection of keywords and phrases as search criteria in Google Scholar. The researcher based the search strings and keywords on the RQ to retrieve as many papers as possible. This process resulted in 142 published papers, which included journal articles, conference papers and e-resources. The researcher stored all the screened articles in *Mendeley* for bibliographic and in-text citations (Harpur 2018). *Mendeley* is a tool that allows researchers to manage PDFs, documents and citations through a desktop client version (Parabhoi, Seth & Pathy 2017).

The planning step requires a clear identification of the purpose and intended goals (Okoli 2015). A review protocol provides a clear review of the procedure to be followed where a confined strategy assists to select primary studies and to conduct the SLR (Kitchenham 2004). It supports the replication of the SLR for further studies, and it minimises the bias of the search (Okoli 2015). This study applied the review protocol in the field of DG in HIS in South Africa. From the Planning – Phase 1, iteration 1 led to a group of 142 articles. The Selection – Phase 2 shows the articles that went through the selection process. The Planning phase included the use of *ATLAS.ti V8* for data analysis, followed by the Extraction – Phase 3. Finally, Execution – Phase 4 sets the stage for writing the results of an SLR.

**Phase 2 Selection: Practical screen and search for literature**

This is the second phase of the SLR strategy, which consists of two steps, namely application of a practical screen and the search for literature. This step is also called the screening for inclusion, whereby certain studies were considered for review and other studies were eliminated (Okoli 2015). The study excluded papers not relevant to this SLR through abstract reading. Excluded are non-English publications relevant to HIS, as well as those that are not full papers. Re-reviewing of papers for the second time, using keywords and abstracts, focuses on the RQ and the objective of the study. Screen-published studies on DG in HIS that provided broader information in healthcare, based on titles, abstracts and date were studied. Only literature published from 2010–2020 was analysed to determine the status of DG in HIS within the South African context. A total number of 142 papers were collected.

The study explained and justified the details of the literature according to how they assured the search’s comprehensiveness (Okoli 2015). Boell and Cecez-Kecmanovic (2010) highlighted that a successful search procedure is not one that occurs in high recollection but rather one that results in high accuracy. Searching for literature provides a clear, in-depth understanding of the field of study. Furthermore, it also improves the way to search literature. Pertinent papers from digital databases and the web search engines were covered. The study reviewed studies published from 2010 to 2020. The investigation assisted in providing a picture of the current state of DG in HIS research in South Africa.

**Phase 3 Extraction: Extraction of data and appraisal of quality**

The researcher used *ATLAS.ti V8*, a computer-assisted qualitative data analysis software (CAQDAS) tool, to import the last selection of articles for the SLR. *ATLAS.ti V8* helped the researcher to link different codes of quotations to create networks (Lewis 2015). Selection of articles included themes that emerged from the research topic, problem, RQ and objectives. After the inclusion of all the identified studies for the review, the researcher systematically extracted appropriate information from each study (Okoli 2015). The data extraction was according to the relevant publications, correctly recording the information acquired from the selected publications. The study extracted phrases, words and quotations from the selected articles during the extraction phase.

The researcher screened the extracted articles for exclusion, whereby quality-oriented criteria were used to determine which studies were included and which were not (Okoli 2015). It is not easy to determine values for all used concepts when extracting data, because the values depend on the contents and studies (Staples & Niazi 2007). The researcher created a codebook from the articles in *ATLAS.ti V8* using *Code in vivo* and Open coding. The researcher grouped the codes in themes, sorted the themes in alphabetical order and prefixed each theme with the theme abbreviation and code number, for example (DG01). After the codebook creation, the researcher moved each code to each relevant code group. The researcher gleaned six concepts from the literature sources that led to the proposal of six categories below:

The list of the categorised items is as follows:

- **Category A**: Transformation – install intelligent technologies;
- **Category B**: Effectiveness – implement proper DG;
- **Category C**: Performance – explore contingency factors;
- **Category D**: Adoption – prepare for a new change and evolve;
- **Category E**: Harmonisation – align healthcare system processes; and
- **Category F**: Dynamics – training in emerging AI technologies.

Figure 1 illustrates an example of a network diagram that links the category to a theme as well as items and authors that contributed to each theme.

Figure 1 is the analysis of the code snippets (quotations) from the articles. From the analysis of the codes, the researcher was able to create a visual representation of the themes, items and
associated authors. The blue dotted lines show a link for the authors who contributed to each item. The purple dotted lines illustrate that the items, ‘DG01 Big Data’, ‘DG02 Healthcare Challenges’, ‘DG03 Improved data governance’, ‘DG04 Internet of Things’ are associated with Category C that is directly connected to the theme DG through the link of the red dotted lines.

Selection criteria of the study

The aim of the criteria for study selection is to classify the primary studies that give direct evidence about the RQ (Kitchenham & Charters 2007). Based on the RQ, the study applied the inclusion (I1–I4) and exclusion (E1–E4) criteria. The inclusion for the 142 articles is in Phase 1 Selection.

The study applied the following inclusion criteria to decide whether the article should be included in the study:

- I1: Addresses the use of DG in HIS;
- I2: Pertains to healthcare contexts;
- I3: Includes articles published between 2010 and 2020;
- I4 Having an abstract available; and
- I5: Includes an academic, peer-reviewed journal or a conference proceeding.

The criteria below serve as the base of the exclusion of articles:

- E3: Is not a suitable recent publication;
- E4: Does not have an abstract; and
- E5: Is not an academic or a peer-reviewed journal.

Iteration 1 included several scans of extracted literature sources. The study followed several methods, which includes snowballing. These methods are a foundation of the first selection defined in Phase 1 – Planning, which consists of 142 studies. During iteration 2, the application of exclusion criteria led to a reduced count of 38 articles. Iteration 3 bridged the noted gaps connected to competitive strategy, DG, DG contingency model, performance strategy and process harmonisation, which led to the addition of nine articles, resulting in 46 publications.

Table 1 is a summary of three iterations (1–3) that include the exclusions as well as the addition of articles. Iterations 1 and 3 consist of a single activity, whilst iteration 2 consists of five exclusion criteria.

Quality assessment of the study

The quality evaluation serves to check whether the end search results have been adequate and offers support for the scope of the review. Kitchenham and Charters (2007) informed that on top of the inclusion or exclusion criteria, it is vital to assess the quality of primary studies:

- To provide a thorough view of exclusion and inclusion criteria;
- To explore whether quality differences give an explanation for differences in the outcome of the study;

FIGURE 1: Network diagram for the category, themes, items and authors.

Source: Adapted from Harpur, P.A., 2018, ‘A framework for ad hoc mobile technology-enhanced learning in a higher education context’, Doctoral dissertation, Cape Peninsula University of Technology
According to Boell and Cecez-Kecmanovic (2010), SLRs are researchers’ methods and systematic literature review detailed report so that other researchers can reproduce the selected publications. In an SLR, the process needs a full synthesis of studies and the writing of the review. Facts extracted from the studies are combined by using a qualitative or a quantitative technique or both (Okoli 2015).

Execution is the fourth phase, comprising two steps, namely the synthesis of studies and the writing of the review. Facts are extracted from the studies are combined by using a qualitative or a quantitative technique or both (Okoli 2015). This step collects, combines and summarises the results of the selected publications. In an SLR, the process needs a full detailed report so that other researchers can reproduce the review results (Okoli 2015).

Data collection methods – In relation to other researchers’ methods and systematic literature review

According to Boell and Cecez-Kecmanovic (2010), SLRs are of specific interest for the significance they have on the process of literature searching. Okoli (2015) points out that researchers when doing research for its predetermined steps that allow the evaluation of search integrity currently choose an SLR. A researcher needs to consider the following important steps when doing an SLR (Gough, Oliver & Thomas 2012). Researchers propose that the following structured phases from Okoli (2015) and Kitchenham (2004) are relevant to the successful construction of SLRs:

- **Identify the purpose:** to identify clearly the intended goals and purpose of the study.
- **Draft the protocol:** to confine a strategy for the selection of primary studies.
- **Search for literature:** to clearly explain and justify literature details to guarantee search completeness.
- **Apply practical screen:** to determine which studies will be included or excluded.
- **Appraise quality:** to rate papers for elimination because of insufficient quality.
- **Extract data:** the applicable data will be systematically extracted from each study;
- **Synthesis of studies:** to combine facts extracted from studies using qualitative or quantitative techniques.
- **Writing the review:** to report the outcome of the review in detail.

This SLR adopted a three-prong strategy. Firstly, it focused on HIS. Secondly, it addressed DG issues and challenges. Finally, it explored DG guidelines via an explicit collection of relevant sources. Finally, the review navigated recently published sources regarding three components of the DG contingency model, namely performance strategy, competitive strategy and process harmonisation. The SLR method adopted in this study served to gather, analyse and interpret previously published data. This study is based on the eight-step approach recommended by Okoli (2015) outlined in Figure 2 below:

The four phases depicted above contain eight steps necessary when conducting an SLR. These phases are planning, selection, extraction and execution (Okoli 2015).

**Research tools**

Table 2 illustrates the research tools used in the study for data collection to achieve the research purpose.

The study used a thematic analysis for data analysis because it followed an interpretive approach. The thematic analysis analyses the categorisations and displays themes that are connected to data. It explains data in detail whilst dealing with various subjects through interpretations (Alhojailan & Ibrahim 2012). It provides a description and understanding of answers through discovering patterns and creating themes. Themes came from the secondary data through an inductive approach.

**Discussion of findings**

The six themes that emerged during the study are transformation – competitive advantage, effectiveness – data
tackle these challenges (Gujral, Shivarama & Mariappan 2019). Although digital transformation has begun in many healthcare organisations, few of them have reached maturity (Gopal et al. 2019). The improvement of technology would lead to a greater competitive advantage.

Effectiveness – Data governance

The entire organisation needs to align its goals to DG. Organisations are increasingly developing advanced DG capabilities (Janssen et al. 2020). The only way to solve the data problem is the implementation of effective DG (Al-ruith, Benkhelifa & Hameed 2018). Data governance assists organisations to ensure data quality and to maintain the value of data as an organisational asset. Many companies see DG as a promising approach to ensuring data quality (Otto 2011a, 2011b). Successful DG may answer certain data challenges of many organisations.

Performance – Data governance contingency model

A contingency is an upcoming event, which is potential but not easily predicted. The most appropriate contingency factors are recognised as culture, industry, maturity and structure (Pereira & Silva 2012). Organisations with high contingency fit are less vulnerable to deviation from the formation of organisation performance (Volberda et al. 2012). Data governance is necessary to safely manage organisational data and success (Lee, Zhu & Jeffery 2018b). Organisations need greater levels of innovativeness to be successful (Boso et al. 2013).

Adoption – Healthcare information systems

A HIS serves as a bridge between IS and the business processes in healthcare in order to bring better healthcare services (Almunawar & Anshari 2012). The proper collection and utilisation of electronic healthcare record (EHR) is the foundation of the digital healthcare (Yang et al. 2015). Electronic healthcare record serves as the main driver of modern healthcare. The effect of technological, social and political factors changed the nature of the healthcare industry eventually (Almunawar & Anshari 2012). This change led organisations to seek organisational change through digital transformation.

Dynamics – Performance strategy

Performance strategy is a method that organisations use to help implement their strategy into their organisation to achieve all goals. The digitisation of patient records opens rich possibilities for medical professionals (Atasoy, Greenwood & Mccullough 2019). The strategies for sharing information on goals, organisational structure and overall performance have a significant positive effect on performance. Managerial competences play a very important role on organisational performance (Vainieri et al. 2019). Healthcare professionals and organisations must be prepared to change (Wiljer & Hakim 2019). Develop human and organisational governance, performance – data governance contingency model, adoption – healthcare information systems, dynamics – performance strategy and harmonisation – process harmonisation.

Transformation – Competitive advantage

Competitive advantage is the feature that identifies the organisation to outdo its competitors. Increasing the competitive advantage is one of the challenges in healthcare (Saeidi et al. 2019). The digital transformation is a key to

---

TABLE 2: Research tools used for data collection adapted from Harpur (2018).

<table>
<thead>
<tr>
<th>Research tool</th>
<th>Purpose</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATLAS.ti V8 <a href="http://atlasti.com/">http://atlasti.com/</a></td>
<td>A qualitative data analysis tool used for coding and linking of quotations to create networks.</td>
<td>Used for analysing literature sources.</td>
</tr>
<tr>
<td>Google Drive <a href="https://www.google.com/">https://www.google.com/</a></td>
<td>A file storage that was developed by Google.</td>
<td>Allows users to store and backup files in Google servers.</td>
</tr>
<tr>
<td>Mendeley <a href="https://www.mendeley.com/">https://www.mendeley.com/</a></td>
<td>A desktop software used for arranging research articles.</td>
<td>Helps to import bibliographic material from Google Scholar.</td>
</tr>
</tbody>
</table>


---

FIGURE 2: A systematic guide to literature review development (Okoli 2015).
skills to adopt cultures and accept changes that promote readiness to face unexpected and expected challenges (Alsharif et al. 2018).

**Harmonisation – Process harmonisation**

Process harmonisation refers to organising and applying standards for a business process to achieve targeted business requirements. Data harmonisation is an important intervention to strengthen health system’s functioning (Schmidt et al. 2018). It is a key intervention to give strength to the functioning of health systems (Schmidt et al. 2018). Data harmonisation enhances the accessibility, production and usability of standard health information for service management and clinical decision-making. Harmonised data quality assessment terms, methods and reporting practices can establish a common understanding (Kahn et al. 2016).

**Limitations of the study**

This study was limited to the period 2010–2020; thus, relevant research studies conducted before this period were excluded, and therefore, important and pertinent information could have been missed. This study was limited to data collected from digital databases and web search engines and thus could have missed relevant research in public libraries and university databases. An SLR is the only method used to gather data and that helped the researcher to search for titles, abstracts, keywords and phrases.

**Conclusion**

The objective of this study was to understand the features of HIS, acquire information about DG success and understand the influence noted on DG. Moreover, the study also analysed competitive strategy, DG contingency model, HIS, process harmonisation and performance strategy. The findings from the study revealed that many organisations have realised that the only method to fix the data problem is the implementation of effective DG.

The study also revealed that EHR is the main driver of the digital healthcare. Therefore, the accurate collection and utilisation of EHR have become the foundation of the digital healthcare. With the rise of cloud computing and increased adoption, DG has gained interest amongst healthcare professionals. Therefore, the competencies and skills from various IT experts and businesses should synchronise. In conclusion, healthcare organisations should align their goals to DG.

**Acknowledgements**

Nomputumo Ngesimani acknowledges the CPUT for their support of this study as part of the master’s degree. The family is acknowledged for all the support that they gave the author during the time of compiling this research article together.

**Competing interests**

The authors declare that they have no financial or personal relationships that may have inappropriately influenced them in writing this article.

**Authors’ contributions**

The first author N.L.N., who is the main author, contributed in conceptualisation, methodology, formal analysis, investigation, writing original report and writing review and editing. The second author E.R. contributed to supervising me on the selection of the methodology I used. The third author P.A.H. contributed in supervising me on the steps of writing the methodology I used, data analysis and writing of the article.

**Ethical considerations**

This article followed all ethical standards for research without direct contact with human or animal subjects.

**Funding information**

The authors received no financial support for the research, authorship and/or publication of this article.

**Data availability**

Data sharing is not applicable to this article as no new data were created or analysed in this study.

**Disclaimer**

The views and opinions expressed in this article are those of the authors and do not necessarily reflect the official policy or position of any affiliated agency of the authors.

**References**


Kitchenham, B., 2004, Procedures for performing systematic reviews, Keele University, Keele.


Okoli, C., 2015, 'A guide to conducting a standalone systematic literature review', Communications of the Association for Information Systems 37(1), 43. https://doi.org/10.17705/1CAIS.03743


Otto, B., 2011b, 'Organising data governance: Findings from the telecommunications industry and consequences for large service providers', Communications of the Association for Information Systems 29(1), 1. https://doi.org/10.17705/1CAIS.02903


Piper, R.J., 2013, 'How to write a systematic literature review: A guide for medical students', National AMR, Fostering Medical Research 1, 1–8.


Russom, P., 2012, Seven reasons why master data management needs data governance, TDWI, Renton.

Ryan, G., 2010, Guidance notes on planning a systematic review, James Hardiman Library, National University of Ireland Galway, Galway.


Siddaway, A., 2014, 'What is a systematic literature review and how do I do one', University of Stirling 1(1), 1–3.


