ISSN 2732-3781 https://doi.org/10.52398/gjsd.2025.v5.i1.pp4-25

GILE Journal of Skills Development

Skills Prediction for the South African Food and Beverages Manufacturing Sector: Towards Addressing the High Youth Unemployment

Megashnee Munsamy

Johannesburg Business School, University of Johannesburg, South Africa
ORCID: https://orcid.org/0000-0001-7403-2580

Arnesh Telukdarie

Johannesburg Business School, University of Johannesburg, South Africa
ORCID: https://orcid.org/0000-0003-0057-3493

Abstract

The high unemployment rate, especially among the youth in South Africa, necessitates addressing the skills requirements to meet the labour market demands. In South Africa, the Sector Education and Training Authorities are mandated to identify and provide the skills demanded. A quantitative approach of big data analytics is proposed for predicting the future skills of the South African food and beverages manufacturing sector. The source of big data is skills-related publications from the Elsevier Scopus database, for the period 1999 to 2020. Natural language processing, and word representation technique was applied to quantify and analyse year-on-year change of author and index keywords, subject areas and number of publications towards the determination of the future skills requirement. The analysis identified four vital skills: (1) entrepreneurship inclusive of digital skills; (2) data analytics inclusive of AI, simulation, and modelling; (3) digital tools and technologies implementation, and operation; and (4) sustainable operations skills of advanced process control, and energy, water, and resource optimisation. The identified skills were comparatively analysed with the World Economic Forum predictions for skills requirements for 2025, with the results aligning to the findings and validating the approach. The approach is applicable to any Sector Education Training Authority in South Africa. This enables enhanced skills planning to better prepare the South African workforce, especially youth, given the persistent challenges to completion of Grade 12 leading to access to higher education. It is acknowledged that youth with tertiary qualifications have a better chance of employment than those without, thus skills planning initiatives is essential. The skills gap cannot be instantaneously addressed but requires a long-term strategic effort from all stakeholders in the skills ecosystem. As a starting point in youth skills development, short courses to workshops to workplace training is a feasible option.

Keywords/key phrases**:** big data analytics, skills prediction, food and beverage, fourth industrial revolution, digital skills

1. Introduction

The ability to accurately predict future skills is a key strategy of governments the world over with the objective of sustaining economies and seeking key differentiators for growth. The highly competitive global skills space is influenced by various domestic and international factors. With the onset of the fourth industrial revolution, the pace of change of skills is accelerating, making predicting skills requirements an interesting opportunity. Country specific protocols of skills prediction, together with regional and global organizations protocols exist and are diverse. The skills predicted by these countries and regional groups have matured over the years, with the prediction cycle requiring many years for verification with multiple variables considered.

Skills forecasting is a systematic method to derive quantified projections of future skills needs by using labour market and skills information (British Council, 2018). Skills forecasting enables informed decision-making by policymakers and improves labour market operations (ILO-OECD, 2018). It enables the education and training systems to strategically respond and prevent skills shortages and mismatches (ILO, 2020). The availability of the required skills by the labour market facilitates economic growth (ILO, 2020).

The key drivers of the rapidly evolving skills requirements of the labour market are the fourth industrial revolution (4IR), globalisation, climate change, and demographic changes (ILO-OECD, 2018). The 4IR and its disruptive nature make future skills forecasting even more critical, but also more complex and challenging (ILO-OECD, 2018). Within the global manufacturing industry, 13.2% of workers are at risk of displacement due to digitization, with only 44.6% of the displaced workers expected to be successfully redeployed (World Economic Forum, 2020). The skills required influence occupational standards, apprenticeship as well as learnership designs, training programmes, national and sectoral employment and skills strategies, curriculum development, and career guidance (ILO-OECD, 2018).

South Africa has one of the highest unemployment rates in the world of 34.5%, which requires systematic and aggressive measures to address (Statistics South Africa (a), 2022). In addressing the high unemployment rate, the provision of future skills is critical. However, South Africa lacks a skills forecasting model (DHET, 2019), which is detrimental to its economic growth and addressing of the present inequalities. Thus, skills forecasting is a priority for South Africa. The South African food and beverages manufacturing sector, a key contributor to Gross Domestic Product GDP and employer of the labour force, lacks a tool for skills forecasting. The sector is in transition driven by the 4IR, the COVID-19 pandemic, climate change and evolving customer requirements. To remain competitive in the global marketplace, the sector needs to transform faster. In some countries the sector has advanced via adoption of digital transformation in addressing climate change and evolving customer requirements. Understanding and learning from these advancements is an opportunity to propel the digital advancement of the South African food beverages manufacturing sector. But this advancement is reliant on the availability of required skills.

The existing skills model are typically country specific, focusing on either skill supply or demand, data intensive, requiring extended periods of quality data, and requiring a high level of technical proficiency, which is limiting to a developing country such as South Africa. In developing a South African skills forecasting model key considerations include practicality and accessibility of the results, and the costs of the exercise (ILO-OECD, 2018). Big data analytics (BDA) was identified as an accepted to forecasting, that is especially applicable to a developing country. This study adopts a BDA approach to predict the skills demand of the food and beverage manufacturing sector.

2. Literature Review

Every country has a specific set of internal and external drivers impacting skills demand; hence skills prediction models are typically country specific. Some countries conduct occupational projections instead of skills forecasting, but in identifying future occupations the skills requirements are inherently defined. Globally, skills forecasting approaches are qualitative, quantitative, or integrated. Qualitative approaches range from surveys to sector studies to expert discussion forums, whilst quantitative approaches are modelling, or simulation based. An integrated approach merges the quantitative and qualitative methods; the use of qualitative results in the quantitative models addresses issues such as missing data or breaks in time series, incorporates the tacit knowledge of experts, and the use of past trends and current drivers (ILO-OECD, 2018). Skills and occupations forecasting can be at a macroeconomic, regional, or sectoral level.

2.1. South African Status

South Africa is experiencing persistently high unemployment rates as demonstrated for the period 2018-2023, where unemployed people increased from 6.1 million in 2018 to 7.9 million in 2023 (Statistics South Africa (b), 2024). Of the 7.9 million unemployed, 3.4 million had matric while 4 million did not have matric. However, only 9.4% of the total unemployed had tertiary education, (Statistics South Africa (b), 2024).

The youth in South Africa are disproportionally affected by the high unemployment rate. For the period 2018 – 2023, youth unemployment was almost twice that of adults (Statistics South Africa (b), 2024). In Quarter 3 of 2024, it was observed that 10.3 million people aged between 15 and 24 were not in employment, education, or training (Statistics South Africa (c), 2024). The youth are primarily employed in low and semi-skilled occupations due to lack of education and skills (Statistics South Africa (b), 2024). For 2023, 44% of unemployed youth did not complete matric, 45.3% completed matric and 10.1% had a tertiary qualification.

Based on the data from above, education and training are critical factors in unemployment, especially among youth, who do not have on-the-job training of adults. Given the importance of education and training on unemployment, it becomes essential that Higher Education Institutions (HEIs) and Sector Education Training Authorities (SETAs) provide the necessary skills set to youth to meet the demands of the labour market. Similarly, policies promulgated by the government should focus on the support and development of youth, facilitating the completion of matric at a minimum and access to tertiary and vocational training and other learning initiatives. Thus, skills forecasting is essential in identifying initiatives to support the skilling of youth to improve employability.

2.2. South African skills forecasting approach

In South Africa, the provision of skills is the mandate of SETAs. There are 21 SETAs in South Africa, and each is specific to a particular economic sector. The mandate of the SETAs is to analyse sector priorities via the collection of data from companies in the sector, interpretation of data, and support of emergent skills profile through learnerships and other mechanisms. Each SETA independently defines its skills requirement for its sector based on the analyses of workplace skills plans submitted by companies comprising the sector and interviews with said companies. The workplace skills plan is submitted annually by companies and defines the current occupations and skills requirements of the said company. Thus, in the current skills cycle, the economic sectors, chambers, and companies influence the skills development process. The primary challenge is the limited future skills forecasting, including digital and technological skills. In South Africa, the current skills do not meet industry demands, resulting in a skills mismatch and intensifying the current high unemployment rate. Based on the findings of the 2020 World Economic Forum (WEF) report, the anticipated future skills in South Africa are analytical and critical thinking, problem-solving, innovation, resilience, quality control, technology use, design and programming, and emotional intelligence (World Economic Forum, 2020).

2.3. Global review of skills forecasting

A review of the approaches adopted by various countries was conducted. The countries were selected based on their established system of skills and/or occupation predictions. Most countries adopt a semi-integrated approach, with qualitative and quantitative approaches conducted independently with none to limited integration. The quantitative approaches adopted typically comprises of:

1. Surveys/questionnaires for employers, workers, and graduates on qualifications, skills, occupations, and occupational growth
2. Census or population data
3. Statistical documents on the labour force (current and future, if available)
4. Business surveys

The quantitative approach is fundamentally based on an economic approach, which varies from regression to econometric models. The models commonly utilise historical data in terms of occupation trends, population data and economic data. The models typically determine the demand of skills/occupation, with some models determining the supply and demand side and other model determining an equilibrium state of supply and demand.

In a review of countries with established skills forecasting approach, as presented in Table 1, it is notable that all countries except Brazil (a developing country) are categorised as first-world countries. The approaches of both developed and developing countries are analysed to identify an approach appropriate to South Africa, which is at the infancy stage of skills forecasting. In review of the skills and/or occupations forecasting approach of India and China, there is limited to no information from either public or private institutions on skills or occupation prediction. Thus, an approach to skills or occupation prediction could not be determined for either country

(British Council, 2018). However, it should be noted that both countries have several initiatives on skills development including upskilling.

TABLE 1. CHARACTERISTICS OF THE VARIOUS COUNTRY SPECIFIC SKILLS/OCCUPATIONS MODEL

Country	Approach	Skill/ occupations	Time span of prediction	Frequency of update	Models feature	Survey Types
Canada	Quantitative (Governmemnt of Canada, 2023), (Employment and Social Development Canada, 2022) and Qualitative (OECD, 2016)	Forecasts 292 occupational groupings and skills	10 years national level	2 years	• Macroeconomic model • Considers 42 industries • Supply side estimation • Data intensive	Employers, and workers, or graduate surveys
USA	Quantitative (U.S. Bureau of Labor Statistics, 2021) and Qualitative (OECD, 2016)	Occupations projections	10 years national level	2 years	• Macroeconomic model • Inclusion of 205 commodities • Follows a six-step process, which examines: o Size and demographic of labour force o Aggregate economic growth o Commodity final demand o Input-output o Industry output & employment o Occupational employment and openings • The model is designed to achieve full employment in the target years • Data intensive	Workers or graduate surveys
UK	Quantitative (Wilson, Barnes, May-Gillings, Patel, & Bui, 2020) and Qualitative (OECD, 2016)	Projects occupational employment for each industry	• 10-year national and regional level • National level - 87 industries • Regional level - 46 industries		• Multisectoral dynamic model (MDM-E3) • A macroeconomic model, which links with a qualification supply model and main occupation model. These models link to other models. • Requires long times series data set • Complex	Employer and national surveys on occupational and qualification structures; labour force survey; annual employment survey; annual business survey, and population and census survey

Australia	Quantitative (National Skills Commission, 2022)	Occupations and skills prediction	5-year employment forecasts by industry, occupation, skill level and region	Yearly	• Derived from time series models. • Combines forecasts autoregressive integrated moving average (ARIMA) and exponential smoothing with damped trend (ESWDT) models	Australian Bureau of Statistics labour force survey
	Quantitative (CoPs, 2022)	341 occupations predictions		Yearly	• General Equilibrium model of the Australian economy • Required up-to-date data • Utilises information from specialist forecasting organisations and historical trends	
	Qualitative (OECD, 2016)					Employers, and workers, or graduate survey
Germany	BIBB-IAB qualification and occupational field projections (Quantitative) (Maier, 2017) and Qualitative (OECD, 2016)	Occupations prediction	15–20 years	2 years	• QINFORGE econometric model • Determines both supply and demand at the occupation level • Data intensive - micro census data is used as well as data from education institutions	Micro census data and employer and workers or graduate surveys
Brazil	Senai Model (Quantitative and Qualitative) (ETF, 2016)	Skills prediction at the national level	5–10 years		• Identifies emerging technologies that will influence the occupations demand and the associated skills • Analysis of occupational trends via two models; an input-output model and analysis of the database of the Ministry of Labour and Employment • The results are compared to the sector occupational demand/changes of USA, Canada, Australia, New Zealand, and the UK	

Source: own compilation

All the countries reviewed adopt both qualitative and quantitative approaches for skills and occupation predictions, but not necessarily an integrated approach. Common characteristics among the models include:

a. Updating the model bi-annually, with Australia updating its model annually. The frequency of the updates enables the system to moderate the impacts of shocks to the system such as COVID-19, whilst allowing the longer-term drivers of skills and occupation changes to be understood and defined such as 4IR, globalisation, and climate change.

b. A macroeconomic structure

c. Requiring time series data for aggregated and disaggregated model input variables

The challenge of these forecasting models is the data requirements, extended periods of data, and the quality of the data. This is especially challenging for developing economies, where data availability and quality are limited. The Brazilian approach is distinctive, with a focus on emerging technologies. The key emerging technologies are identified together with their impact on occupations (current and new), thereby allowing the identification of the required skills. To verify the results, comparisons to the sector and occupational change predictions of the USA, Canada, Australia, New Zealand, and the United Kingdom were conducted. This is an interesting approach that could have applicability in developing countries, as key drivers of skills are globalisation, and technological advancements. This approach also negates the intensive data requirements of macroeconomic skills forecasting models.

2.4. Big data analytics

Informed decision making is essential for organisational success, which impacts a country's growth and competitiveness (Elgendy, Elragal, & Päivärinta, 2022). The fourth industrial revolution presents tools for facilitating skills forecasting, such as big data analytics (BDA). BDA is applied to define the unknown in data (Elgral & Klischewski, 2017) and to facilitate decision-making. BDA has been applied in various sectors in support of decision-making; retail, hospitality, entertainment, government, health, finance, manufacturing, and transport (Sabharwal & Miah, 2021). There are three BDA classes; descriptive, predictive, and prescriptive. Descriptive describes what has happened or what is happening, predictive defines what is to happen, and prescriptive defines what should be done, to achieve what should be happening (Sabharwal & Miah, 2021). Big data analytics (BDA) together with artificial intelligence (AI) enables the analysis of large volumes of data in short time periods of time to extract relevant labour market information (ILO, 2020). The BDA process can be broken down into the following phases: (1) big data acquisition, (2) data pre-processing, (3) data analysis, and (4) information interpretation (Elgral & Klischewski, 2017). Big data acquisition is the gathering of data. The data can be obtained from various sources such as social media platforms, IoT devices, online job vacancy sites, and databases. Information extracted from text is referred to as text mining, while information obtained from unstructured audio data is audio analytics. Similarly, information extracted from video streams and social media channels are video analytics and social media analytics respectively (Gandomi & Haider, 2015). Once the data is gathered it must be cleaned in the pre-processing stage. This encompasses all steps for ensuring consistency of data, such as removal of duplicates, and missing values, and consolidation and integration of data (Gandomi & Haider, 2015). In the data analysis step, various techniques are applied to analyse the data, ranging from simple regression algorithms to various AI techniques, such as random forests to artificial neural networks. In the final step of interpretation, the information extracted from the analysis is interpreted and visually presented to aid in decision-making.

BDA has been applied in skills and occupation forecasting, typically adopting a text mining approach of usage of data gathered from online job vacancy sites to predict skills demand. The

European Centre of the Development of Vocational Training (CEDEFOP) investigated the use of online job vacancies (OJVs) to gather intelligence on skills (ILO, 2020). It has established the skills online vacancy analysis tool for Europe (OVATE), which presents the information extracted from OJVs. The information, presented in a dashboard format, includes online job advertisements by occupation and most requested skills. It is advised that OJV analysis should not replace conventional methods but support it. The advantages of OJVs include a large volume of data, time, and cost-effectiveness as well as in-depth information on skills, whilst the disadvantages include unstructured information, non-representativeness, and advanced analytical skills (ILO, 2020). Mezzanzanica and Mercorio (2019) identified big data for predicting future skills, new and emerging occupations, and the rate of digital skills adoption by occupation. They further state that big data is critical for future skills prediction as the vacancies advertised on the internet is the only alternative source of information provided by skills survey and big data provides detailed and granular data to illustrate the variation in skills demand across occupations, sectors, and regions. Big data from an online job vacancy portal, together with AI techniques, were used to analyse the labour market and demand for jobs and employment in Malawi. The results provided various insights, including trending jobs in the labour market (Mezzanzanica & Mercorio, 2019). LinkedIn, a professional network that is employment orientated, utilizes its own data in determining how occupations are changing, the skills required by workers and identifies the key skills on a yearly basis. The similarity of the scores between the two years indicates the overlap of common skills and the relative importance of the skills. Six years of data were aggregated and analysed (LinkedIn, Future of skills, 2021). The results are presented at three resolutions of country, sector, and occupation on their website Economic Graph (LinkedIn, Future of skills, 2021).

Currently, a key characteristic of big data in skills prediction is the use of online job vacancies as the primary source of data. This can be a limitation, as there is a potential of bias in coverage of occupations, firm sizes, and regions (CEDEFOP, 2019). In developing countries, especially in Africa, the online job posting may be skewed to the requirements of businesses based in urban areas, whilst the skill demand of the rural areas is lost. Thus, necessitating the use of an unbiased source for data acquisition.

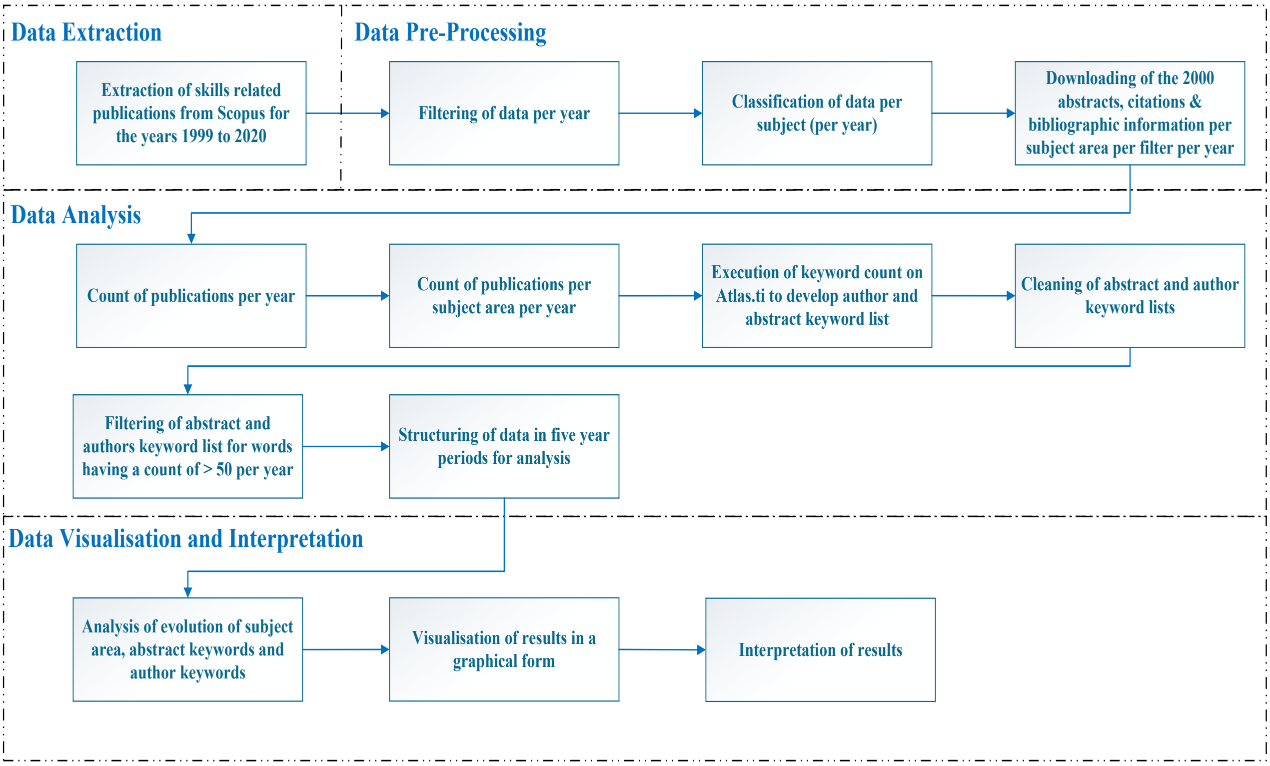
3. Methods

A text mining BDA approach was adopted as it enabled forecasting of future skills without the intensive data, time, and cost requirements of traditional quantitative methods. (Mezzanzanica & Mercorio, 2019), (ILO, 2020) and (LinkedIn, Future of skills, 2021) utilised online vacancy sites in their text mining BDA approach. Due to the potential biases of the South African online job sites that focuses on economic hubs and following the principle of learning from what has been achieved globally, this study adopted the Elsevier Scopus database as the source of big data.

The Elsevier Scopus database defines the skills evolution in the food and beverages manufacturing sector, the current status, and the future trajectory, which could influence the South African labour market. The information can be considered to be valid, since all Elsevier Scopus publications are peer reviewed. Further to this, the approach aligns to the methodology adopted by Brazil, which validates its occupational predictions against those of the USA, Canada, Australia, New Zealand, and the UK. The Elsevier Scopus database defines the skills trajectory of the countries that have digitally transformed enabling alignment of the South African food and beverages manufacturing sector.

The Elsevier Scopus database contains publications from other acknowledged publishers such as IEEE, Taylor and Francis, Wiley, and Emerald Insight. Thus, Elsevier Scopus is a comprehensive source for publication on skills. The Elsevier Scopus database contains peer-reviewed abstracts on various knowledge areas spanning from engineering to health science to social science. For each abstract, the bibliographic details, citation information, publication references, and the link to the publisher are provided. The Elsevier Scopus database is filtered for publications on skills, which serves as the big data for analysis. The text mining BDA technique adopted comprises data acquisition, pre-processing, processing, and presentation, as illustrated in Figure 1.

FIGURE 1. BDA APPROACH



Source: own compilation

- ***Data extraction***

The data is extracted from Elsevier Scopus for a period of 20 years from 1999 to 2020, with the search term "skills/skill" in the title, the keywords, and abstract fields. The inclusion criteria for the search were for publications in English only.

- ***Filtering of data per year***

Utilising the "analyse results" functionality in Elsevier Scopus, the publications were filtered per year.

- ***Classification of data per subject area***

For each year the publications were further classified per subject area, as per the subject area analysis in the “analyse results” functionality. The subject areas include social sciences, psychology, medicine, engineering, computer science, business management and accounting, nursing, health professionals, energy, and chemical engineering, among others. This enabled analysis of the evolution of research in a particular area, indicating growth (increased publications) or decline (decreased publications). An increase in publications in a subject area, indicates an increased focus on said subject area, driven by advancements, and technological innovation. A decrease in publications in a specific subject area, indicates waning interest or a divergence to a specialised interest or a new subject area driven by disruptions such as the global financial crisis in 2010, COVID-19, or the fourth industrial revolution.

- ***Downloading of abstracts and keywords***

Elsevier Scopus has a limitation of a maximum download of 2,000 abstracts and associated citations and bibliographic information per filter. The filters include newest publications, oldest publications, highest cited, lowest cited, relevance, the first author (A-Z), first author (Z-A), source title (A-Z), and source title (Z-A). For each filter, the first 2,000 publications were downloaded to obtain a comprehensive publication list per subject area per year.

- ***Formatting of data to Atlas format***

The abstract, author keywords, and abstract keywords extracted from Elsevier Scopus in CSV format were converted to Microsoft Excel format, which is compatible with Atlas.ti. The data extracted for each of the filters (per subject area per year) were merged and duplicates removed.

- ***Atlas cleaning***

For each year, per subject area, the extracted abstracts were analysed to produce a word list and associated word count, with the results exported to Microsoft Excel. Similarly, a count of the occurrence of the author keywords (per year) weas also conducted.

- ***Data cleaning***

The abstract and author keyword count lists were cleaned to remove duplicates, This was followed by removal of stop words such as is, the, and that. This was proceeded by the aggregation of synonyms and plural forms of the same word.

- ***Filtering of word count results***

For each of the word count lists (abstract and author keywords), words having a count of less than 50 per year were removed.

- ***Data structuring***

Each word count list was transformed into a n*m data matrix, with n representing the year, and m representing either the abstract keyword list or the author keyword list. The n*m data matrix detailed the count of either the abstract keyword or author keyword per year.

- ***Data interpretation and visualisation***

Each matrix was individually analysed in clusters of five years to determine the trends in the abstract and author keywords. Similar to the subject area count of publications, the evolution in the abstract and authors keywords indicated growth or decline. The abstract and author keywords were skills associated and, thus, were used as an indication of the skills trends and demands.

4. Results

The food and beverages manufacturing sector is a core economic activity in South Africa. The sector is in a state of transition, driven by 4IR, the COVID-19 pandemic, globalisation, and technological advancement. Thus, the ability to accurately predict and develop the appropriate skills in the food and beverage manufacturing sector is a priority. The current approach adopted by the Food and Beverages Manufacturing (FoodBev) SETA is limited and not fully representative. Thus, the BDA approach was applied to the FoodBev sector.

TABLE 2. NUMBER OF SKILLS PUBLICATIONS PER YEAR

Year	Number of Publications	Year on Year % Difference
1999	10,810	
2000	11,535	6.71
2001	12,807	11.03
2002	13,835	8.03
2003	15,789	14.12
2004	17,394	10.17
2005	20,655	18.75
2006	22,594	9.39
2007	24,588	8.83
2008	26,905	9.42
2009	29,433	9.40
2010	33,157	12.65
2011	35,310	6.49
2012	37,638	6.59
2013	41,863	11.23
2014	43,575	4.09
2015	46,286	6.22
2016	49,531	7.01
2017	53,577	8.17
2018	58,995	10.11
2019	68,093	15.42
2020	73,341	7.71
Total	**747,711**	

Source: own compilation

The number of publications on skills was erratic from 1999 to 2013. Since 2014, there has been a steady year-on-year (YoY) increase of 4.09% in 2014 to 15.42% in 2019. A decrease in YoY growth was observed in 2020, which is attributed to the disruption brought upon by the COVID-19 pandemic and can thus be viewed as an outlier. Applying the YoY growth pattern from 2014 to 2019, as it can be considered the normal growth trajectory, the projected YoY increase for 2020 was 15.45%, which would have resulted in 78,616 publications. Thus, the COVID-19 pandemic reduced the number of publications by 5,275.

4.1. Subject area analysis

Table 3 details the year-on-year (YoY) percentage change per subject area, for the subject areas most relevant to the food and beverages manufacturing sector. The cells highlighted in red indicate the highest (top 33%) year-on-year increase, while the cells highlighted in green indicate the lowest (bottom 33%) year-on-year decrease.

TABLE 3. TRENDS IN PUBLICATIONS PER SUBJECT AREA

Year	Year on Year % Difference									
	Engineering	Business management & accounting	Computer science	Economic, econometrics & finance	Environmental science	Agricultural & biological sciences	Biochemistry, genetics & molecular biology	Mathematics	Materials science	Decision science
2019	-	-	-	-	-	-	-	-	-	-
2000	7.33	0.77	6.47	–29.60	–13.05	–18.98	5.76	10.31	27.67	23.39
2001	16.98	2.18	12.50	30.63	6.08	6.89	2.55	38.69	31.27	0.33
2002	0.93	–2.92	3.56	–9.97	–22.28	0.56	–11.86	–10.62	1.35	–21.11
2003	6.96	5.26	8.13	11.48	13.22	21.20	8.43	9.01	15.08	36.69
2004	19.00	3.64	–28.54	1.16	–1.33	–11.99	–6.55	–3.58	–2.32	7.07
2005	4.09	0.18	45.19	2.34	–7.03	–18.43	1.11	14.94	3.68	12.89
2006	–11.12	–12.14	12.51	–11.32	0.36	8.85	25.80	20.38	–14.89	–10.56
2007	–4.56	6.04	10.33	9.70	8.27	1.61	–13.49	–3.73	–1.30	10.67
2008	–3. 94	–2.28	8.17	2.14	–12.20	2.37	8.23	–3.36	–19.47	13.19
2009	–6.45	–7.05	6.27	–2.24	–1.13	6.57	6.53	10.85	7.31	–7.91
2010	1.71	0.85	4.91	25.81	–3.34	–13.46	–19.09	4.00	–25.70	–1.44
2011	–3.45	–2.78	–5.84	–12.88	10.17	24.93	10.07	7.06	–18.15	–12.94
2012	–4.01	–1.35	–2.64	10.95	9.05	–2.55	23.72	–6.71	–100.00	–6.52
2013	–5.91	1.69	–5.24	22.34	–7.64	10.43	11.53	3.92	-	–1.39
2014	3.58	–6.90	0.48	4.47	10.84	4.57	2.95	3.08	-	11.46
2015	–4.77	–1.43	6.52	–11.01	–5.14	–11.98	–4.77	4.59	-	9.26
2016	5.45	10.65	3.62	–3.76	27.82	–2.02	–11.24	7.36	-	15.68
2017	6.36	8.68	4.28	12.43	3.03	6.80	–6.27	20.58	-	16.01
2018	4.75	–10.67	4.91	–22.96	8.47	–4.62	18.73	–11.35	-	19.76
2019	11.33	6.39	12.75	–13.66	8.91	–0.92	–15.85	6.26	-	15.18
2020	–17.60	–13.03	–18.79	–8.35	27.55	–12.23	14.58	–12.69	-	–22.88

Source: own compilation

The results were analysed in clusters of five years; 2000 to 2004, 2005 to 2009, 2010 to 2014, and 2015 to 2019. The year 2020 was analysed separately, as it is an outlier due to the disruption caused by the COVID-19 pandemic. For the 2000 to 2004 period, the publications in the relevant subject areas were erratic, with no subject area having a steady increase or decline. During this period the subject areas with the highest growth were mathematics at 38.69% in 2001, followed by decision science at 36.69% in 2003, and material science at 31.27% in 2001. Advances in material sciences influence the type of packaging utilized by food and beverage manufacturers, which impacts the process and maintenance operations, operational costs, and supply chain, as potentially new suppliers would need to be sourced to enable the use of advanced food packing material options. Decision science influences the approach to decision-making, from investment to operational strategies, hence is critical to business sectors. Thus, for this period, the key skills focus areas were decision-making, material science, and engineering, with mathematics an inherent requirement of all the aforementioned skill areas.

From 2005 to 2009, the subject areas with the highest growth were computer science at 45.19% in 2005, biochemistry, genetics, and molecular biology at 25.80% in 2006, and mathematics at 20.38% in 2006. The growth in computer science highlights the increased adoption of software in all aspects of a business.

- Human resources – For personnel information, leave, performance and training
- Maintenance – task scheduling, assets management
- Finance – invoice and payroll management

The re-occurrence of mathematics emphasises its importance as a fundamental knowledge area. Process control automation and optimization, which is fundamentally based on mathematics is a key knowledge area for personnel in the FoodBev sector, as improved control can result in significant financial, operational, and environmental benefits. The growth in biochemistry, genetics, and molecular biology publications indicated an increasing interest in genetically modified foods. This significantly impacted the food and beverage sector value chain from suppliers to the manufacturing processes to the final customer. This had ancillary effects in terms of skills requirements from operational personnel to engineers to technicians to sales and marketing.

During the 2010 to 2014 period, the subject areas achieving the highest YoY growth were multi-disciplinary at 35.03% in 2013, economic, econometrics, and finance at 25.81% in 2010, and agricultural and biological sciences at 24.93% in 2011. During this period, there was an increase in IT applications, with various systems being deployed to manage the different business operations, such as enterprise resource planning (ERP) systems to manage planning and strategic activities, manufacturing execution systems (MES) to manage operations, and plant control networks (PCN) to manage plant operations at the shop floor level. This necessitated a technically proficient workforce with multidisciplinary skills; operation personnel understanding aspects of the IT system that manages the control and instrumentation on the plant and finance personnel navigating the various functionalities of an ERP finance module. This necessitated upskilling of the current workforce, whilst those entering the workforce were expected to demonstrate a proficiency in computer skills. This increased demand in computer associated skills (both hardware and programming skills) resulted in a boom in students studying IT, IT-related, and programming qualifications. Following the 2008 financial crisis, as economies endeavoured to emerge from the economic recession, an increase in economic, econometrics, and finance publications was observed in the subsequent years. The world's population has increased exponentially in the past century in comparison to previous centuries. This has resulted in an associated increased demand for resources, including food. Newer and efficient methods of agricultural production were being developed to address the rising food demand, which has a downstream effect on the food and beverage manufacturers, as the agricultural sector is the primary raw material supplier in the value chain. Changes in raw materials may necessitate changes in the manufacturing process operations to achieve the final product quality and required throughput. This may require the adoption of new technologies, requiring the upskilling of the workforce. Thus, the increase in agricultural and biological sciences publications.

During the 2015 to 2019 period, decision science demonstrated an average 14.56% year-on-year increase, driven by the fourth industrial revolution of digitalization. Decision science, driven by BDA and artificial intelligence (AI) tools, continues to drive the demand for cognitive, critical thinking, problem-solving, analytical, and technology proficiency skills. This is especially relevant to the food and beverage manufacturers, as operations become digitalized, integrated, and "smart," a dynamic workforce with muti-disciplinary (technical, cognitive, analytical, critical-thinking) skills are required. Environmental science also experienced significant growth during this period, with a peak growth of 27.82% in 2016. This is driven by the global climate change challenge and the goal to reduce the global temperature rise by 2.5 degrees Celsius. As the manufacturing industry is one of the key contributors to greenhouse gas (GHG) emissions, manufacturers are being legislated to reduce GHG emissions, driving the demand for green skills. The top 10 green skills required by employers in 2021, as identified by LinkedIn (LinkedIn, 2022), were sustainability, remediation, occupational safety and health, climate, renewable energy, environmental awareness, environmental health and safety, solar energy, corporate social responsibility, and recycling. Mathematics experienced significant growth in 2017 of 20.58%. This is associated with a rise in digitalization and AI, both of which are mathematically based.

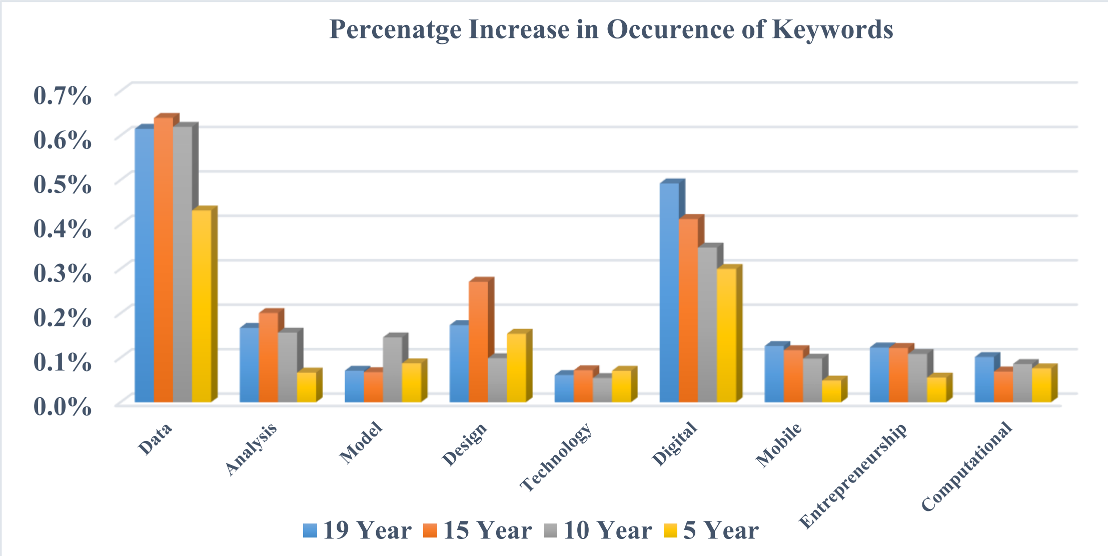
The COVID-19 pandemic disrupted the research focus in 2020, with an exponential shift to the medical sciences as evidenced by increased publications in pharmacology, toxicology, and pharmaceuticals (40.02% increase from 2019), nursing (21.35% increase from 2019) and medicine (10.35% increase from 2019). Interestingly the publications in nursing and medicine have mostly been in decline in the past 20 years. Environmental science (27.55% increase from 2019) was also a focus area in 2020.

The analysis of the change in subject areas provides a suitable indication of the skills required by the associated sectors.

4.2. Analysis of abstract keywords

The abstracts of the various subject areas were consolidated per year and a word count was executed in Atlas.ti. The word count list was cleaned to remove duplicates and stop words and merge synonyms and plurals of the same word. This resulted in a word count of 664 for the 1999 to 2019 period. The year 2020 was excluded as most publications focused on the COVID-19 pandemic, and thus an outlier. The 664 words were analysed to identify the words associated with skills specific to the food and beverages manufacturing sector. This resulted in a final count of 65 abstract keywords, which were used in the analysis. The year-on-year change in keyword count was calculated and aggregated for a period of five year five years (2014–2019), 10 years (2009–2019), 15 years (2004–2019), and 19 years (1999–2019). The skills-related keywords words that demonstrated constant growth across the 19-year period are illustrated in Figure 2.

FIGURE 2. INCREASING ABSTRACT KEYWORDS ACROSS THE PERIOD

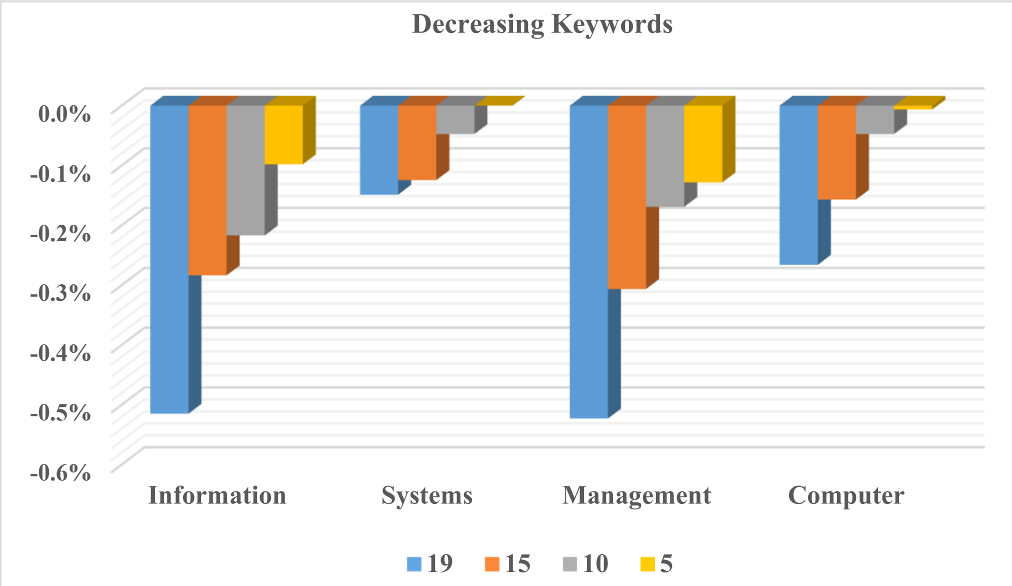


Source: own compilation

The words data and digital demonstrated the highest increases across the five-year clusters. The keyword data had the highest growth across the nineteen-year period of 0.61% followed by digital with 0.49%. The increased occurrence of the abstract keywords of data and digital during the 1999 to 2013 (15 year) period were driven by automation (partial), process optimization, and increased information, and communication technology (ICT) application, while 4IR drove the increase in the last five years. The slower growth in the last five years is attributed to the increased and specific applications of data and digital, resulting in more detailed descriptions of AI, machine learning, natural language processing, robotics, virtual reality, etc.

Similarly, for the 1999 to 2013 period, the increased occurrence of the words analysis, model, design, technology, and mobile were driven by automation (partial), process optimization, and increased ICT application, while 4IR drove the increase during the 2014 to 2019 period. SME growth and sustainability are key drivers of the economic growth of a country. Hence, the skills required by entrepreneurs and SMEs are critical. SMEs and entrepreneurs, especially in developing countries like SA, have limited access to resources (finance, time, people, and skills) to harness the full potential provided by the fourth industrial revolution, thus digital skill enablement of SMEs is critical. The analysis of the abstract keywords defines that skills development initiatives should be focused on (1) data gathering, handling, and analysis, inclusive of AI and modelling, (2) digital tools and technologies, (3) cognitive and critical thinking skills, and (4) digital enablement skills of SMEs. This information can be strategically used by the FoodBev SETA to plan its mandated annual and five-year skill interventions and initiatives for the sector.

FIGURE 3. DECREASING ABSTRACT KEYWORDS ACROSS THE PERIOD



Source: own compilation

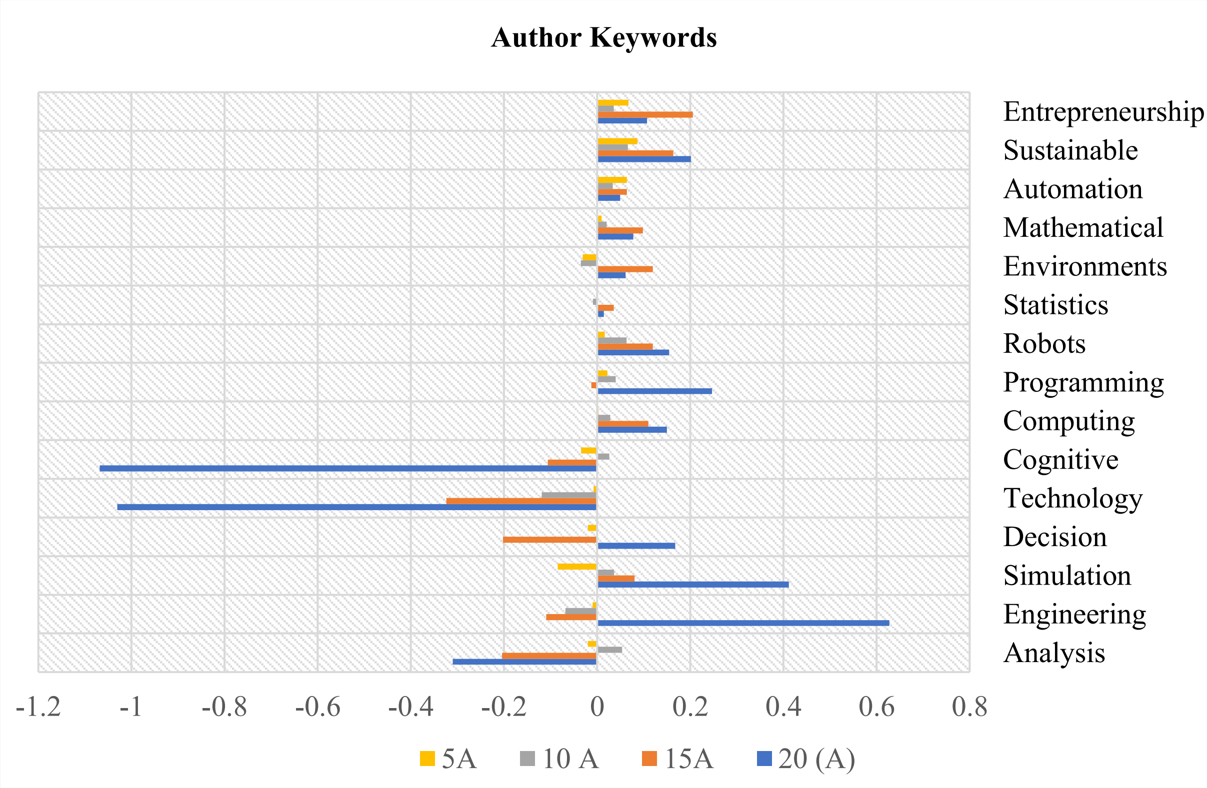
Figure 3 illustrates the skills-related words that have been in continuous decline during the 1999 to 2019 period. The declining abstract keywords are generic, and applicable to a wide range of subject matters, contributing to its decline, as skills and occupations are currently specific and specialised. Past job postings define a broad skills category, such as computer programming skills, but current job postings are highly specific, such as programming skills in Python, data analytic skills in natural language processing, or blockchain skills.

4.3. Analysis of author keywords

Articles indexed in Elsevier Scopus have keywords defined by the author, which are referred to as author keywords. The process, as explained for the abstract keyword analysis is followed for the author keywords. The key difference is that the analysis of the author keywords includes the results of 2020, as consistency in results were observed. The percentage sum of the YoY change is analysed in periods of five years (2015–2020), 10 years (2010–2020), 15 years (2005 –2020) and 20 years (1999–2020), with the results presented in Figure 4.

Sustainable is a broad term that is applicable to various areas, from health and the environment to society. Sustainable or sustainability is synonymous with climate change, a global challenge. The manufacturing sector is one of the largest emitters of greenhouse gases, the key contributor to global warming, and is strictly regulated in terms of permissible volumes of greenhouse emissions. Thus, the manufacturing sector, inclusive of the food and beverages manufacturing, is continuously improving operations, from the adoption of new technologies to cleaner production initiatives to 4IR to lean manufacturing, to reduce environmental impacts towards sustainable operations. Hence, skills associated with environmental impact reduction and process optimization such as cleaner production, green economy, circular economy, value chain optimization, six sigma, lean, and total quality management are in demand. Sustainable is also closely aligned to environmental science, a subject area that has shown significant growth since 2016, as detailed in the subject area analysis.

FIGURE 4. TRENDS IN AUTHOR KEYWORDS ACROSS THE PERIOD



Source: own compilation

The keywords of technology, cognitive, and analysis demonstrate a notable decrease during the 1999 to 2020 period. As ICT and 4IR adoption rapidly expands, the generic terms of technology and analysis are being replaced by specific technology applications and analysis techniques. Technology applications include Internet of things (IoT), robotics, cyber-physical systems, and virtual reality, while analysis techniques include natural language processing, supervised and unsupervised learning, decision forest and machine learning. Similarly, for the word cognitive, employers are looking for specific skills sets in relation to the ability to analyse, interpret, and visualise. Engineering as an author keyword has been in decline for the last 15 years, which could be attributed to the use of specific application terms such as genetic engineering, 3D and 4D printing, and robotic engineers.

The analysis of the subject areas, abstract keywords, and author keywords are interpreted to identify skills for the food and beverages sector, demonstrating the practical application of the approach. The skills identified for the FoodBev sector from this analysis include:

- Entrepreneurship: Inclusive of digital skills to support the growth and development of SMEs. This is especially relevant to the food and beverages sector, as it comprises 90% of SMEs.
- Data analytics: Data gathering, cleaning, analysis, and visualisation, inclusive of AI, simulation, and modelling.

- Digital tools and technologies application: Implementation and operational skills for various technologies such as IoT, virtual reality, 3D printing, and cyber-physical systems.
- Sustainable operations: Process optimization includes green technology implementation and operation, as well as energy, water, and resource optimisation.

To validate the findings of the study, the results are compared to the increasing jobs and the top 10 skills for 2025 as identified by the World Economic Forum (2020). The WEF findings are detailed in Table 4.

TABLE 4. TRENDS IN JOBS AND SKILLS (WEF, 2020)

Increasing Job Roles	Top 10 Skills for 2025
Data analysts and scientists	Analytical thinking
AI and machine learning specialist	Active learning and learning strategies
Big data specialist	Complex problem solving
Digital marketing and strategy specialist	Critical thinking and analysis
Process automation specialist	Creativity, originality and initiative
Business development specialist	Leadership and social influence
Digital transformation specialist	Technology use, monitoring and control
Information security analysts	Technology design and programming
Software applications developer	Resilience, tolerance and flexibility
Internet of things specialist	Reasoning, problem solving and ideation

Source: World Economic Forum (2020)

This study identified the skills of data analytics and digital tools and technologies application as important. This is validated by the WEF findings as all the increasing job roles (as per Table 4) require the skills of data analytics and digital tools and technologies application to some extent. Further to this, the aforementioned skills are aligned with the skills of analytical thinking, complex problem solving, critical thinking and analysis, technology use, monitoring and control, and technology design and programming (as per Table 4).

A comprehensive entrepreneurship skill set would include the skills of leadership and social influence, creativity, originality, and initiative. Similarly, a comprehensive sustainable operations skills set would require the demonstration of the use and application of relevant technologies, applicable monitoring and control capacities and reasoning, problem-solving, and ideations. The alignment of the study findings to the World Economic Forum 2020 report on skills (World Economic Forum, 2020) support the proposed BDA approach for skills prediction.

The identifications of the skills demand of the sector enables the FoodBev SETA, HEIs and government to strategically plan initiatives. This would aid a proactive approach, a shift from the current reactive approach. These initiatives include the bursaries to offer, training programmes to be developed, existing training programs requiring updating, upskilling programmes for the current workforce, and required partnerships. The training programmes and partnerships should facilitate practical training, workshops, and short courses at a minimum. While this would be essential steps in addressing the current skills mismatch and high youth unemployment it is an initial point, as addressing the skills gap is a long-term strategic effort from all stakeholders in the skills ecosystem.

4.4. Assumptions, Limitations, and Bias

A key assumption made is that Elsevier Scopus provides a comprehensive source of data that can be analysed to derive findings. This is based on Elsevier Scopus having 330 disciplines, over 7000 publishers, over 24.6 million open access items, over 368 000 books, and 97.3 million records (Elsevier Scopus, 2025).

The limitations of the study include:

- A single source of information was adopted for the BDA analysis. Utilising multiple sources of information would provide a more comprehensive and holistic understanding of the skills ecosystem. This would be essential in developing a skills ecosystem that is interconnected and dynamic, towards better informed policy and decision making.
- The focus is on the South African Food and Beverage Manufacturing sector only, this would inherently neglect some of the skills requirements of the other sectors of the South African economy. Thus, an economy wide skills approach is recommended as the work force is not restricted to a specific sector.
- The review was limited to publications in the English language only.

Bias is usually present in small to medium sample sizes and has a high prevalence with non-peer reviewed sources. For this study, our sample is every available paper amounting to 747,711 publications, which is a huge sample. We further only used peer reviewed sources implying good data. The subject areas were as determined by Elsevier Scopus, while the keywords utilised were author and index keywords as defined by the authors and the journal respectively. With the above in mind, we declare limited to no bias.

5. Conclusion

Skills determination, current and future, is essential for the economic growth of a country. South Africa, a developing country, is experiencing one of the highest unemployment rates globally. The high unemployment rate intersection with poverty and social decline, necessitates a response. One of the key contributors to high unemployment is the skills mismatch; the mismatch between the graduate skills set and workplace demand. To strategically address this skills mismatch, both current and future, the skills demand must be accurately predicted. The current SA approach is fragmented, as it is per economic sub-sector via the respective SETAs, with no integration to provide an economy-wide perspective on skills requirement. Further, the SETAs approach, based on the analysis of submitted work sector skills plans, by the companies within the sector, is limited.

The skills forecasting models of countries that have been successful in predicting skills demand were analysed. The challenge of these forecasting models is the data requirements, extended periods of data, and the quality of the data. This is especially challenging for developing economies, where data availability and associated quality may be limited.

This study adopted a BDA approach, to predict the SA skills need. The approach entailed extraction of skills related publications from the Elsevier Scopus database, an established peer-reviewed academic database. The publications were extracted and analysed in three categories of subject area evolution, abstract keyword occurrence and author keyword occurrence. This was done to identify the skills demand of the food and beverages manufacturing sector. The skills associated publications for the 1999 to 2020 period were downloaded and analysed. The results were disaggregated per year and analysed in clusters of five years based on YOY change of keywords. The interpretation of the results identified four key areas of skills demand for the South African food and beverages manufacturing sector: 1. entrepreneurship skills inclusive of digital skills; 2. data analytics inclusive of AI, simulation, and modelling; 3. digital tools and technologies implementation, and operational skills; and 4. sustainable operations skills of advanced process control, of energy, water, and resource optimization. The results obtained were compared to the findings of the World Economic Forum 2020 report on skills and was found to align (World Economic Forum, 2020). Further, the skills of sustainable operations align to the findings of the LinkedIn report on global green skills (LinkedIn, Global Green Skills Report, 2022). The proposed BDA approach is generic and can be applied to any of the 21 SETAs in South Africa.

Skills forecasting is essential in reducing the youth unemployment in South Africa. The ability to define future skills planning, enables proactive planning and collaboration by stakeholders in the skills ecosystem. This proactive planning and collaboration will enable the youth to develop the required skills at the appropriate time, facilitating employment. Skills forecasting will enable HEIs and SETAs to strategically plan skills initiatives such that it meets the needs of the required position from artisan to engineer to manager instead of a one fit all skills approach. This would lead to a workforce that is adaptive and sustainable.

References

British Council. (2018). *Futureskills in India foundation.* British Council.

CEDEFOP. (2019). *The online job vacancy market in the EU.* CEDEFOP.

CoPs. (2022). *The Monash model.* University of Melbourne. https://www.copsmodels.com/monmod.htm

DHET. (2019). *Skills, supply and demand in South Africa.* DHET.

Elgendy, N., Elragal, A., & Päivärinta, T. (2022). DECAS: a modern data-driven decision theory for big data and analytics. *Journal of Decision Systems, 31*(4), 337–373. https://doi.org/10.1080/12460125.2021.1894674

Elgral, A., & Klischewski, R. (2017). Theory-driven or priocess-driven prediction? Epistemological challenges of big data analytics. *Journal of Big Data*, 1–20.

Elsevier Scopus. (2025, February). *Scopus content*. Retrieved from Elsevier: https://www.elsevier.com/products/scopus/content#0-content-coverage

Employment and Social Development Canada. (2022). *Canadian Occupational Projection System(COPs).*

ETF. (2016). *Developing skills foresights, scenarios and forecasts. Guide to anticipating and matching skills and jobs (Vol 2).* ETF.

Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods and analytics. *International journal of infromation management, 35*(2), 137–144. https://doi.org/10.1016/j.ijinfomgt.2014.10.007

Governmemnt of Canada. (2023). *Canadian Occupational Projection Systems (COPs)*. Retrieved from https://occupations.esdc.gc.ca/sppc-cops/l.3bd.2t.1ils@-eng.jsp

ILO. (2020). *The feasibility of using big data in anticipating and matching skills needs.* ILO.

ILO-OECD. (2018). *Approaches to anticipating skills for the future of work.* ILO-OCED.

LinkedIn. (2021). *Future of skills*. Retrieved from LinkedIn: https://linkedin.github.io/future-of-skills/

LinkedIn. (2022). *Global green skills report.* LinkedIn.

Maier, T. (2017). The BIBB-IAB Qualification and Occupational Field Projections (QuBe-Projekt.de).

Mezzanzanica, M., & Mercorio, F. (2019). *Big data for labour market intelligence.* ETF.

National Skills Commission. (2022). *Employment outlook: Industry and occupation trends over the five years to November 2026.* Government of Australia. https://doi.org/10.1021/cen-v028n030.p2503

OECD. (2016). *Getting skills right: Assessing and anticipating changing skills need.* OECD.

Sabharwal, R., & Miah, S. (2021). A new theoretical understanding of big datra analytics capabilities in organizations: a thermatic analysis. *Journal of Big Data, 8*(1), 1–23. https://doi.org/10.1186/s40537-021-00543-6

Statistics South Africa (a). (2022). *Quarterly labour force survey.* Statistics South Africa.

Statistics South Africa (b). (2024). *Labour market dynamics in South Africa 2023.* Statistics South Africa.

Statistics South Africa (c). (2024). *Quarterly Labour Force Survey Quarter 3–2024.* Statistics South Africa.

U.S. Bureau of Labor Statistics. (2021). *Employment projections - 2012–2022.*

Wilson, R., Barnes, S.-A., May-Gillings, M., Patel, S., & Bui, H. (2020). *Working Futures 2017–2027: Long-run labour market and skills projections for the UK.* Department of Education.

World Economic Forum. (2020). *The Future of Jobs Report 2020.* http://www3.weforum.org/docs/WEF_Future_of_Jobs_2020.pdf

Declaration Statements

Acknowledgements

The authors thank the Food and Beverages Manufacturing SETA and University of Johannesburg for support of this work.

Conflict of Interest

The author reports no conflict of interest.

Funding

The author received no financial support for this article's research, authorship, and/or publication.

Data Availability

No dataset is associated with this article.

AI Declaration

AI was not used in the write-up of this paper.

Open Access Agreement

Corresponding Author

The corresponding author for this manuscript is Megashnee Munsamy who can be contacted by email via mmunsamy@uj.ac.za.