Integration of Industry 4.0 technologies into Lean Six Sigma DMAIC: a systematic review

Tanawadee Pongboonchai-Empla, Jiju Antonyb, Jose Arturo Garza-Reyes\*c, Tim Komkowskia & Guilherme Luz Tortorellad

a Edinburgh Business School, Heriot-Watt University, Edinburgh, Scotland; bDepartment of Industrial and System Engineering, Khalifa University, Abu Dhabi, United Arab Emirates; cCentre for Supply Chain Improvement, University of Derby, Derby, UK; dUniversidade Federal de Santa Catarina, Florianópolis, Brazil

\*Corresponding author: J.Reyes@derby.ac.uk

Integration of Industry 4.0 technologies into Lean Six Sigma DMAIC: a systematic review

Integration of Industry 4.0 technologies into Lean Six Sigma DMAIC: a systematic review

This review examines which Industry 4.0 (I4.0) technologies are suitable for improving Lean Six Sigma (LSS) tasks and the benefits of integrating these technologies into improvement projects. Also, it explores existing integration frameworks and discusses their relevance. A quantitative analysis of 692 papers and an in-depth analysis of 41 papers revealed that “Analyse” is by far the best-supported DMAICs phase through techniques such as Data Mining, Machine Learning, Big Data Analytics, Internet of Things, and Process Mining. This paper also proposes a DMAIC 4.0 framework based on multiple technologies. The mapping of I4.0 related techniques to DMAIC phases and tools is a novelty compared to previous studies regarding the diversity of digital technologies applied. LSS practitioners facing the challenges of increasing complexity and data volumes can benefit from understanding how I4.0 technology can support their DMAIC projects and which of the suggested approaches they can adopt for their context.

Keywords: Lean Six Sigma, Industry 4.0, DMAIC 4.0, Big Data Analytics, Data Science

# List of abbreviations

|  |  |
| --- | --- |
| Abbreviation | Meaning |
| 5S | Sort, Set in order, Shine, Standardise, Sustain |
| AI | Artificial Intelligence |
| ANN | Artificial Neural Network |
| ANOVA | Analysis of Variance |
| BDA | Big Data Analytics |
| DOE | Design of Experiments |
| DFSS | Design for Six Sigma |
| DMAIC | Define, Measure, Analyse, Improve, Control |
| DM | Data Mining |
| I4.0 | Industry 4.0 |
| IoT | Internet of Things |
| LSS | Lean Six Sigma |
| ML | Machine Learning |
| PCA | Principal Component Analysis |
| PLS | Partial Least Squares |
| PM | Process Mining |
| RQ | Research Question |
| RSM | Response Surface Method |
| SLR | Systematic Literature Review |
| SMED | Single-Minute Exchange of Die |

# Introduction

Lean Six Sigma (LSS) is a business process improvement strategy that combines the virtues of Lean and Six Sigma and has been advocated by many academics over the past two decades. The Lean toolkit, which includes, among others, 5S, Just-in-Time, Value Stream Mapping, SMED, Kanban and Poka-Yoke, is based on experience and judgment and focuses on increasing productivity while knowing the goal of improvement (Antony, Snee, and Hoerl 2017). In contrast, Six Sigma tools rely on data-driven, statistical techniques, for example, Hypothesis Testing, ANOVA, Regression Analysis, and Statistical Process Control (SPC) (Hoerl 2001, Goh 2011) while aiming to reduce variations and defects to improve process performance (Snee 2010, Antony 2011). The DMAIC (Define, Measure, Analyse, Improve, Control) roadmap provides a structured approach to achieving these goals and is an integral part of Six Sigma. The benefits of the integrated approach commonly cited in the existing literature are (1) reduced cost; (2) reduced errors/defects (3) improved product/service quality; (4) reduced non-value-added steps; (5) improved cycle time; (6) increased customer satisfaction; (7) improved employee morale (Albliwi, Antony, and halim Lim 2015, Singh and Rathi 2019).

However, in the age of the fourth industrial revolution, new challenges have arisen. For example, obtaining process robustness is often difficult in multiple-process manufacturing because hundreds of factors can impact the output, and part-specific quality constraints further increase complexity. Furthermore, the vast amounts of heterogeneous data recorded by interconnected machines and systems imposed by the fourth industrial revolution require a paradigm shift of the LSS methodology, as traditional tools and software programs may not be effective any longer. George et al. (2019, 102) argue that companies must meet contemporary challenges, such as increasing complexity and huge data volumes, to remain competitive and, therefore, should fully exploit the spectrum offered by digital technologies. There is ample evidence of LSS limitations that practitioners in the field struggle to address (Pongboonchai-Empl, Stemann, and Antony 2021). Some crucial limitations such as unavailability of data (Albliwi et al. 2014), long implementation cycles (Sony, Naik, and Therisa 2019), incorrect selection and prioritisation of projects (Snee 2010), and unsustainable results (Aboelmaged 2011) are strong motivating factors for leveraging Industry 4.0 (I4.0) technologies. Therefore, the first research question is:

* RQ1: Which I4.0 technologies were applied in conjunction with LSS tools in previous studies?

Representing the fourth industrial revolution, Industry 4.0 (I4.0) is understood as a manufacturing environment in which sensors and robots enable autonomous systems. The core idea of I4.0 is the “convergence of real and virtual production in all sectors” (Deuse et al. 2020). Although no universal definition of I4.0 has been agreed upon, it is typically associated with cyber-physical systems, Internet of Things (IoT), big data analytics (BDA), cloud computing, additive manufacturing, augmented reality, autonomous robots, cybersecurity and horizontal & vertical integration (Rojko 2017, Ghobakhloo 2018, Butt 2020). While the connectedness between machines and between humans and machines generates immense data volumes, data science provides techniques to gain insights into processes and draw conclusions. Additionally, cloud computing enables the storage and analysis of large amounts of data. I4.0 technologies and data science techniques offer new opportunities for data-driven quality improvement strategies, such as LSS.

The integration of LSS with I4.0 is an emerging trend (Antony et al. 2019, Antony and Sony 2020) that enjoys growing attention in academia. Several researchers have already followed this trend, and numerous empirical studies exist on this topic, for instance, by Chiarini and Kumar (2021) or Koppel and Chang (2021). Furthermore, several literature review papers have been published on this topic. However, they either focus on organisation theories instead of practical implementation or investigate a particular I4.0 technology or LSS tool. For example, Gupta, Modgil, and Gunasekaran (2020) composed a comprehensive review of theory building and empirical works involving Big Data integration in LSS. In contrast, the reviews provided by Wang et al. (2021) and Widjajanto, Purba, and Jaqin (2020) are limited to the integration of I4.0 with one LSS tool each, namely, VSM and Poka Yoke. Appendix A outlines the results of the analysis of the review papers.

In conclusion, existing review papers do not provide a comprehensive overview of how I4.0 technologies support LSS tools and DMAIC phases or remain at a high-level. Given the demands that today’s world places on operational excellence teams, the potential of modern technologies must be extensively explored. Therefore, this review seeks to systematically collect evidence from previous research and provide a clearer picture of support opportunities. The systematic literature review methodology is suitable for consolidating the evidence published on a particular topic (Tranfield, Denyer, and Smart 2003). To investigate the applicability and impact of I4.0 tools for LSS projects, we pose the following research questions:

* RQ2a): How have I4.0 technologies enhanced LSS projects?
* RQ2b): What were the benefits of integrating I4.0 technologies in LSS projects?

In traditional LSS research, frameworks were commonly developed to facilitate the implementation of the proposed solutions for organisations. Especially when combining two different fields, a framework is crucial to clarify their interactions (Yadav, Seth, and Desai 2017). While frameworks that integrate I4.0 tools and LSS have emerged from several reviews, these are also either only high-level or limited to a specific technology. Details are presented in Appendix A. RQ3 assists in gaining an overview of existing frameworks:

* RQ3: Which frameworks or roadmaps for integrating I4.0 technology in LSS DMAIC phases have been presented?

Similarly to previous authors (Sordan *et al.*, 2021; Vinodh *et al.*, 2021), we conducted a systematic review to capture the existing approaches for integrating I4.0 technologies in LSS projects and their benefits. Likewise, we also analysed the limitations and research gaps to provide future direction for research. In addition, we aim to examine whether existing approaches are transferable to other environments to help practitioners find a suitable approach for their environment. In addition, Finally, answering questions 1 to 3 will enable the development of a DMAIC 4.0 framework that is both comprehensive and detailed in terms of tool suggestions.

This article is organised as follows. The next section provides the methodological foundation of this systematic literature review. Section 3 presents the SLR findings, followed by a discussion of the three-stage analysis in section 4. Section 5 presents a proposed DMAIC 4.0 framework resulting from this review. Finally, Section 6 concludes this paper with contributions, implications, and a future research agenda.

# Methodology

The systematic literature review (SLR) was initially developed for medical researchers and practitioners and is a scientific methodology to investigate existing literature in a particular field (Tranfield, Denyer, and Smart 2003, Briner and Denyer 2012). The strength of the SLR method is that it enables researchers to develop an evidence base and foundation of knowledge by imposing a predefined protocol for systematically extracting knowledge from previous studies (Tranfield, Denyer, and Smart 2003, Briner and Denyer 2012, Williams et al. 2020). According to the authors, SLR minimises researcher bias and adds more rigour than a traditional review, where studies are selected purposively and thus cherry-picked. There are several approaches to conducting SLR, and various disciplines seem to have established a preferred methodology that pioneers in the respective field have adopted. This SLR follows the phases proposed by Tranfield, Denyer, and Smart (2003), who introduced this methodology for management research. The need for gathering evidence in disciplines where qualitative studies are common made it necessary to develop SLR methods that accommodate both quantitative and qualitative analyses and syntheses (Tranfield, Denyer, and Smart 2003, Rousseau, Manning, and Denyer 2008, Okoli and Schabram 2010). Figure 1 illustrates the review process of this study.

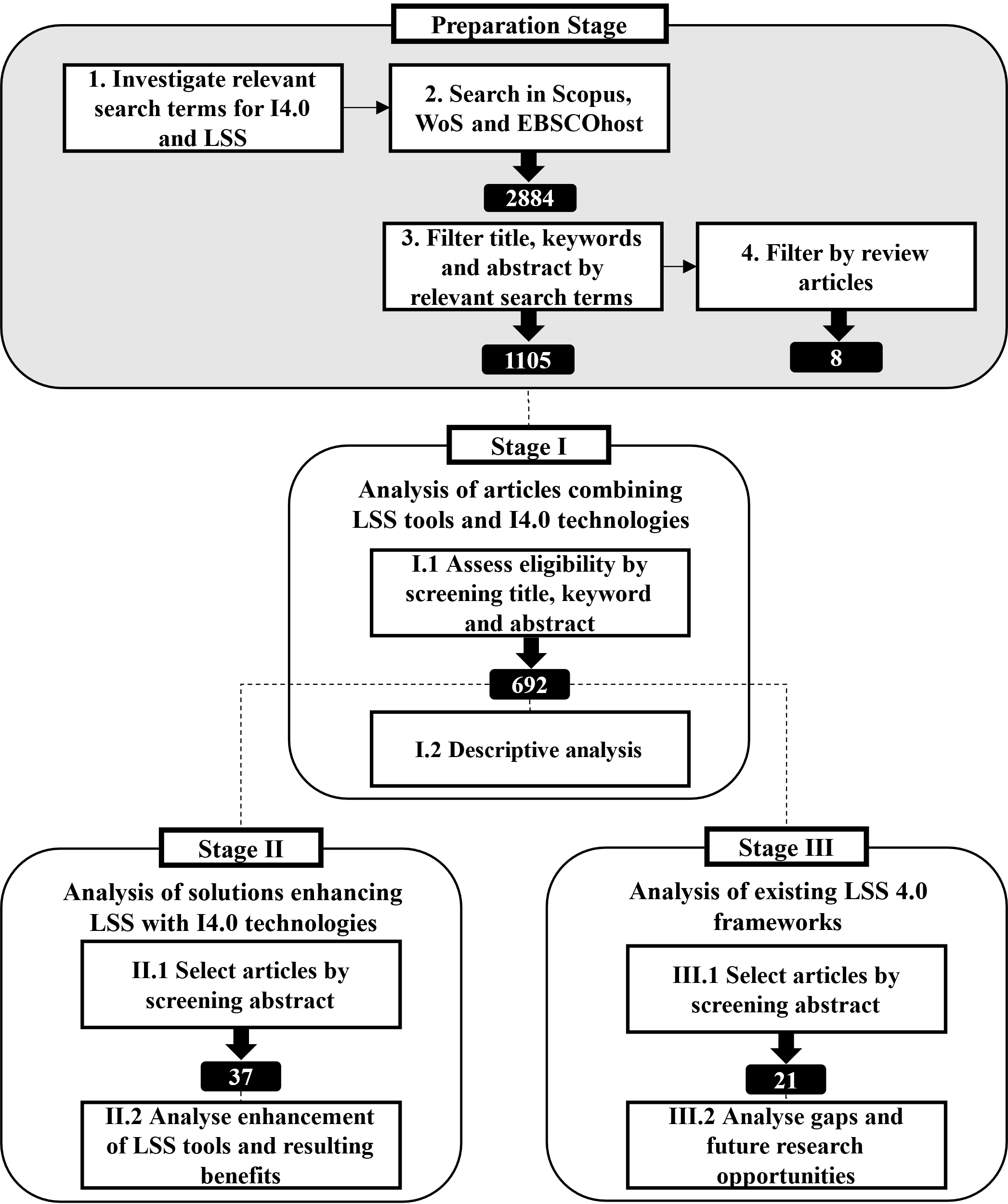


Figure . SLR process (inspired by Tranfield, Denyer, and Smart 2003 and Goienetxea Uriarte, Ng, and Urenda Moris (2020)).

Various definitions of I4.0 can be found in the literature, depending on the country and the time of publication. Therefore, the relevant search terms were extracted from 5213 journal articles related to I4.0. An analysis based on text mining techniques in the analytics software KNIME yielded a list of I4.0 technologies used in conjunction with LSS tools (Table 1). The LSS terms were adopted from an empirical investigation of the most frequently used LSS tools by Uluskan (2019). According to the researchers, the empirical findings also largely correspond to the most frequently used LSS tools reported in the literature.

Table . List of search terms for the systematic literature review.

|  |  |
| --- | --- |
| **LSS terms** | **Industry 4.0 terms** |
| **Core search terms:**   |  |  | | --- | --- | | Six Sigma | DMAIC | | |  |  | | --- | --- | | Big data | Machine Learning | | Industry 4.0/Industrie 4.0 | Analytics | | Digital Transformation | Artificial Intelligence | | Predictive Maintenance | Neural Network | | Preventive Maintenance | Internet of Things | | Data Mining | Data Science | | Smart Manufacturing | Robotics | | Process Mining | Cyber Physical System | |
| **Basic tools (most frequently used):**   |  |  | | --- | --- | | Pareto chart/Pareto diagram | Process sigma | | Brainstorming | Process map/Process flow diagram | | Fishbone diagram/Ishikawa | Measurement system analysis | |
| **Analysis tools (most frequently used):**   |  |  | | --- | --- | | Voice of the Customer | Box Plots | | Operational definition | ANOVA | | Data collection plan | FMEA | | Run chart | PDCA | | Control chart | Mistake proofing | | Data normality test | Poka yoke | |

The search was performed on October 26th, 2021. Papers published after that are not in the scope of this review. The query in three leading scientific databases, Scopus, Web of Science and EBSCOhost, yielded 2884 unique articles. Subsequently, the title, abstract and keywords were screened to assess the eligibility and define the subsets for each stage. Inclusion and exclusion criteria (Table 2) add transparency to the selection of relevant studies (Briner and Denyer 2012) and their allocation to each stage.

Table . SLR Inclusion and Exclusion Criteria (author’s compilation).

|  |  |  |
| --- | --- | --- |
|  | **Inclusion Criteria** | **Exclusion Criteria** |
| **Stage I** | (+) Articles published in academic or specialist journals | (-) Conference papers, book chapters, grey literature, websites |
| (+) Research papers containing most frequently used LSS tools in combination with data-driven technologies | (-) Articles not related to process or quality improvement |
| (+) Articles in English or German | (-) Articles presented in other languages |
|  |  |  |
| **Stage II**  (subset of Stage I) | (+) Articles in which Six Sigma, DMAIC or PDCA projects are enhanced by data-driven technologies | (-) LSS articles without reference to data-driven technologies or frameworks |
| (-) I4.0 articles not related to Lean Six Sigma |
| (-) Articles solely focusing on DFSS and not DMAIC |
| (+) Articles presenting approaches for enhancing or complementing Lean Six Sigma with I4.0 technology | (-) Articles in which LSS is used as a methodology to implement or enhance I4.0 approaches  (-) Literature Review papers |
|
|  |  |  |
| **Stage III**  (subset of Stage I) | (+) Articles presenting a framework, model, approach, roadmap or method integrating LSS with I4.0 | (-) Articles not covering all phases of DMAIC |

The objective of stage I is to provide a quantitative analysis of studies that integrated tools and techniques from both LSS and I4.0. This method makes it feasible to analyse a large number of papers covering a long time period (Puram and Gurumurthy 2021). Papers selected for stages II and III are analysed in-depth using NVivo12. While Stage II intends to investigate approaches incorporating I4.0 technologies in LSS DMAIC projects and their benefits, Stage III aims to analyse existing frameworks and their limitations. Due to the qualitative nature of literature reviews, steps should be taken to prevent researcher bias and promote rigour (Tranfield *et al.*, 2003; Rousseau *et al.*, 2008). To this aim, the first author kept a research diary to recapitulate the decisions made during the research process if questions arose. Also, a process manual was defined, so data collection and analysis steps were executed consistently. Other authors were involved in the systematic review to add validity to the study.

# Findings

The screening process resulted in 692 papers for stage I, 37 for stage II, and 21 for stage III. All other papers were excluded, for example, if they only referred to LSS or solely to I4.0 technologies.

## Stage I: Articles integrating LSS with I4.0 technologies

Stage I contains research papers that involve both LSS tools and I4.0 technologies, i.e. one of the LSS tools plus an I4.0 related search term appears in the title, keyword or abstract. This stage aims to provide an overview of which tools and technologies of these two independent fields were already used collaboratively. A text mining software, VOSviewer (van Eck and Waltman 2010), is used to visualise relevant topics and how they relate to each other. Figure 2 illustrates which key terms are jointly mentioned in the 692 papers, and the size of the dots represents the term’s frequency. The related terms are clustered in colour-coded groups. The red and the blue group represent two of the most dominant I4.0 themes, machine learning and artificial neural networks (ANN). Unsurprisingly, the red machine learning cluster terms are closely related to the terms from the blue ANN and the yellow data mining cluster. For example, “classification” and "pattern recognition” are connected to “artificial neural network”, while “classification” and “statistical process control” are linked to “data mining”. The predominant LSS tool is ANOVA, which is saliently related to ANN as part of the blue cluster. This relationship deserves further investigation in a future study. Similarly, other complex, traditional LSS tools such as RSM and Taguchi appear to leverage ANN techniques. The green cluster, on the other hand, includes more general terms such as “Industry 4.0,” “Big Data,” and “artificial intelligence,” along with “lean six sigma,” “six sigma,” and “dmaic,” suggesting that these are higher-level, conceptual works rather than case studies that apply specific tools and technologies. Overall, the repeated appearance of themes in different clusters shows they are closely interrelated.

Diagram

Description automatically generated

Figure . Co-occurrence network of frequent keywords in papers combining LSS tools and I4.0 technologies (developed for research).

The 692 articles that emerged from the selection process exhibit various combinations of topics from the two disciplines, as shown in Figure 3. Not all of them explicitly mention LSS or Six Sigma, but instead, LSS tools are used for quality or process improvement. Articles mentioning more than one technology were assigned to the most prevalent topic, so there are no duplicate counts. The legend shows LSS tools that occurred in more than ten articles. The predominance of the combination of ANOVA with ANN is salient, as the combination of control charts and ANN. Research that included additional terms besides ANN, such as “data mining”, “machine learning”, and “big data”, was accounted for as ANN because ANN is a specific technique. Although other I4.0 terms, such as Process Mining, occur far less often, they can play an essential role in stage 3, where the articles in which I4.0 technology is used to enhance LSS will be investigated in depth. In contrast, Data Mining alone (without ANN) plays an important role in supporting LSS tools, such as ANOVA and control charts, while a closer look in stage 3 will reveal that this technique is not applied across all DMAIC phases. Another important point to notice is that studies labelled as 6s/DMAIC do not only cover a specific LSS tool but entail a complete Six Sigma or DMAIC project. These are the papers selected for stage 2, and they are also partially included in stage 3. Also noteworthy are the 6s/DMAIC items in the data mining bar which represent articles integrating DMAIC and the CRISP-DM (Cross Industry Standard for Data Mining) model.

Figure . Articles integrating I4.0 Technologies with LSS Tools (developed for research).

Figure . Publications per year by I4.0 technology (developed for research).

Figure 4 indicates that technologies like ANN, DM, ML and AI have been studied in conjunction with LSS tools since 2001, long before the term ‘Industry 4.0’ was coined. In contrast, the first BDA, IoT, and Robotics articles emerged around 2012. Although PM first appeared in academia in 2003 (van der Aalst et al. 2003), it took over a decade to be combined with LSS. In general, there has been increasing interest over the past few years in combining LSS tools and smart technologies in research.

## Stage II: Articles presenting approaches to enhancing LSS with I4.0

Stage II includes 37 studies in which I4.0 technologies enhanced LSS projects. Qualitative content analysis techniques were applied to answer RQ2a), RQ2b) and RQ3. Before discussing the in-depth analysis results, here are some facts and figures on the articles that belong to this stage. 20 out of 37 articles were published in business management journals, followed by computer science (9), engineering (8) and social sciences (1). Three articles were published in each of the journals “International Journal of Lean Six Sigma” and “Production Planning & Control”. Two articles were found in each of the journals “Computers & Industrial Engineering”, “International Journal of Six Sigma and Competitive Advantage”, “Processes” and “ZWF Zeitschrift fuer Wirtschaftlichen Fabrikbetrieb". The remaining 23 articles were found across 23 different journals. The research was predominantly conducted in the manufacturing sector, but there was also interest in healthcare, higher education, and other sectors (Table 3).

Table . Stage II articles by sector.

|  |  |
| --- | --- |
| **Sector** | **Count** |
| Manufacturing | 26 |
| Healthcare | 5 |
| Higher Education | 2 |
| Finance | 1 |
| Service | 1 |
| Supply Chain | 1 |
| Unknown | 1 |

Regarding which I4.0 technology or concept are predominantly covered by the articles ANN, together with ML, outperforms the others, followed by DM/CRISP-DM and BDA (Figure 5). Process Mining and IoT seem to play a minor role when comparing the number of papers by technology. However, the analysis by DMAIC coverage shows that PM can consistently support each phase and takes a similarly important role as DM and ANN (Figure 6).

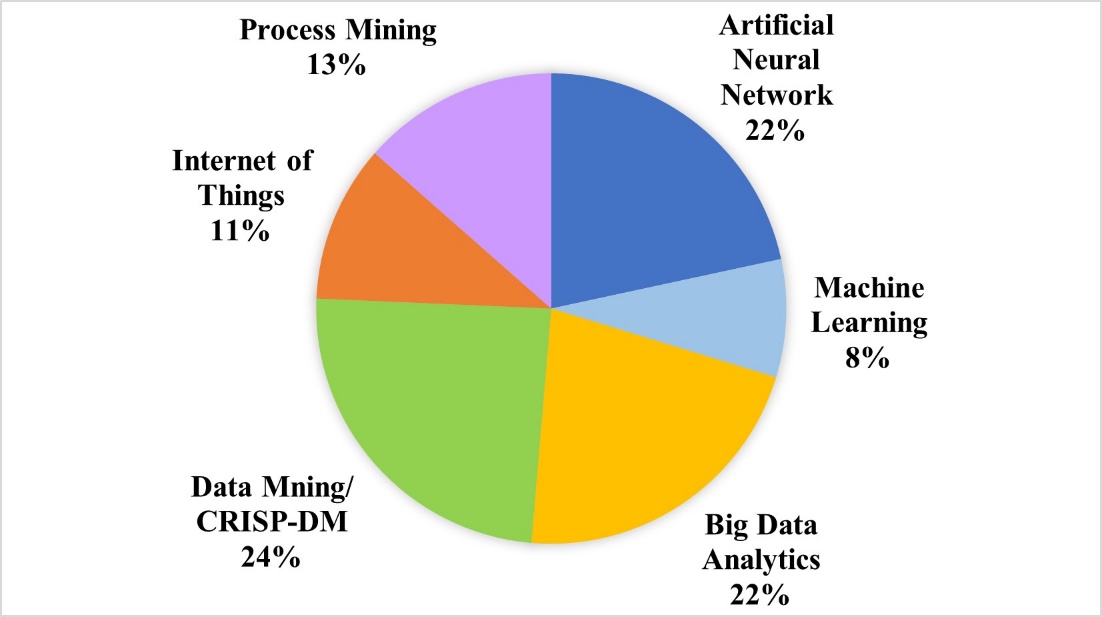


Figure . I4.0 Technologies enhancing LSS (developed for research).

When considering which DMAIC phase is enhanced best by I4.0 technologies (Figure 6), Analyse is far ahead, followed by Control and Improve. Define and Measure are far behind, although digital support also has much potential here. For example, Text Mining and Video Mining, which are considered BDA techniques, can provide valuable input for Voice of the Customer (Dogan and Gurcan 2018, Gupta, Modgil, and Gunasekaran 2020). Currently, most case studies seem to see more value-add in analysis tasks than in project definition and measure tasks. However, this does not imply that those phases cannot be enhanced in various ways. Rather, it is because the challenges for the other phases, especially Analyse, are presently greater and require support more urgently. The benefits of data collection through sensors and interconnected machines (and humans) are immense and can significantly alleviate and speed up the Measure Phase. Also, knowing what customers want without having to conduct market research is of immense value.

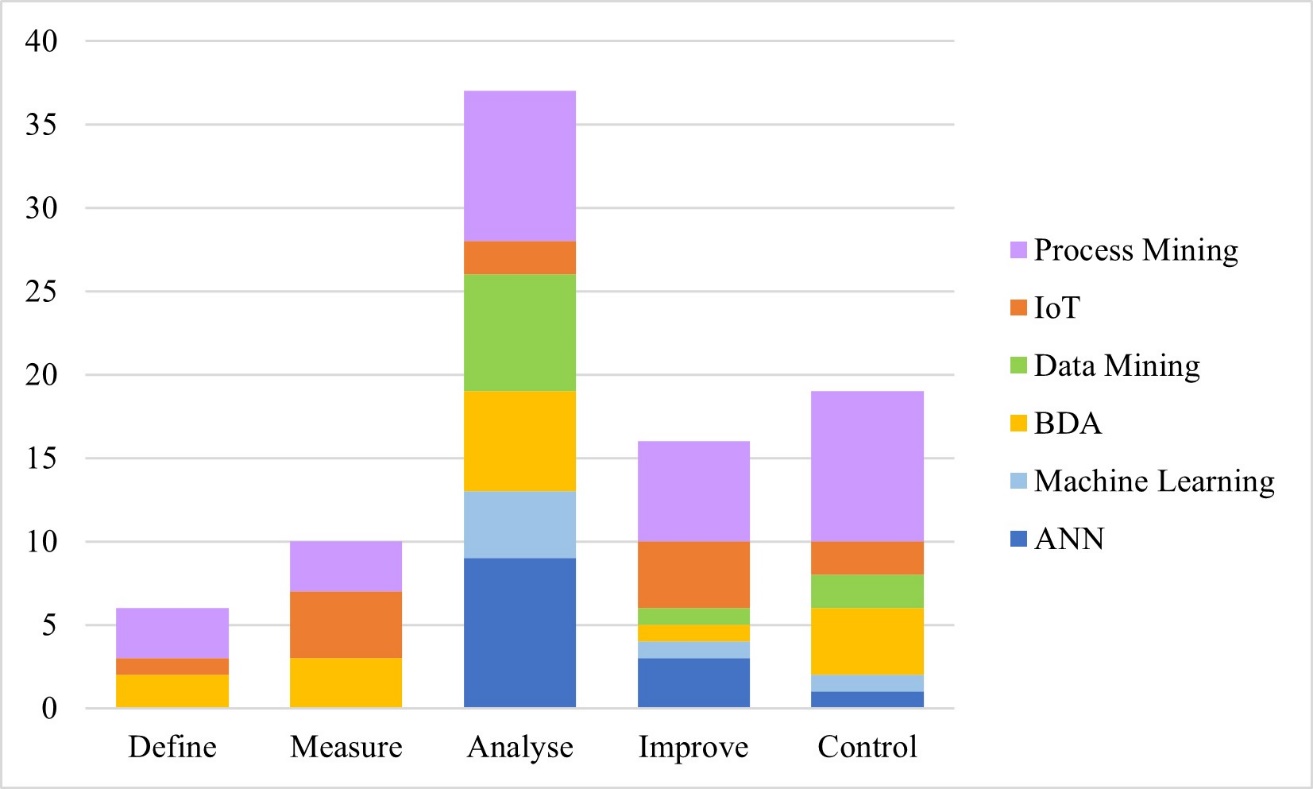


Figure . DMAIC phases enhanced by I4.0 technologies (developed for research).

As there are no official rules for assigning tools to DMAIC phases, the researchers allocated them according to the content of each study. Process Mining takes a similarly important role as Data Mining and Artificial Neural Networks when assessing the supported tools or tasks per phase. Table 4 shows which LSS tools or tasks were enhanced or potentially replaced by which I4.0 technologies. The specific tools and techniques applied are listed in Appendix B.

Table . LSS tools and tasks supported by I4.0 technologies.

| LSS tool/task | I4.0 technology | References |
| --- | --- | --- |
| 5S | Internet of Things | Chiarini and Kumar (2021) |
| Cause and Effect diagram | Data Mining, Big Data Analytics | Hsiao et al. (2016), Fahey et al. (2020) |
| Control Chart | Big Data Analytics | Belhadi et al. (2021), Koppel and Chang (2021) |
| Design of Experiments | Machine Learning, Data Mining, Artificial Neural Network | Chi et al. (2007), Gaudard et al. (2009), Ghosh and Maiti (2014), Mishra and Rane (2019), Morlock and Boßlau (2021) |
| Data Analysis | Big Data Analytics | Laux et al. (2017), Shanshan et al. (2021) |
| Data Collection | Internet of Things, Big Data Analytics, Process Mining | Arcidiacono and Pieroni (2018), Shivajee et al. (2019), Abed et al. (2020), Graafmans et al. (2021), Koppel and Chang (2021), Kregel et al. (2021), Ramires and Sampaio (2021) |
| Data Visualisation | Internet of Things | Arcidiacono and Pieroni (2018) |
| Multivariate Regression | Process Mining | Ramires and Sampaio (2021) |
| Parameter Tuning | Big Data Analytics | Belhadi et al. (2021) |
| Poka Yoke | Artificial Neural Network, Internet of Things | Abed et al. (2020), Chiarini and Kumar (2021), Tay and Loh (2021) |
| Principal Component Analysis | Artificial Neural Network, Machine Learning | Sen (2015), Palací-López et al. (2020) |
| Process Map | Process Mining, Internet of Things | Shin et al. (2019), Graafmans et al. (2021), Kregel et al. (2021), Ramires and Sampaio (2021), Singh et al. (2021), Tay and Loh (2021) |
| Process Monitoring | Internet of Things, Process Mining, Big Data Analytics, Machine Learning | Arcidiacono and Pieroni (2018), Shin et al. (2019), Fahey et al. (2020), Palací-López et al. (2020), Graafmans et al. (2021), Kregel et al. (2021), Ramires and Sampaio (2021), Singh et al. (2021) |
| Process Redesign | Process Mining, Internet of Things | Shin et al. (2019), Fernandez et al. (2021), Graafmans et al. (2021), Kregel et al. (2021), Ramires and Sampaio (2021), Singh et al. (2021) |
| Response Surface Method | Machine Learning, Artificial Neural Network | Chi et al. (2007), Su et al. (2019) |
| Regression | Artificial Neural Network, Data Mining, Machine Learning, Big Data Analytics | Johnston et al. (2009), Ghosh and Maiti (2014), Giannetti and Ransing (2016), Sanchez-Marquez and Jabaloyes Vivas (2020), Uluskan (2020), Zgodavova et al. (2020) |
| Root-cause Analysis | Data Mining, Internet of Things, Artificial Neural Network, Process Mining | Gaudard et al. (2009), Arcidiacono and Pieroni (2018), Mishra and Rane (2019), Shin et al. (2019), Graafmans et al. (2021), Kregel et al. (2021), Singh et al. (2021) |
| SIPOC | Process Mining | Graafmans et al. (2021), Kregel et al. (2021) |
| SMED | Process Mining | Singh et al. (2021) |
| Statistical Process Control | Big Data Analytics, Artificial Neural Network | Sanchez-Marquez and Jabaloyes Vivas (2020), Zgodavova et al. (2020) |
| Simulation | Data Mining, Big Data Analytics | Yang et al. (2009), Fahey et al. (2020) |
| Statistical Analysis | Data Mining, Big Data Analytics | Yang et al. (2009), Sungbum Park and Kang (2016), Zwetsloot et al. (2018), Belhadi et al. (2021) |
| Taguchi Design | Artificial Neural Network | Lin et al. (2012), Satsangi et al. (2013) |
| Taguchi Loss Function | Artificial Neural Network | Uluskan (2020) |
| Value Stream Mapping | Internet of Things | Chiarini and Kumar (2021), Tay and Loh (2021) |
| Visual Management | Internet of Things | Fernandez et al. (2021) |
| Voice of the Customer | Big Data Analytics | Laux et al. (2017), Tay and Loh (2021) |

## Stage III: Articles presenting frameworks, roadmaps, models, methods, or approaches

This stage sought to answer the last research question in this study (RQ3) by investigating existing frameworks that integrate I4.0 solutions into LSS and DMAIC. The 21 articles selected for this stage included 15 frameworks, three methods, two approaches, one model, and two roadmaps (two articles contain both a framework and a roadmap). These are outlined in Appendix C and discussed in section 4.3. Most approaches proposed combined DMAIC with the CRISP-DM model, followed by BDA techniques and hardware, such as sensors and smart devices (Figure 7). Four articles explore the integration of LSS and I4.0 technologies, three with and one without consideration of DMAIC. 3S LSS is a new Simple, Speedy and Smart paradigm propagated by Park, Dahlgaard-Park, and Kim (2020) that entails a simple roadmap for problem-solving. Instead of executing 15 steps as is typical in LSS projects, the authors suggest five 3S LSS steps. Finally, one paper (Vinodh et al. 2021) proposes a conceptual framework for integrating four continuous improvement strategies, such as LSS, Lean, KAIZEN and Sustainability, with I4.0 concepts and technologies.

Figure . Approaches proposed in framework (developed for research).

The existing approaches identified by this review were also analysed according to their practical relevance (Appendix C). Eighteen of the 21 approaches were developed and applied based on a real-world case. However, only one approach was verified in a multi-case study, i.e. the Integrated CRISP-DM and DMAIC roadmap and selection matrix developed by Zwetsloot et al. (2018), whereas the rest were evaluated within one organisation only. Four frameworks, for example, by Sordan et al. (2021), remain purely conceptual.

# Discussion

This section presents a qualitative analysis of articles selected for stages II and III. First, the studies engaging I4.0 technologies in LSS DMAIC projects are grouped and discussed according to the respective techniques or technologies. Subsequently, the existing frameworks, methods, approaches, models, and roadmaps (for simplification, the term “approach” is used) identified in stage III are analysed and grouped based on their relevance and transferability. The authors also assessed the limitations of each framework. Finally, future research opportunities are derived from gaps and limitations indicated in the studies.

## I4.0 technologies supporting LSS DMAIC projects

### Data Mining for LSS

Data Mining recognises patterns in large amounts of unstructured data (Köksal, Batmaz, and Testik 2011, Ghosh and Maiti 2014) and is one of the prominent data science techniques for enhancing LSS DMAIC. In most studies reviewed, data mining is applied for LSS tasks executed during Measure and Analyse phases of the problem-solving methodology. Data mining techniques were either employed to supplement tools or render them completely obsolete. For example, decision tree algorithms such as classification and regression trees help to identify factors with the most significant impact on a particular product characteristic or a defect (Gaudard, Ramsey, and Mia 2009, Hsiao et al. 2016, Schäfer et al. 2019) and can therefore replace traditional Six Sigma tools, such as “Cause and effect matrix” or “Ishikawa diagram”. Also, these techniques can provide inputs for planning the optimal use of resources, which positively impacts costs and service quality (Yang et al. 2009).

Furthermore, Data Mining techniques, for example, recursive partitioning (Gaudard, Ramsey, and Mia 2009) and other decision tree algorithms such as CART (Classification and Regression Trees) or CHAID (Chi-square Automatic Interaction Detection) (Ghosh and Maiti 2014), can complement the LSS tool “Design of experiments” (DOE) by highlighting the most relevant factors out of many and calculating their optimal levels. Consequently, the reduced complexity requires fewer runs, thus saving time and money (Ghosh and Maiti 2014, Gaudard, Ramsey, and Mia 2009). Several authors proposed a new framework, integrating the CRISP-DM (Cross Industry Standard for Data Mining) model with the LSS DMAIC framework. These are discussed as part of stage III. Some papers on Data Mining and Six Sigma were already published before 2011, i.e. before “Industry 4.0” was coined (Rojko 2017), which is not a surprise as both Data mining and Six Sigma were both established several decades ago.

To sum up, the main contributions of Data Mining techniques for LSS are the reduction of complexity and root-cause identification by providing insights into large amounts of historical and real-time data. In addition, a better understanding of the process through reduction of process complexity leads to fewer errors and cost savings.

### Machine Learning for LSS

Similar to Data Mining, Machine Learning is suitable for solving complex problems in LSS, whereby the boundaries between the two techniques are often blurred. Machine learning is a subset of Artificial Intelligence and uses algorithms to train systems to learn from data and become smarter without further human intervention (Zgodavova et al. 2020). Chi et al. (2007) demonstrated that machine learning models like support vector machines and other algorithms could be applied to optimise system settings and configurations and thus increase process performance. They suggest this method to make DOE obsolete because experiments are costly and time-consuming, and most companies avoid them whenever possible. However, the results might not be as precise as a physical experiment. Giannetti and Ransing (2016) applied a regression tree algorithm, a bootstrap method and a principal component analysis to overcome the complexity caused by noise factors that regression analysis could not handle. Similarly, Palací-López et al. (2020) performed a principal component analysis to resolve a complex problem in the chemical industry. The authors used a Matlab toolbox to identify the factors with the most significant impact on output performance or response. They also applied a partial least squares regression model to monitor the batch quality by predictive analysis.

### Artificial Neural Networks (ANN) for LSS

ANN is a sophisticated machine learning technique that uses multiple levels of algorithms and is very powerful in solving problems with great complexity (Uluskan 2020). The literature review revealed that articles in which LSS was enhanced by ANNs were published before 2011 as well. Case studies involving ANN to enhance DMAIC or process improvement typically target predictions for quality improvement (Lin et al. 2012, Sen 2015, Uluskan 2020) and defect prevention (Satsangi, Kumar, and Prajapati 2013, Mishra and Rane 2019, Su et al. 2019). Reduction of variation can serve as an intermediate goal of defect prevention, as evidenced by Johnston, Maguire, and McGinnity (2009) and Zgodavova et al. (2020). Abed et al. (2020) presented a DMAIC based “deep lean approach” that applies a neural network that processes real-time sensor data to predict anomalies and, thus, process deviations. Additionally, other positive effects, such as cost savings through optimal use of resources (Sen 2015) or shorter cycle time (Johnston, Maguire, and McGinnity 2009, Mishra and Rane 2019), were reported. Compared to traditional LSS tools, such as DOE, Taguchi method or RSM, ANNs can outperform these (Uluskan 2020) provided the algorithms and the construct of learning rate, momentum coefficient, and number of nodes are carefully chosen (Su et al. 2019). Uluskan (2020) demonstrated that ANN outperformed other non-linear functions, such as maximum likelihood, logistics regression or Taguchi loss function, in terms of prediction accuracy. Therefore, ANNs could be applied instead and achieve the same goals. More studies are needed to corroborate this conclusion. Other studies have shown that ANNs and Taguchi experimental design can complement each other to determine optimal process parameters by comparing the respective results (Satsangi, Kumar, and Prajapati 2013) or by applying Taguchi orthogonal arrays to determine key factors and their optimal levels and feeding them to a back-propagation network (Lin et al. 2012).

The studies above show that the application of ANN in DMAIC projects primarily addresses the Analyse and Improve phase. For example, prediction models can help identify root causes quickly and thus accelerate problem-solving (Mishra and Rane 2019). However, neural networks can also improve statistical process control by automatically detecting when a process is out of control and predicting when corrective action is needed based on the CTQ characteristics of the input factor (Zgodavova et al. 2020).

### Big Data Analytics (BDA) for LSS

Big Data refers to huge masses of primarily unstructured data collected through smart systems, i.e., sensors, devices, and applications, stored in so-called ‘data lakes’, a repository for large volumes of data (Stefanowski, Krawiec, and Wrembel 2017). Similar to Data Mining and ANNs, BDA utilises machine learning techniques to support LSS solutions. However, BDA solutions show integrations with other phases such as Measure (Koppel and Chang 2021) and Control (Sanchez-Marquez and Jabaloyes Vivas 2020, Belhadi et al. 2021, Koppel and Chang 2021) and can also support project selection (Koppel and Chang 2021). Moreover, they are used more broadly across industries, while Data Mining and ANNs are primarily adopted in the manufacturing industry. For example, (Laux et al. 2017) developed a framework combining Six Sigma and BDA principles to enhance students’ success. On the other hand, Shanshan, Wenfei, and Lijuan (2021) used text mining to analyse web data to create an intelligent curriculum system following the DMAIC approach. This review mainly examines cases where LSS benefits from BDA, while they are mutually beneficial in some cases. On the one hand, BDA can help LSS overcome the limitations of traditional statistical methods and tools, such as the lack of ability to handle big masses of data of different types and with non-linear relationships, by incorporating analytics techniques, for example, a Random Forest algorithm (Fahey, Jeffers, and Carroll 2020). On the other hand, the DMAIC method provides the structure and tools for applying BDA successfully (Laux et al. 2017, Shivajee, Singh, and Rastogi 2019). A key benefit of BDA for LSS is to analyse large amounts of data, gain insights, and ideally predict relevant process parameters.

### Internet of Things (IoT) for LSS

Internet of Things (IoT) or, in the context of I4.0, Industrial Internet of Things (IIoT), refers to interconnected physical industrial assets, such as machine sensors, technical equipment and computer hardware, but also digital models of processes, products and plants (Ghobakhloo 2018). IoT technology enables real-time data collection for improved process control and performance measurement. In the IoT architectures presented, data is automatically recorded through an interconnected network of sensors or smart devices (Arcidiacono and Pieroni 2018, Chiarini and Kumar 2021, Fernandez et al. 2021). Intelligent systems analyse the data using machine learning algorithms and provide insights and guidance for taking appropriate measures based on predicted results. The authors also proposed how IoT technologies can support LSS DMAIC phases and supported their claims through case studies. In this context, Arcidiacono and Pieroni (2018) introduced the term “Lean Six Sigma 4.0” (LSS 4.0) to represent their novel concept. Thus far, there is no officially agreed definition of LSS4.0, but academics and practitioners might argue that there is more to LSS4.0 than the integration of LSS with IoT or a few I4.0 technologies.

Quality improvement methodologies, such as Lean, Kaizen or Lean Six Sigma, can benefit from IoT technology in many ways, particularly in data collection and error prevention. The pervasive use of interconnected devices and sensors ensures seamless tracking of the entire process and therefore supports Lean tools, such as VSM or 5S (Chiarini and Kumar 2021), and also Six Sigma tools, such as root-cause analysis or SPC (Arcidiacono and Pieroni 2018).

### Process Mining for LSS

Studies reviewed have shown that Process Mining tools qualify for supporting LSS. Process Mining is an analytical approach situated between data science and process science and aims to build an exhaustive and objective vision of processes based on data from event logs delivered by IT systems (van der Aalst 2016, 18, 31-33). In contrast to the other I4.0 approaches described above, Process Mining finds a broader application outside manufacturing. The existing case studies are not primarily concerned with manufacturing processes optimisation (Shin et al. 2019, Kregel et al. 2021) but also with administrative processes, such as invoicing (Graafmans et al. 2021) or procurement (Ramires and Sampaio 2021), or with processes in healthcare, such as resource planning for eye surgeries (Singh, Verma, and Koul 2021). Besides Shin et al. (2019), all authors mapped Process Mining functions to LSS or Six Sigma, albeit with slight variations. Singh, Verma, and Koul (2021) applied a Plan, Do, Study, Act (PDSA) approach instead of DMAIC. In contrast, Kregel et al. (2021) implemented Process Mining for Six Sigma using an “Agile” DMAIC methodology, i.e., the team re-iterated between the DMAIC phases instead of performing one phase after the other like in a waterfall approach. The studies by Graafmans et al. (2021), Kregel et al. (2021), and Ramires and Sampaio (2021) show that Process Mining functions can support all DMAIC phases. The “Exploration” function provides an overview of existing process flows (Kregel et al. 2021) and thus can enhance or replace process mapping. Also, it can assist in analysing root causes for process related issues and identifying bottlenecks and idle time (Kregel et al. 2021). “Conformance checking” detects unwanted process deviations and thus provides input for process improvements or a new process model (Kregel et al. 2021). “Enhancement”, a function for autonomously improving processes, has unfortunately been little researched so far. Some Process Mining applications also offer dashboards that enable process monitoring (Graafmans et al. 2021, Ramires and Sampaio 2021). The articles by Shin et al. (2019) and by Singh, Verma, and Koul (2021), on the other hand, are more focused on finding improvement opportunities by identifying idle times, redundant steps and bottlenecks to improve the process flow and do not fully leverage the potential of a Process Mining integration with LSS like the other authors. What all the studies have in common, however, is that they used standard process mining software for the implementation (see Appendix B).

## Integration Benefits

The benefits of integrating I4.0 technologies with LSS are manifold. Figure 8 shows a word cloud of the main benefits identified from stage II papers. The findings concur with the expectations George et al. (2019, 109) raised that the usage of AI in production environments would achieve multiple competitive advantages, such as minimised costs and reduced cycle time (70, 109), by eliminating waste. The authors also argue that the application of neural networks can help reduce the complexity that would otherwise not be manageable (76-80), which Belhadi et al. (2021) demonstrated in their study. Similarly, BDA and IoT can help detect defects and analyse their root causes in real-time, improving product and service quality and reducing waste (Arcidiacono and Pieroni 2018, Belhadi et al. 2021). These conclusions echo other authors positing an integration between LSS and I4.0 (Antony, Sony, and Gutierrez 2020, Sordan et al. 2021, Vinodh et al. 2021). Appendix B outlines stage II papers, their approaches, and associated benefits.

Timeline

Description automatically generated with low confidence

Figure . Benefits word cloud (created on worditout.com).

The analysis of stage II articles also showed that technologies commonly associated with I4.0 are not equally suitable for improving LSS. For example, Sordan et al. (2021) argue that technologies such as cloud computing and cybersecurity are enabling technologies and therefore only indirectly enhance LSS. Also, there is no consensus on which technology is the most suitable. In their systematic review, Gupta, Modgil, and Gunasekaran (2020) concluded that BDA techniques were prominent for supporting LSS. However, this investigation shows that slightly more case studies employ Data Mining and ANN (Figure 5), indicating that these techniques are equally relevant.

## LSS4.0 Approaches

Stage III analysed articles presenting a framework, model, approach, roadmap or method integrating LSS with I4.0. The proposed concepts were grouped into into five categories according to their characteristics:

1. Comprehensive, high-level approaches such as presented by Gupta, Modgil, and Gunasekaran (2020), Park, Dahlgaard-Park, and Kim (2020), Tay and Loh (2021), Sordan et al. (2021), or Vinodh et al. (2021). They are conceptual only and do not provide technical details for implementation.
2. Specialised approaches as designed by Arcidiacono and Pieroni (2018), Azadeh-Fard, Megahed, and Pakdil (2019), Laux et al. (2017), or Zgodavova et al. (2020), are developed for a specific industry or to solve a specific problem and are typically not transferable to other industries.
3. Well-balanced approaches are generic enough to be relevant for different industries and detailed enough to guide organisations in their implementation. They involve either BDA techniques (Belhadi et al. 2021, Koppel and Chang 2021) or Process Mining solutions (Graafmans et al. 2021, Kregel et al. 2021).
4. Partial approaches emerging from studies by Fahey, Jeffers, and Carroll (2020), Ghosh and Maiti (2014), Hsiao et al. (2016), or Shivajee, Singh, and Rastogi (2019) utilise digital technologies in only one or two DMAIC phases.
5. Approaches with mutual support demonstrated by Chiarini and Kumar (2021), Morlock and Boßlau (2021), Schäfer et al. (2019), or Zwetsloot et al. (2018) are approaches in which the LSS methodology and I4.0 capabilities support each other.

In conclusion, high-level approaches (i) offer a holistic integration and are a good starting point for organisations looking to leverage I4.0 technologies for LSS projects. However, the well-balanced approaches (iii) provide more detailed directions for organisations to follow. Also, they are open enough to be transferable to other sectors, processes, and problems. Lastly, even though the other groups’ approaches are less comprehensive, they may still be worth adopting in a new framework. Appendix C presents an overview of the 21 approaches and discusses their limitations.

## Gaps and future research opportunities

This section considers both the research gaps identified by the authors of the articles reviewed and the limitations revealed by the analysis of this study regarding existing approaches, as listed in Appendix C. The major gaps and limitations translate into opportunities for future research, as outlined in Table 5.

Studies around integrating LSS and I4.0 have been conducted predominantly in the manufacturing sector. This imbalance could be a concern, and therefore, the current approaches should also be evaluated in other sectors (Koppel and Chang 2021, Morlock and Boßlau 2021). Similarly, frameworks designed for a specific domain, such as hospitals (Arcidiacono and Pieroni 2018, Azadeh-Fard, Megahed, and Pakdil 2019) or higher education (Laux et al. 2017) should be tested for transferability to other environments. There is plenty of room for further studies to discover the usefulness of an integration in diverse business sectors. Frameworks that remain high-level should be tested in various settings and explain how LSS tools can be enhanced or replaced. As a result, sector and company size specific frameworks and guidelines could be provided.

Table . Gaps and future research opportunities.

|  |  |  |
| --- | --- | --- |
| **Gap/Limitation** | **Author(s)** | **Future research opportunities** |
| Gap1: Lack of empirical evidence | Laux et al. (2017) Vinodh et al. (2021) Sordan et al. (2021) | Apply the proposed framework/approach in a real-world setting |
| Graafmans et al. (2021) Kregel et al. (2021) | Use Process Mining tool in real-life settings |
| Gap2: Single-case studies | Belhadi et al. (2021) Kregel et al. (2021) Zwetsloot et al. (2018) | More case studies or action research for more generalisability |
| Gap3: Sector bias | Sordan et al. (2021)  Azadeh-Fard et al. (2019) Arcidiacono and Pieroni (2018) Koppel and Chang (2021) Morlock and Boßlau (2021) | Test if the respective framework/approach can be transferred to other environments |
| Schäfer et al. (2019) | Create a guideline that is applicable for various sectors. |
| Gap4: More data-driven tools/Exploit opportunities | Ghosh and Maiti (2014) Arcidiacono and Pieroni (2018) | More data science tools to be exploited |
| Belhadi et al. (2021) Shivajee et al. (2019) Azadeh-Fard et al. (2019) Graafmans et al. (2021) | Enhance with more LSS tools |
| Kregel et al. (2021) | Combination of PM with RPA |
| Gap5: More comprehensive frameworks | Fahey et al. (2020) Koppel and Chang (2021) Hsiao et al. (2016) Morlock and Boßlau (2021) Schäfer et al. (2019) | Holistic and adaptive framework on the integrated approach (combining the tools from both LSS and I4.0) for different sectors. |
| Gap6: Lack of detail | Chiarini and Kumar (2021) Laux et al. (2017) Park et al. (2020) Tay and Loh (2021) Zwetsloot et al. (2018) | Provide details on how LSS tools and DMAIC phases can be enhanced |
| Vinodh et al. (2021) | Create action roadmaps |
| Gap7: More data, real-time | Zgodavova et al. (2020) Azadeh-Fard et al. (2019) | Test with more real-time data and/or a larger data set |
| Fahey et al. (2020) | Add more data to improve model accuracy |
| Gap8: Improve prediction capability | Hsiao et al. (2016) | Include more diverse service context features that impact customer complaints (e.g. noise level) into the model. |
| Gap9: Exploit potential for a greener process | Shivajee et al. (2019) | Define sustainability goals and aim to achieve them. |

Another vital gap that should urgently be addressed is the lack of exploitation of digital technologies. Opportunities offered by I4.0 technologies should be exploited by incorporating more diverse technologies into one framework. Unfortunately, many of the frameworks integrate only one technique, such as Data Mining (Ghosh and Maiti 2014) or Process Mining (Graafmans et al. 2021, Kregel et al. 2021) or Big Data (Laux et al. 2017). Consequently, a future study should develop a comprehensive LSS4.0 framework that contextualises results from existing studies. This new framework should subsequently be evaluated in the real world. Finally, the integration of I4.0 with the Design for Six Sigma methodology is another opportunity for future research.

# Proposed DMAIC4.0 conceptual framework

The 5 DMAIC phases served as a starting point for the framework development. The approaches presented in the previous sections were synthesised according to the underlying I4.0 concept, such as Big Data Analytics, Machine Learning, Data Mining, Process Mining and Internet of Things. The LSS tools supported by I4.0 technologies were assigned to the DMAIC phase in which they are typically used. The I4.0 tools and techniques that can support the respective LSS tool are attached beneath. The framework is illustrated in Figure 9.

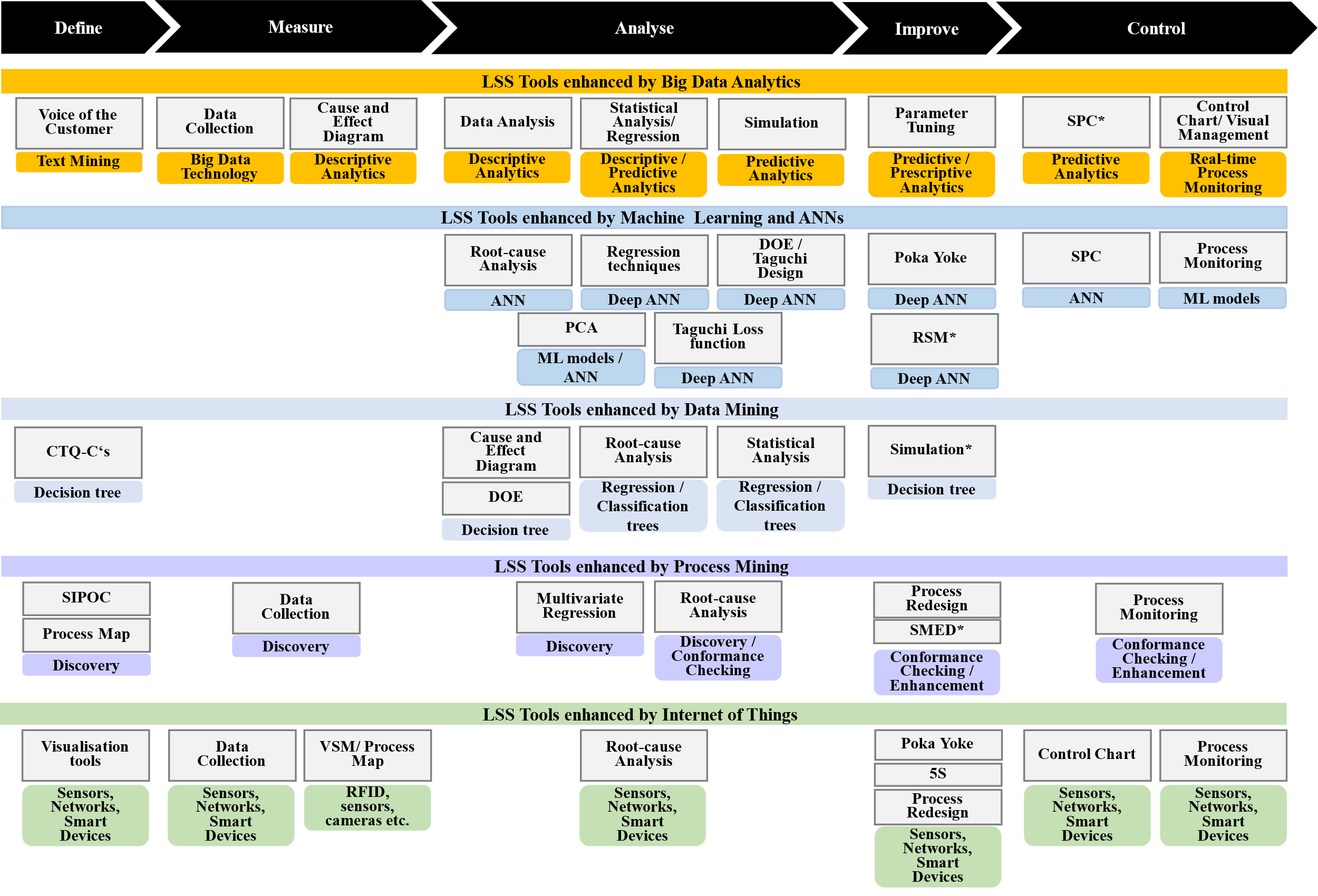


Figure . DMAIC 4.0 conceptual framework.

This conceptual framework shows five I4.0 concepts with multiple techniques that support 27 LSS tools or tasks. While BDA, PM and IoT cover all DMAIC phases, ML, ANN and DM are especially helpful for the Analyse phase. Although the articles reviewed found that DM techniques supported only six LSS tasks, this does not imply that DM is the least beneficial I4.0 concept for DMAIC projects. The advantage of DM-based tools is that they can be applied relatively quickly because they require only a few technical prerequisites, such as an analytics software tool. The implementation of Big Data and IoT architectures or PM applications, on the other hand, is more complex and time-consuming.

In general, the I4.0 tools and concepts identified for this framework can mostly be applied individually or in combination. The tool selection depends on the problem’s nature and the data availability and type. For example, where large amounts of data are available, BDA techniques can boost LSS tools by gaining insights into the process and predicting future events. However, leveraging BDA and IoT for LSS projects depends on these technologies already being implemented within the company. For both DM, ML, ANN and PM, the software market offers free standard software applications that can be implemented to enhance improvement projects. Nonetheless, all techniques have in common that they require the relevant expertise for their deployment.

# Conclusion

LSS is still the most popular business process improvement methodology (Sony, Antony, and Naik 2020) witnessed for several decades, and its integration with I4.0 is an emerging trend (Antony et al. 2019). There have been several literature reviews that examine the integration of smart technologies and LSS. However, this review differs from previous work in several aspects. First, it specifically focuses on I4.0 technologies and LSS tools without being limited to specific technologies or industries. In Stage I, we found that 692 papers have examined various I4.0 solutions such as ANN, DM, ML, BDA, and AI in conjunction with LSS tools. This shows that there is much interest in exploring integration opportunities. Second, it thoroughly explores how I4.0 technologies can enhance LSS projects. Stage II provides an in-depth assessment of 37 studies. A comprehensive picture emerged of which I4.0 technologies are suitable to enhance LSS tools and which benefits can be achieved. Further, Stage III analyses 21 papers presenting frameworks, methods, approaches, models, and roadmaps for integrating I4.0 technologies with LSS DMAIC. These were grouped and discussed according to their relevance and transferability. Their gaps and limitations were discussed as well. Researchers are encouraged to explore this hybrid area further using the future research directions outlined in section 4.4

Finally, the captured methods and approaches were synthesised into a novel DMAIC 4.0 framework. It differs from existing frameworks, for example, by Belhadi *et al.* (2020) or Sordan *et al.* (2021), because it includes several I4.0 technologies and allocates them to LSS tools along the DMAIC cycle.

## Practical and Managerial implications

The implications of this research for LSS practitioners are manifold. While several review articles on integrating LSS and I4.0 exist, no author has specifically addressed the DMAIC cycle and top 10 LSS tools thus far. The focus on LSS tools is of particular interest to practitioners facing limitations with traditional LSS tools and seeking remediation through modern technologies. The results of this study show how I4.0 technologies can be applied to boost LSS projects. Many companies have started with digital transformation; for instance, they implemented sensors, IoT and various other smart technologies generating large amounts of data (Butt 2020, Veile et al. 2020). The benefits highlighted, such as faster and more comprehensive cost reduction and the ability to optimise quality with real-time data, should encourage managers and practitioners to exploit the potential of modern technologies to benefit LSS.

The designed framework shows how 27 common LSS practices can each be supported by a specific digital technology that may already exist in the organisation. The design is intended to be generic rather than specific to enable practitioners to adapt it according to their context and needs. The framework promotes managerial awareness of areas where employees need to be upskilled. Moreover, LSS practitioners can identify which digital skills they might want to develop to keep up with the latest technological advancements.

## Agenda for future research

The scope of this study is determined by the search terms and databases used. The top 10 LSS tools were considered for this review. However, other LSS tools could also benefit from I4.0. For example, hundreds of studies exist on synergies between Design of Experiments and smart manufacturing, which deserves a separate investigation. Another limitation of this study is that conference proceedings that might provide additional relevant and more recent insights were not considered. The limitations in the scope consequently limit the comprehensiveness of the framework. Therefore, future reviews could extend the research scope and include conference papers and additional LSS tools.

Furthermore, the proposed DMAIC 4.0 framework is mainly conceptual, as most approaches analysed in this review were not tested in practice or only applied in a single-case study. To demonstrate that the framework is valid for LSS projects, it should be applied and evaluated in an industrial setting, for example, through case studies. Also, experiments or simulations may be suitable approaches to test the framework’s effectiveness, especially in manufacturing. Similarly, the framework could be applied across different sectors to assess its practical relevance. However, since the framework encompasses many different technologies, it is neither realistic nor reasonable to implement it in its entirety. Therefore, selecting a set of tools based on specific problems and circumstances is advisable. These customised tool bundles will guide organisations in digitalising LSS projects. In conclusion, we hope the framework presented here will benefit future research, projects, and LSS practitioners looking into digitalisation.

# Disclosure statement

No potential conflict of interest was reported by the author(s).

**References**

Abed, A. M., S. Elattar, T. S. Gaafar, and F. M. Alrowais. 2020. "The neural network revamping the process's reliability in deep lean via internet of things." *Processes* 8 (6). doi: 10.3390/PR8060729.

Aboelmaged, M. G. 2011. "Reconstructing Six Sigma barriers in manufacturing and service organisations." *International Journal of Quality & Reliability Management*.

Albliwi, S., J. Antony, and S. A. halim Lim. 2015. "A systematic review of Lean Six Sigma for the manufacturing industry." *Business Process Management Journal*.

Albliwi, S., J. Antony, S. A. halim Lim, and T. van der Wiele. 2014. "Critical failure factors of Lean Six Sigma: a systematic literature review." *International Journal of Quality & Reliability Management* 31 (9):1012-1030. doi: 10.1108/ijqrm-09-2013-0147.

Antony, J. 2011. "Six Sigma vs lean." *International Journal of Productivity and Performance Management*.

Antony, J., R. Snee, and R. Hoerl. 2017. "Lean Six Sigma: yesterday, today and tomorrow." *International Journal of Quality & Reliability Management*.

Antony, J., and M. Sony. 2020. "An empirical study into the limitations and emerging trends of Six Sigma in manufacturing and service organisations." *International Journal of Quality and Reliability Management* 37 (3):470-493. doi: 10.1108/IJQRM-07-2019-0230.

Antony, J., M. Sony, M. Dempsey, A. Brennan, T. Farrington, and E. A. Cudney. 2019. "An evaluation into the limitations and emerging trends of Six Sigma: an empirical study." *The TQM Journal* 31 (2):205-221. doi: 10.1108/tqm-12-2018-0191.

Antony, J., M. Sony, and L. Gutierrez. 2020. "An Empirical Study Into the Limitations and Emerging Trends of Six Sigma: Findings From a Global Survey." *IEEE Transactions on Engineering Management*:1-14. doi: 10.1109/tem.2020.2995168.

Arcidiacono, G., and A. Pieroni. 2018. "The revolution Lean Six Sigma 4.0." *International Journal on Advanced Science, Engineering and Information Technology* 8 (1):141-149. doi: 10.18517/ijaseit.8.1.4593.

Azadeh-Fard, N., F. M. Megahed, and F. Pakdil. 2019. "Variations of length of stay: A case study using control charts in the CRISP-DM framework." *International Journal of Six Sigma and Competitive Advantage* 11 (2-3):204-225. doi: 10.1504/IJSSCA.2019.101418.

Belhadi, A., S. S. Kamble, A. Gunasekaran, K. Zkik, D. K. M., and F. E. Touriki. 2021. "A Big Data Analytics-driven Lean Six Sigma framework for enhanced green performance: a case study of chemical company." *Production Planning & Control*:1-24. doi: 10.1080/09537287.2021.1964868.

Briner, R. B., and D. Denyer. 2012. "Systematic review and evidence synthesis as a practice and scholarship tool." *Handbook of evidence-based management: Companies, classrooms and research*:112-129.

Butt, J. 2020. "A strategic roadmap for the manufacturing industry to implement industry 4.0." *Designs* 4 (2):11.

Chi, H. M., O. K. Ersoy, H. Moskowitz, and K. Altinkemer. 2007. "Toward automated intelligent manufacturing systems (AIMS)." *INFORMS Journal on Computing* 19 (2):302-312. doi: 10.1287/ijoc.1050.0171.

Chiarini, A., and M. Kumar. 2021. "Lean Six Sigma and Industry 4.0 integration for Operational Excellence: evidence from Italian manufacturing companies." *Production Planning & Control* 32 (13):1084-1101. doi: 10.1080/09537287.2020.1784485.

Deuse, Jochen, Uwe Dombrowski, Fabian Nöhring, Jürgen Mazarov, and Yannick Dix. 2020. "Systematic combination of Lean Management with digitalisation to improve production systems on the example of Jidoka 4.0." *International Journal of Engineering Business Management* 12:1847979020951351. doi: 10.1177/1847979020951351.

Dogan, O., and O. F. Gurcan. 2018. "Data perspective of Lean Six Sigma in industry 4.0 Era: a guide to improve quality." Proceedings of the international conference on industrial engineering and operations management Paris.

Fahey, Will, Paul Jeffers, and Paula Carroll. 2020. "A business analytics approach to augment six sigma problem solving: A biopharmaceutical manufacturing case study." *Computers in Industry* 116:103153.

Fernandez, J. R., Y. T. Prasetyo, S. F. Persada, and A. A. N. Perwira Redi. 2021. "Automation of predictive maintenance using internet of things (Iot) technology at university-based o&m project." *International Journal of Information and Education Technology* 11 (7):332-336. doi: 10.18178/ijiet.2021.11.7.1531.

Gaudard, M., P. Ramsey, and Stephens Mia. 2009. "Interactive Data Mining Informs Designed Experiments." *Quality and Reliability Engineering International* 25 (3):299-315. doi: 10.1002/qre.971.

George, M. L., Daniel K Blackwell, M. George Jr, and Dinesh Rajan. 2019. *Lean Six Sigma in the age of artificial intelligence: Harnessing the power of the fourth industrial revolution*: McGraw-Hill Education.

Ghobakhloo, M. 2018. "The future of manufacturing industry: a strategic roadmap toward Industry 4.0." *Journal of Manufacturing Technology Management* 29 (6):910-936. doi: 10.1108/JMTM-02-2018-0057.

Ghosh, S., and J. Maiti. 2014. "Data mining driven DMAIC framework for improving foundry quality-a case study." *Production Planning and Control* 25 (6):478-493. doi: 10.1080/09537287.2012.709642.

Giannetti, C., and R. S. Ransing. 2016. "Risk based uncertainty quantification to improve robustness of manufacturing operations." *Computers and Industrial Engineering* 101:70-80. doi: 10.1016/j.cie.2016.08.002.

Goh, TN. 2011. "Six Sigma in industry: some observations after twenty‐five years." *Quality and Reliability Engineering International* 27 (2):221-227.

Goienetxea Uriarte, A., A. H. C. Ng, and M. Urenda Moris. 2020. "Bringing together Lean and simulation: a comprehensive review." *International Journal of Production Research* 58 (1):87-117.

Graafmans, T., O. Turetken, H. Poppelaars, and D. Fahland. 2021. "Process Mining for Six Sigma: A Guideline and Tool Support." *Business and Information Systems Engineering* 63 (3):277-300. doi: 10.1007/s12599-020-00649-w.

Gupta, S., S. Modgil, and A. Gunasekaran. 2020. "Big data in lean six sigma: a review and further research directions." *International Journal of Production Research* 58 (3):947-969. doi: 10.1080/00207543.2019.1598599.

Hoerl, R. 2001. "Six Sigma black belts: what do they need to know?" *Journal of Quality Technology* 33 (4):391-406.

Hsiao, Y. H., L. F. Chen, Y. L. Choy, and C. T. Su. 2016. "A novel framework for customer complaint management." *Service Industries Journal* 36 (13-14):675-698. doi: 10.1080/02642069.2016.1272592.

Johnston, AB, LP Maguire, and TM McGinnity. 2009. "Downstream performance prediction for a manufacturing system using neural networks and six-sigma improvement techniques." *Robotics and computer-integrated manufacturing* 25 (3):513-521.

Köksal, G., I. Batmaz, and M. C. Testik. 2011. "A review of data mining applications for quality improvement in manufacturing industry." *Expert Systems with Applications* 38 (10):13448-13467. doi: 10.1016/j.eswa.2011.04.063.

Koppel, S., and S. Chang. 2021. "MDAIC – a Six Sigma implementation strategy in big data environments." *International Journal of Lean Six Sigma* 12 (2):432-449. doi: 10.1108/IJLSS-12-2019-0123.

Kregel, I., D. Stemann, J. Koch, and A. Coners. 2021. "Process Mining for Six Sigma: Utilising Digital Traces." *Computers and Industrial Engineering* 153. doi: 10.1016/j.cie.2020.107083.

Laux, C., N. Li, C. Seliger, and J. Springer. 2017. "Impacting Big Data analytics in higher education through Six Sigma techniques." *International Journal of Productivity and Performance Management* 66 (5):662-679. doi: 10.1108/IJPPM-09-2016-0194.

Lin, W. T., S. T. Wang, M. H. Li, C. T. Huang, and H. Y. Huang. 2012. "An analysis of an optical coating process capability prediction model in a Six Sigma procedure by integrating APIBPN and K-means." *International Journal of Materials and Product Technology* 44 (1-2):17-34. doi: 10.1504/IJMPT.2012.048193.

Mishra, N., and S. B. Rane. 2019. "Prediction and improvement of iron casting quality through analytics and Six Sigma approach." *International Journal of Lean Six Sigma* 10 (1):189-210. doi: 10.1108/IJLSS-11-2017-0122.

Morlock, F., and M. Boßlau. 2021. "Expert-supported Data Science Projects for Production Integration of Expert Knowledge through Process Optimization Methods." *ZWF Zeitschrift fuer Wirtschaftlichen Fabrikbetrieb* 116 (6):438-441. doi: 10.1515/zwf-2021-0097.

Okoli, Chitu, and Kira Schabram. 2010. "A guide to conducting a systematic literature review of information systems research."

Palací-López, D., J. Borràs-Ferrís, L. T. da Silva de Oliveria, and A. Ferrer. 2020. "Multivariate six sigma: A case study in industry 4.0." *Processes* 8 (9). doi: 10.3390/PR8091119.

Park, Sung Hyun, Su Mi Dahlgaard-Park, and Dong-Chun Kim. 2020. "New Paradigm of Lean Six Sigma in the 4th Industrial Revolution Era." *Quality Innovation Prosperity* 24(1) (1):1-16. doi: 10.12776/qip.v24i1.1430.

Pongboonchai-Empl, T., D. Stemann, and J. Antony. 2021. "LSS4.0-Creating Opportunities from existing Limitations."

Puram, P., and A. Gurumurthy. 2021. "Celebrating a decade of International Journal of Lean Six Sigma–a bibliometric analysis to uncover the “as is” and “to be” states." *International Journal of Lean Six Sigma*.

Ramires, Francisco, and Paulo Sampaio. 2021. "Process mining and lean six sigma: a novel approach to analyse the supply chain quality of a hospital." *International Journal of Lean Six Sigma* ahead-of-print (ahead-of-print). doi: 10.1108/IJLSS-12-2020-0226.

Rojko, Andreja. 2017. "Industry 4.0 concept: Background and overview." *International Journal of Interactive Mobile Technologies* 11 (5).

Rousseau, Denise M, Joshua Manning, and David Denyer. 2008. "11 Evidence in management and organisational science: assembling the field’s full weight of scientific knowledge through syntheses." *Academy of Management Annals* 2 (1):475-515.

Sanchez-Marquez, R., and J. Jabaloyes Vivas. 2020. "Multivariate SPC methods for controlling manufacturing processes using predictive models – A case study in the automotive sector." *Computers in Industry* 123. doi: 10.1016/j.compind.2020.103307.

Satsangi, P. S., S. Kumar, and D. R. Prajapati. 2013. "A six sigma methodology for reducing casting defects in foundry operations." *International Journal of Technology, Knowledge and Society* 8 (4):163-176. doi: 10.18848/1832-3669/CGP/v08i04/56312.

Schäfer, F., A. Mayr, A. Hess, C. Zeiselmair, H. Otten, and J. Franke. 2019. "Six sigma 4.0 - Data mining as supporting technique within the zero defect management." *ZWF Zeitschrift fuer Wirtschaftlichen Fabrikbetrieb* 114 (3):140-144. doi: 10.3139/104.112047.

Sen, P. 2015. "Application of ANN in Six Sigma for CO modeling and energy efficiency of blast furnace: A case study of an Indian pig iron manufacturing organisation." *International Journal of Six Sigma and Competitive Advantage* 9 (2-4):109-125. doi: 10.1504/IJSSCA.2015.074957.

Shanshan, S., L. Wenfei, and L. Lijuan. 2021. "Applying lean six sigma incorporated with big data analysis to curriculum system improvement in higher education institutions." *International Journal of System Assurance Engineering and Management*:1-16.

Shin, Sung-chul, Seon Yeob Kim, Chun-Myoung Noh, Soon-sup Lee, and Jae-chul Lee. 2019. "Manufacturing process improvement of offshore plant: Process mining technique and case study." *Ocean systems engineering* 9 (3):329-347.

Shivajee, V., R. K. Singh, and S. Rastogi. 2019. "Manufacturing conversion cost reduction using quality control tools and digitisation of real-time data." *Journal of Cleaner Production* 237:117678.

Singh, M., and R. Rathi. 2019. "A structured review of Lean Six Sigma in various industrial sectors." *International Journal of Lean Six Sigma*.

Singh, S., R. Verma, and S. Koul. 2021. "A collaborative method for simultaneous operations: case of an eye clinic." *Opsearch*. doi: 10.1007/s12597-021-00513-9.

Snee, R. D. 2010. "Lean Six Sigma – getting better all the time." *International Journal of Lean Six Sigma* 1 (1):9-29. doi: 10.1108/20401461011033130.

Sony, Michael, Jiju Antony, and Subhash Naik. 2020. "How do organisations implement an effective LSS initiative? A qualitative study." *Benchmarking: An International Journal* 27 (5):1657-1681. doi: 10.1108/bij-10-2019-0451.

Sony, Michael, Subhash Naik, and K. K. Therisa. 2019. "Why do organisations discontinue Lean Six Sigma initiatives?" *International Journal of Quality & Reliability Management* 36 (3):420-436. doi: 10.1108/ijqrm-03-2018-0066.

Sordan, J. E., P. C. Oprime, M. L. Pimenta, S. L. da Silva, and M. O. A. González. 2021. "Contact points between Lean Six Sigma and Industry 4.0: a systematic review and conceptual framework." *International Journal of Quality & Reliability Management* ahead-of-print (ahead-of-print). doi: 10.1108/IJQRM-12-2020-0396.

Stefanowski, J., K. Krawiec, and R. Wrembel. 2017. "Exploring complex and big data." *International Journal of Applied Mathematics and Computer Science* 27 (4):669-679.

Su, C. T., F. M. Su, T. Y. Chen, and L. F. Chen. 2019. "Enhancing the Structural Strength of an ODD Laptop via Six Sigma Approach." *IEEE Transactions on Components, Packaging and Manufacturing Technology* 9 (11):2200-2209. doi: 10.1109/TCPMT.2019.2922728.

Tay, Huay Ling, and Hui Shan Loh. 2021. "Digital transformations and supply chain management: a Lean Six Sigma perspective." *Journal of Asia Business Studies*.

Tranfield, David, David Denyer, and Palminder Smart. 2003. "Towards a methodology for developing evidence‐informed management knowledge by means of systematic review." *British journal of management* 14 (3):207-222.

Uluskan, Meryem. 2019. "Analysis of Lean Six Sigma tools from a multidimensional perspective." *Total Quality Management & Business Excellence* 30 (9-10):1167-1188. doi: 10.1080/14783363.2017.1360134.

Uluskan, Meryem. 2020. "Artificial Neural Networks as a quality loss function for Six Sigma." *Total Quality Management & Business Excellence* 31 (15-16):1811-1828. doi: 10.1080/14783363.2018.1520597.

van der Aalst, W. 2016. *Process Mining: Data Science in Action*. 2nd ed. Berlin: Springer.

van der Aalst, W., B. van Dongen, J. Herbst, L. Maruster, G. Schimm, and A. Weijters. 2003. "Workflow mining: A survey of issues and approaches." *Data & knowledge engineering* 47 (2):237-267.

van Eck, Nees Jan, and Ludo Waltman. 2010. "Software survey: VOSviewer, a computer program for bibliometric mapping." *Scientometrics* 84 (2):523-538. doi: 10.1007/s11192-009-0146-3.

Veile, Johannes W., Daniel Kiel, Julian Marius Müller, and Kai-Ingo Voigt. 2020. "Lessons learned from Industry 4.0 implementation in the German manufacturing industry." *Journal of manufacturing technology management.* 31 (5):977-997. doi: 10.1108/JMTM-08-2018-0270.

Vinodh, S., J. Antony, R. Agrawal, and J. A. Douglas. 2021. "Integration of continuous improvement strategies with Industry 4.0: a systematic review and agenda for further research." *TQM Journal* 33 (2):441-472. doi: 10.1108/TQM-07-2020-0157.

Wang, Hao-nan, Qi-qi He, Zheng Zhang, Tao Peng, and Ren-zhong Tang. 2021. "Framework of automated value stream mapping for lean production under the Industry 4.0 paradigm." *Journal of Zhejiang University-SCIENCE A* 22 (5):382-395.

Widjajanto, S., H. H. Purba, and S. C. Jaqin. 2020. "Novel POKA-YOKE approaching toward industry-4.0: A literature review." *Operational Research in Engineering Sciences: Theory and Applications* 3 (3):65-83.

Williams, Ralph I., Leigh Anne Clark, W. Randy Clark, and Deana M. Raffo. 2020. "Re-examining systematic literature review in management research: Additional benefits and execution protocols." *European Management Journal*. doi: 10.1016/j.emj.2020.09.007.

Yadav, G., D. Seth, and T. N. Desai. 2017. "Analysis of research trends and constructs in context to lean six sigma frameworks." *Journal of Manufacturing Technology Management* 28 (6):794-821. doi: 10.1108/jmtm-03-2017-0043.

Yang, C. C., W. T. Lin, H. M. Chen, and Y. H. Shi. 2009. "Improving scheduling of emergency physicians using data mining analysis." *Expert Systems with Applications* 36 (2 PART 2):3378-3387. doi: 10.1016/j.eswa.2008.02.069.

Zgodavova, K., P. Bober, V. Majstorovic, K. Monkova, G. Santos, and D. Juhaszova. 2020. "Innovative methods for small mixed batches production system improvement: The case of a bakery machine manufacturer." *Sustainability (Switzerland)* 12 (15). doi: 10.3390/su12156266.

Zwetsloot, I. M., A. Kuiper, T. S. Akkerhuis, and H. de Koning. 2018. "Lean Six Sigma meets data science: Integrating two approaches based on three case studies." *Quality Engineering* 30 (3):419-431. doi: 10.1080/08982112.2018.1434892.

# Appendices

Appendix A. Analysis of Review Papers

Appendix B. Stage II analysis results

Appendix C. Existing Frameworks

# List of Tables

Table . List of search terms for the systematic literature review (developed for research).

Table . SLR Inclusion and Exclusion Criteria (developed for research).

Table . Stage II articles by sector (developed for research).

Table . LSS tools and tasks supported by I4.0 technologies (developed for research).

Table . Gaps and future research opportunities (developed for research).

# List of Figures

Figure 1. SLR process (inspired by Tranfield et al. (2003) and Goienetxea Uriarte et al. (2020)).

Figure 2. Co-occurrence network of frequent keywords in papers combining LSS tools and I4.0 technologies (developed for research).

Figure 3. Articles integrating I4.0 Technologies with LSS Tools (developed for research).

Figure 4. Publications per year by I4.0 technology (developed for research).

Figure 5. I4.0 Technologies enhancing LSS (developed for research).

Figure 6. DMAIC phases enhanced by I4.0 technologies (developed for research).

Figure 7. Approaches proposed in framework (developed for research).

Figure 8. Benefits word cloud (created by the author on worditout.com).

Figure 9. DMAIC 4.0 conceptual framework.