An Exploration of Digitalization and Information Communication Technology in the Crop Planting Industry in the Western Cape.

Sean Teuchert

A dissertation submitted in partial fulfilment of the requirements of Varsity College,
for the award of Postgraduate in Data Analytics and Machine Learning

11/2022

Declaration

I hereby declare that the work presented in this thesis has not been submitted for any other degree or professional qualification, and that it is the result of my own independent work.

Sean Teuchert ST10041890

28/10/2022 

Abstract

The South African economy relies on the agricultural sector to produce food, export goods, and employ a large portion of its population. The South African crop planting industry is faced with a problem of increases emanating from input costs which change yearly. Information Communication Technology (ICT) and digitalization are new concepts in the farming industry. Studies have shown that these two concepts can have an influence on the industry by reducing input costs and improving efficiency. The goals of this interpretivist research project aim to add to the field of knowledge concerning ICT and digitalization in the crop planting industry. Additionally, it identifies any factors that may discourage farmers from investing into the new digital era. This study posed the following research questions:

- How does digitalization improve decision-making for planting crops?
- What are the limiting factors to adopting the digital era?

To achieve the project goals, the design of the study adopted a mixed method strategy. The target population was farmers in the crop planting industry. The sample population was farmers from the Tulbagh and Malmesbury region in the Western Cape. The methodology comprised of data collection techniques included a case study and semi-structured interviews where content analysis was used to analyse the data. To handle ethical dilemmas, the researcher needed to adopt a professional code of conduct and ensure participants experienced no harm in any form during the project. The results of this research project included a framework of guidelines that; inform farmers of the best practices for crop planting and identify the limiting factors preventing farmers from adopting the digital era. This research project comprised of four limitations: lack of a budget, short completion period, potential language barrier with participants and availability of participants. Future research could explore this concept in other rural areas of South Africa.

Keywords: Agriculture, Crops, Crop-based decision-making, digital twin, digitalization, Information Communication Technology (ICT), Internet of Things (IoT), Information systems.

Acknowledgements

I would like to thank my parents, Denise Teuchert and Paul Teuchert, for assisting and supporting me during this project. My passion for this topic stems from the upbringing I was able to experience as a child which wouldn't have been possible without my parents. I would also like to thank my brother, Dean Teuchert, for emotionally assisting me during this project. Finally, I would like to thank my lecturer Dr Patricia Harpur and my supervisor Dr Rudolf Holzhausen for their guidance throughout this project.

Table of contents

Contents

List of figures

List of tables

Chapter 1: Introduction

1.1 Background and aims

This study explored Digitalization and Information Communication Technology (ICT) in the crop planting industry in the Western Cape. The South African economy relies on the Agricultural sector to produce food, export goods, and employ a large portion of its population. ICT and Digitalization are new concepts in the farming industry, but studies have shown that these two concepts can have an influence on the industry. This study aimed to explore how ICT and Digitalization can be implemented in the process of planting crops and crop-based decision-making. This study also aimed to identify any factors that may cause farmers to be reluctant to adopting ICT and Digitalization. The farming industry plays a large role in the South African economy and has room for improvement. The current technology used by farmers to plant crops lacks digitalization, this increases the difficulty of identifying problems and making decisions. The easiest way to improve profits is to reduce input costs and improve efficiency. The goal of this research was to study how newer technology combined with digitalization can be used to improve planting crops.

The researcher grew up on a farm in the Tulbagh area and has a good understanding of farming practices. Using this domain knowledge and the researcher's personal attachment to the topic as motivation, the researcher aimed to uplift South African farmers by increasing the profitability of planting crops.

1.2 Research Problem

The lack of digitalization for planting crops can result in a poor harvest. Traditional farming and farming equipment in South Africa lacks data tracking which prohibits digitalization. This has resulted in farmers making decisions based on estimations or traditional methods. The input costs of farming are continuously increasing due to inflation and resource scarcity. Farmers must improve their farming methods and decision-making to continue making a profit. If farmers do not adapt to the changing economic environment, it can result in profit loss.

The researcher has performed a preliminary literature review and the following is already known about the topic:

- Digitalization can be used to create crop models and provide real-time information on crop growth, predict phenology and yield (Shekar, 2021).
- Digitalization provides farmers with a diverse set of tools to address food production challenges such as crop loss and environmental impact (Abbasi, et al., 2022).
- Digital technologies enable farmers to gain new real-time information about farming techniques (Kremer, 2020).

Studies have shown that digitalization can be used to improve farming productivity and decision-making. After performing the preliminary literature review, the researcher has highlighted three potential themes to explore. What still needs to be researched is how digitalization can be used to improve planting crops, how can it improve crop-based decision-making and what are the limiting factors to adopting this new technology.

The intended audience of this study was:

- South African farmers
- Farm equipment manufacturers
- Any researcher interested in the planting of crops.
- Academics

This study can help improve farmer decision-making by maximizing digitalization and increasing knowledge about planting crops. This study explored how Information communication Technology combined with Digitalization can be used to improve planting crops.

The problem stated highlights the slow decline of traditional farming methods due to the continuous increase of input costs and resource scarcity. The lack of digitalization can be a result of farmer's experiencing poor crop harvest. The research problem was satisfied by answering the research questions and achieving the research objectives.

1.3 Research Objectives

This study aimed to:

- Add to the field of knowledge about digitalization in the crop planting industry.
- Investigate how ICT and digitalization can improve crop-based decision-making.
- Identify any factors that may discourage farmers to invest into ICT and digitalization.

1.4 Research Questions:

The researcher used the following two research questions:

- How does Digitalization improve decision-making for planting crops?
- What are the limiting factors to adopting the digital era?

1.5 Dissertation structure

This section discusses the structure of this dissertation. Additionally, it compromises of the following chapters:

- Chapter 1: Introduction – Outlines the background, research problem, and research objectives and questions
- Chapter 2: Literature review – Outlines the information already known about the topic
- Chapter 3: Methodology – Outlines the methods used in this project
- Chapter 4: Empirical analysis and results – Discusses the results obtained during data collection
- Chapter 5: Conclusions and recommendations – Presents conclusions drawn from results and recommendations

Chapter 2: Literature review

2.1 Introduction

This preliminary literature review serves several purposes. It:

- Outlines the theoretical foundations of the study.
- Provides knowledge on what is already known about the topic.
- Guides the reader to what is still not known.
- Links to confirmation that the research is valid (Bezuidenhout, et al., 2014).
- This review is written by addressing the topic associated with Digitalization, Agriculture, and planting crops. Five categories of literature have been identified.

2.2 Internet of Things (IoT)

Internet of things refers to an interlinked system of computer devices, sensors, appliances, and machines connected with the internet, each having unique identities and capabilities for performing remote sensing and monitoring (Abbasi, et al., 2022). IoT has empowered farmers to gather real-time data or advice via a quick phone call or web browser search (Kremer, 2020). It enables farmers with the ability to improve on the current farming methods and techniques. (Abbasi, et al., 2022) state that IOT empowers farmers with decision tools and automation technologies that integrate knowledge, products, and services to improve productivity, quality, and profit.

Farming has always been a skill-based industry. IoT aims to transform traditional skill-based agriculture into knowledge-based and technology-driven digital agriculture (Himesh, 2018). This will transform the farming industry into a state-of-the-art industrial sector.

Abbasi and co-authors discuss how remote sensors can be implemented to gather data about crops and soil. The example used is sensors and monitors on tractors. These sensors and monitors can gather real-time soil and crop data. The soil data can be used to identify what nutrients are needed to ensure a healthy crop.

2.3 Precision Farming

Precision farming is closely linked with Digital farming. Digital technology empowers farmers to be more precise. Traditionally farmers would look at a large field as a whole. Satellite and Drone imagery allows farmers to view afield as an area of land of 100 or more smaller fields and automates disease detection (Menne, 2022).

An example of how this technology can be implemented is spraying fungicides. Traditionally farmers would view afield as one large area and spray the same quantity of fungicide across the whole field. Precision farming allows farmers to split the field into sections and spray the exact quantities of fungicide needed per section of the field (Menne, 2022). This ensures that the crop receives the correct amount of fungicide and reduces input costs by reducing wastage.

Abbasi et al. (2022) discusses the use of sensors to collect soil information and weather data. Clapp and co-authors reinforce this statement. Digital farming technologies can be combined with site-specific data and data analytics to assist farmers with decision-making (Clapp & Ruder, 2020). Traditional farming methods that lack digitalization cannot reach this level of precision.

Clapp and co-author discuss how precision farming promotes environmental sustainability. Digital technologies can be implemented to improve crop performance in hostile climate conditions. Digital farming enables decision-making regarding seeds and chemical use based on soil conditions and weather patterns to maximize crop yields (Clapp & Ruder, 2020). Climate conditions are forever changing. Precision farming provides farmers with a tool to adapt to climate change.

2.4 Unmanned Vehicles

The use of Unmanned Aerial Vehicles (UAVs) impacts the planting process by performing efficient monitoring and spraying activities, thus optimizing the capabilities of pesticides and fertilizers (Maddikunta, et al., 2021). Drones or UAVs can be used as a data collection method. UAVs reduce time spent on collecting data and cultivate better yields based on analysed data (Saha, et al., 2018).

Saha et al. (2018) suggest that drones can be implemented to plant seeds and spread fertilizer. This method of planting and fertilizing is less time-consuming and has fewer input costs. Alka reinforces this statement, drones can be used for planting crops which eliminates labour and tractors. Eliminating tractors saves fuel and reduces carbon emissions (Rani, et al., 2019).

Drones can be implemented to control Irrigation. Drones that are equipped with sensors have the capability to pinpoint heat and water stress in crops (Saha, et al., 2018). Drones can detect and apply irrigation to crops when required. This will reduce water wastage and overwater crops.

Maddikunta and co-authors discuss the challenges facing the use of UAVs. The success of UAVs depends on the user's knowledge and skills. UAVs are sophisticated and high-tech. Farmers with limited or no skills will find the usage of UAVs to be challenging (Maddikunta, et al., 2021).

2.5 Digital Twin

The concept of a digital twin in agriculture refers to creating a digital copy of a farm. (Angin, et al., 2020) propose a project that involves a digital twins-based smart agriculture framework. This system makes use of sensor networks in the farm fields and intelligent processing of aerial images for plant disease and nutrient deficiency detection. They conclude that these technologies can be used to help farmers increase crop yield at reduced production costs.

According to the study performed by Pyrandis et al. (2021), Digital Twins are still in early development. Digital Twins can streamline operations, and fuse information about physical twins by taking multi-angle perspectives and quantifying the possible outcomes to be used for decision making.

2.6 Digital Agriculture Development Constraints

Digitalization and information technologies can have a high input cost but the return on investment is worth it. A constraint for the development of digital agriculture is the digital divide in terms of internet access and trained personnel in digital technologies (Rasputina, 2022).

A study was performed in the Smolensk (Russia) region (Mironkina, et al., 2020). Mironkina and co-authors conclude that there are four limiting factors for the development of digital agriculture. First, Farmers are unaware of the benefits digital technology can provide. Second, the high input cost of modern technology makes them reluctant. Third, reliable mobile internet coverage in Smolensk is poor. Fourth, a lack of financial and human resources is common in the region. Many regions in the world suffer from similar problems to the Smolensk region. Due to these factors, the development of Digitalization and adoption of digital technologies in Agriculture has been stunted.

2.7 Theoretical Approach

The theoretical approach selected by the researcher for this study was the Technology Acceptance Model (TAM). The Technology Acceptance Model is an information systems theory that models how users come to accept and use technology (Davis, 1989). This theory was developed by Fred Davis and Richard Bagozzi. It was further improved by Davis and Venkatesh in more recent times. The basis of this study involves information systems and human interaction with these systems; therefore, it is an appropriate theoretical framework for the intended study.

Figure 1 outlines the concepts of TAM and graphically illustrates it:

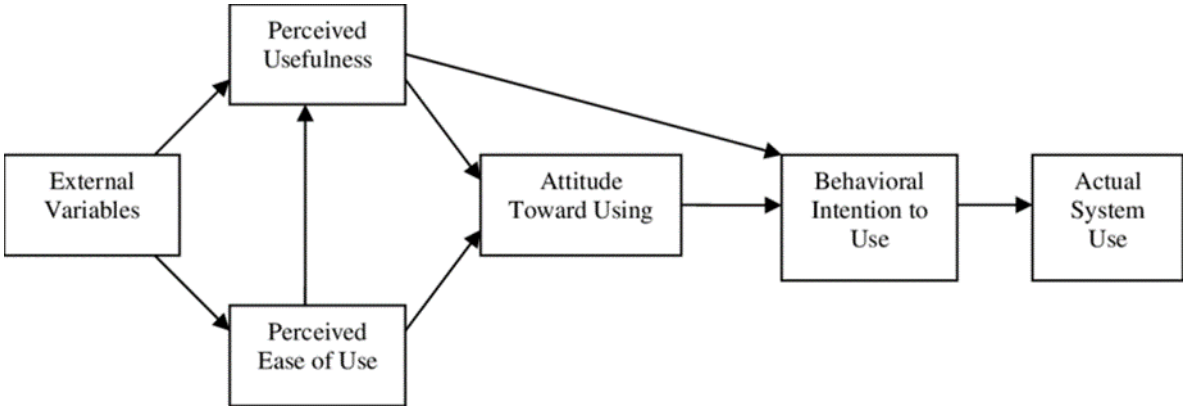


Figure 1 Technology Acceptance Model Diagram(Khera,2010)

(Davis, 1989) states that when a user is presented with a new technology, several factors influence the user's decision about how and when to use it. These external factors can be split into two categories: Perceived usefulness and Perceived Ease of Use. Perceived usefulness refers to the degree to which a person believes that using a particular system would enhance their job performance (Davis, 1989). Perceived Ease of Use refers to the degree to which a person believes that using a particular system would be free from effort (Davis, 1989). These categories have an influence on the user's attitude towards using the system and how they actually use the systems compared to its intended use. This study aims to explore ICT and Digitalization in the crop industry; therefore, it is important to understand how farmers view and interact with these systems.

2.8 Conclusion

Planting crops is a multi-level process that requires farmers to pay attention to finer details. Once the seed has been planted in the ground the process is not over. Depending on the climate conditions, fertilizer or pesticides may be required. The Internet of Things has provided farmers with the technology to monitor crops and farm more precisely. The more information that is gathered via the use of sensors and drones empower farmers to make better crop-based decisions. Drones and UAVs are high tec or sophisticated technology but have the potential to significantly reduce input costs when planting. Digital Twins are still in early development but can be implemented to automate operations and disease detection. The three major factors prohibiting the adoption of Digitalization and digital technologies in farming are: a lack of knowledge, a digital divide and high installation costs. Digitalization in planting crops has the potential to increase profits by reducing input costs and eliminating unnecessary labour. A noticeable gap is forming between the new digital era and the older generation of farmers. The older generation lacks digital technology skills or does not recognize the full potential of digital technology (OECD, 2019). This study aims to increase the field of knowledge about digitalization and planting crops to bridge the gap between the new digital era and traditional farming.

Chapter 3: Research Method

3.1 Introduction

This section outlines the research methods used for the study. Additionally, it discusses:

- Research Design
- Research Methodology
- Research Ethics

3.2 Design

This section outlines the research paradigm and design used for the study. Additionally, it discusses ethical issues, and it comprises the following sections:

- Research Paradigm
- Research Design

The construction of the research methodology is based on the theoretical concept called the Research Onion, proposed by (Saunders, et al., 2019).

Figure 2 illustrates the Research Onion:

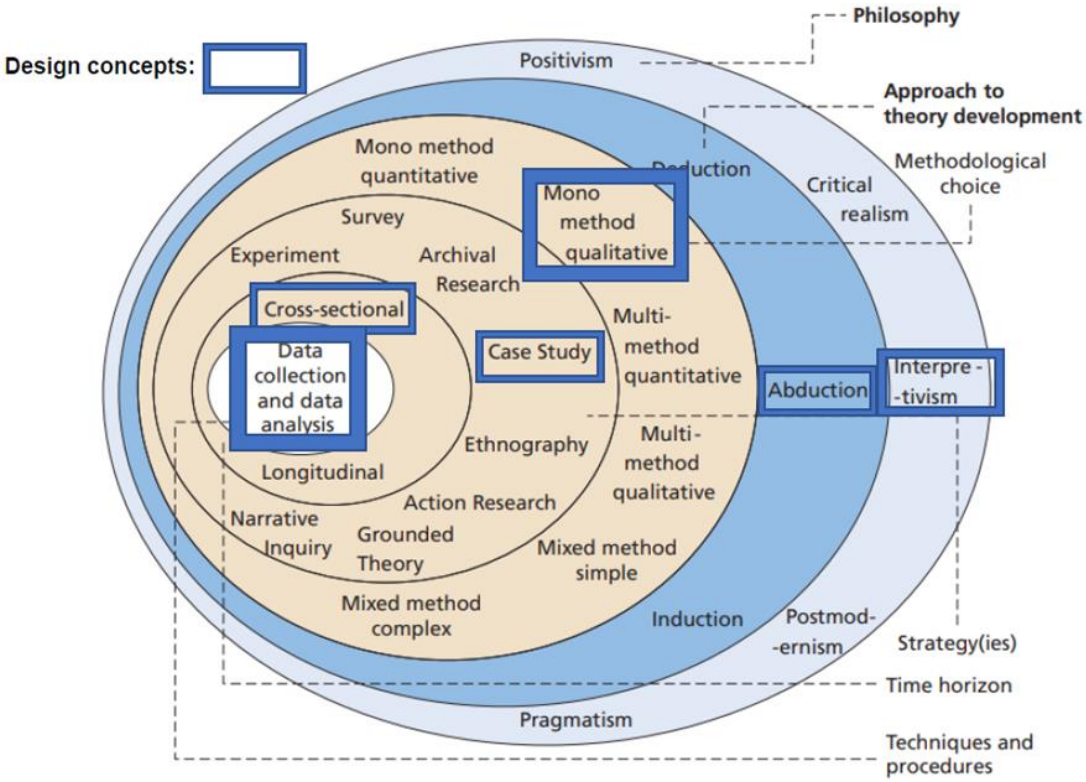


Figure 2 The Research Onion (Saunders, et al., 2019)

Saunders, Lewis and Thornhill (2019) break down the research methodology into six sections: philosophy, approach, strategies, methodological choice, time horizons, and Techniques and procedures of data collection and analysis. The research design and paradigm has been constructed using the principles of the Research Onion created by Saunders, et al., 2019.

3.2.1 Philosophy

This was an information technology (IT) study; therefore, the focus of the study was on the exploration of digitalization and Information Communication Technology (ICT) in planting crops in the Agricultural sector. Interpretivists believe that truth is dependent on people's interpretation of facts (Bezuidenhout, et al., 2014). The author aimed to conduct the study by adopting the principles of Interpretivism and Epistemology.

3.2.2 Approach

The literature review performed in chapter 2 highlighted that the field of knowledge about Agriculture and Digitalization is still developing but historical data and theories could be found. A Deductive approach aims to develop a hypothesis based on existing theory (Saunders, et al., 2019). This study explored the existing field of knowledge and added to it by using an abductive approach. Section 3.2.3 highlights how the researcher applied this approach.

3.2.3 Strategy

The research strategy was a guideline to determine how the researcher collected data to answer the research questions. The purpose of this Interpretivist study was to explore digitalization and ICT in planting crops. The researcher applied triangulation to collect and analyse the data. The initial approach to theory development was deductive by performing a literature review (Saunders, et al., 2019). The researcher adopted an inductive approach when collecting and analysing the data. The research strategy implemented by the researcher was a combination of a case study and survey, specifically semi-structured interviews.

3.2.4 Choices

This study explored how digitalization and ICT can be implemented to improve planting crops and what factors are limiting the adoption of digitalization in the Agricultural sector. To answer the problem statement, an Interpretivist approach was adopted to understand the qualitative factors behind the problem, a mono method approach was applied. The researcher performed qualitative data analysis but presented the results visually using quantitative methods in Excel.

3.2.5 Time Horizons

This study has adopted an Epistemology approach; therefore, it was important to perform the study without influencing the data or participants. Section 3.1.2 highlighted the researcher using an abductive approach which aligns with the concepts of Cross-Sectional study. The researcher aimed to answer the research problem using a Cross-sectional time horizon. This is the most appropriate approach due to the limited completion period of this project.

3.2.6 Techniques and Procedures

A mixed-method data collection approach has been applied to answer the research question. The researcher explored the research problem by using a Case Study and Interviews. The researcher has used convenience sampling to select participants. The researcher analysed the data using content analyses to identify the hidden interpretation behind the transcribed data (Bezuidenhout, et al., 2014).

3.3 Methodology

This section outlines the research methodology used for the study. It comprises the following sections:

- Population sampling.
- Data collection methods.
- Outline of Discussion list.
- Data analysis methods.
- Proposed contribution, validity, reliability, trustworthiness.

3.3.1 Population and Sampling

This study satisfied the research questions by selecting a population from the Western Cape (South Africa), specifically from the Tulbagh and Malmesbury region. The population consisted of five farmers. The researcher had access to farmers in the Tulbagh and Malmesbury region that were willing to participate in the study. The population was selected using Convenience sampling, this study had no budget and limited time to be completed therefore it was important to save costs and time.

3.3.2 Data Collection Methods

This study solved the research questions by using a Case Study. The case study was be performed on a farm that mainly uses traditional planting methods. The case study method that will be used was semi- structured interviews. Interviews provided the researcher with an insight into the practicality of the problem. The researcher performed five interviews with five different farmers. The duration of the interviews was between ten and twenty minutes. The interview was recorded using a mobile phone and laptop microphone. The interviewee was made comfortable by ensuring the researcher was wearing a face mask, offered tea or coffee, and catered to any possible needs of the interviewee that were within reason.

3.3.3 Outline of Discussion List

This section outlines the questions that the researcher asked participants in the interview. The researcher asked the following questions during the interview:

The researcher started with two probing questions:

- What is your farming background or experience?
- What kind of technology do you use to plant crops?

The researcher followed up with four topic questions:

- Do you use any form of ICT or Digitalization when planting crops?
    - Inquire if answered yes
        - What type of ICT or Digitalization?
        - Improved or decreased farming ability in anyway?
    - If participant answered no to question above:
        - Can you please give me a summary of your planting process?
        - What kind of labour is needed (number of workers)?

- Do you see any business value in investing in ICT and Digitalization?
    - If answered yes:
        - Inquire what value it adds?
    - If answered no:
        - Inquire their opinion on the topic

- Before making a crop-based decision, do you refer to any source of knowledge?
    - How does this source improve your decision making?

- Are there any factors that are restricting you from involving ICT and Digitalization when planting crops?
    - Inquire about the factors and how affects them

The interview was a semi-structured interview with a few open-ended questions. The researcher adopted a standardised, open-ended interview approach (Bezuidenhout, et al., 2014). Each interview consisted of the same set of questions but depending on the answer provided by the participant the follow up question was adapted to meet the flow of the interview, this ensured that the participants remained on topic and helped to avoid repeated answers.

3.3.4 Data Analysis Methods

This study solved the research problem by using content analysis to analyse the data. According to Saunders et al (2019), when text or content analysis is used to analyse text, it is employed for both text and transcribed data. Performing content analysis provided the researcher with the tools to identify hidden interpretations in the transcribed data and summarize the data.

(Miles & Huberman, 1994) suggest that qualitative data analysis consists of three procedures: data reduction, data display and conclusion drawing/verification. The researcher adopted this approach when analysing the data collected from the interviews. To ensure the interviewees' comfortability, the interviews were conducted in the interviewee's home language (Afrikaans). The researcher translated and transcribed the interviews in English. To ensure trustworthiness by eliminating bias when translating the interviews, a third party was asked to confirm that the translation was a true reflection of the interviewee's answers.

The researcher attempted to perform content analysis using Python, Appendix D contains screenshots of the code used. A pipeline was used to clean the data of null values, stop words and punctuation. The researcher used a Wordcloud to visually represent the frequencies of common words used by farmers during the interview. The researcher used an N-gram to visually represent the frequencies of common phrases used by farmers during the interview. The researcher identified a flaw with this method, due to the interviews being translated similar words or phrases were used by farmers but worded differently. This resulted in the Wordcloud and N-gram being unable to accurately capture the patterns and trends in the data.

After identifying the flaw in using Python to perform content analysis, the researcher adopted the principles of content analysis set by Miles and Huberman (1994). The researcher has reduced the data by selecting relevant information to research questions and performed the five stages of data coding. The researcher adopted three coding techniques: open coding, axial coding, and selective coding. Appendix A contains the translated interviews from Afrikaans to English and Appendix B contains the three different coding techniques used during analysis.

The first stage of data coding consisted of identify all statements relating to the research question and assigning it a code or category. The second stage of data coding consisted of searching for statements that fit into the specific categories. The third stage consisted of eliminating unnecessary codes by combining codes that were very similar. The fourth stage consisted of selecting data that explains or backs up the codes. The last stage consisted of organising the data and creating a diagram to visually represent the codes and categories. The researcher identified six categories with its corresponding codes, refer to Appendix B: Figure 14 & 15.

The researcher identified six categories from the collected data that linked to the preliminary literature review, refer to Chapter 4: Table 1. The researcher has used content analysis to identify the underlying themes of the data and presented the data in a quantitative manner by using Excel to generate graphs and tables.

3.3.5 Proposed Contribution and Trustworthiness

This section outlines the proposed contribution for this study. Additionally, it discusses the principles of trustworthiness stated by (Bezuidenhout, et al., 2014). The proposed contribution of this study that aims to add to the field of knowledge of:

- Agricultural Crop planting.
- Crop-based decision making.

Bezuidenhout et al. (2014) break down trustworthiness into four concepts. These four concepts have been implemented by the researcher to ensure trustworthiness:

- Credibility.
- Transferability.
- Dependability.
- Confirmability.

The researcher addressed credibility by implementing triangulation into the study and ensured all research findings were made available to participants. The researcher applied known concepts in the research design and methodology, this increased the transferability of the study. The researcher collected thick and rich data and analysed the data using content analysis, this ensured the findings were dependable. The researcher implemented a well-known data analyses method (Content analysis). This ensured that the data analyse process was thorough and the findings properly represented the data which increased the confirmability of the study. By using established data collection and data analyses methods enabled the researcher to describe the research process fully and be open to scrutiny.

3.4 Ethics

Ethics are the researcher's moral or professional code of conduct that sets a standard for the researcher's attitude and behaviour (Saunders, et al., 2019). A researcher's ethics affects all stakeholders in the research. To ensure that no harm was experienced by individuals involved in this research (physically, mentally, or spiritually). It was important to address the following ethical concepts (Saunders, et al., 2019):

- Anonymity.
- Confidentiality.
- Informed Consent.
- Participation is voluntary.

The researcher promised to uphold these ethical principles by using aliases for all business and individual names involved in the research to ensure Anonymity; requesting permission before publishing any information or using dummy figures about said business or individual to ensure Confidentiality; provided an in-depth explanation of what the research was about and how the information provided by said business or individual will be used to ensure Informed Consent; making sure that said business or individual understood the research was voluntary and can request to stop being involved/retract any information provided to ensure participants are not under duress.

The researcher identified three possible limitations for the study. The three limitations were:

- Time.
- Budget.
- Availability of participants.

This study had no budget and limited time to be completed within. Due to the time of the farming season, the researcher was able to interview five farmers and overcome all the limitations for the study.

3.5 Conclusion

Using the Research Onion designed by Saunders (2019), the researcher has developed a research design to satisfy the research question by using triangulation. The methodology adopted by the researcher consists of a case study, specifically semi-structured interviews where content analysis was used to analyse the collected data. The researcher has addressed research ethics by adopting a professional code of standards set by Bezuidenhout et al. (2014). The next chapter will discuss the results obtained after performing content analysis.

Chapter 4: Empirical analysis and results

4.1 Introduction

This section outlines the results that were uncovered during qualitative data analysis. It compromises of three sections:

- Introduction to research problem
- Role of Digitalization in Crop-Planting Industry
- Digital Agriculture Development Constraints

The South African economy relies on the agricultural sector to produce food, export goods, and employ a large portion of its population. The South African crop planting industry is faced with a problem of increases emanating from input costs which change yearly. The aim of this study was to add to the field of knowledge about digitalization in the crop planting industry, investigate how ICT and digitalization can improve cop-based decision-making, and identify any factors discouraging farmers to invest in new technology.

4.2 Role of Digitalization in Crop Planting Industry

This section outlines the role that ICT and Digitalization plays in the crop-planting industry by using descriptive statistics and content analysis. Additionally, it compromises of the following sections:

- Technology used by Farmers
- Sources of Knowledge used by Farmers
- Themes identified from Interviews

4.2.1 Technology used by Farmers

Figure 3 illustrates the types of technology famers used while planting:

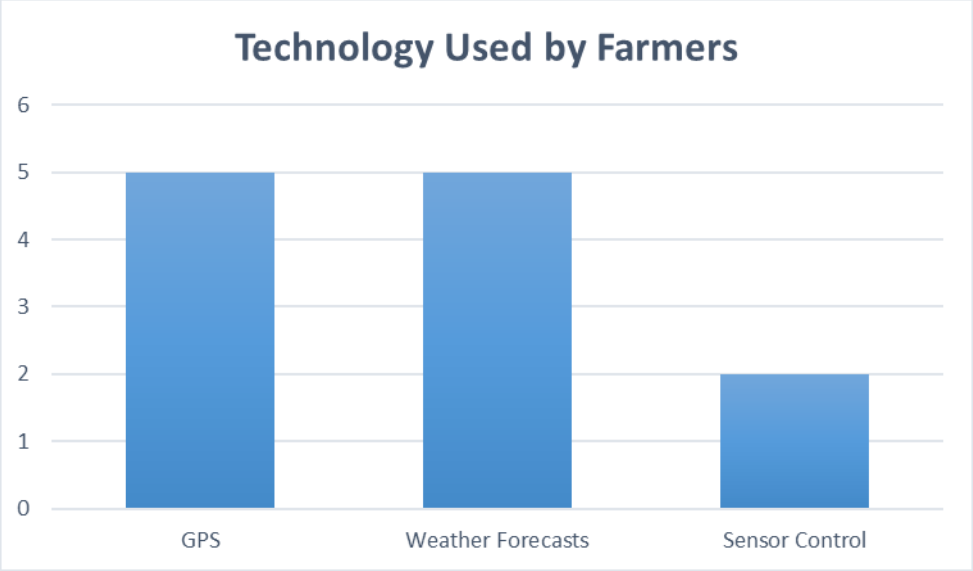


Figure 3 Technology used by Farmers (Teuchert,2022)

According to Figure 3 and the data collected during the research, the most used technology by farmers in the Tulbagh and Malmesbury was: Global Positioning System (GPS), followed by sensor control and weather forecasts. This information was recorded by analysing each interview. Each time a farmer referred to a type of technology they used it was recorded. Each farmer that was interviewed has used GPS and weather forecasts when planting crops while only two farmers included sensor control in their farming practice. (Kremer, 2020) states that the Internet of Things empowers farmers to gather real-time data, this statement is backed up by figure 3. Farmers in the area are accustomed to using some form of ICT and digitalization during crop planting. Abbasi and co- authors discuss how remote sensors can be implemented to gather data about crops and soil. Two farmers stated that they use sensor control during crop planting, this confirms the statement made by (Abbasi, et al., 2022).

4.2.2 Sources of Knowledge used by Farmers

Figure 4 illustrates the sources of knowledge each farmer referred to before planting crops:

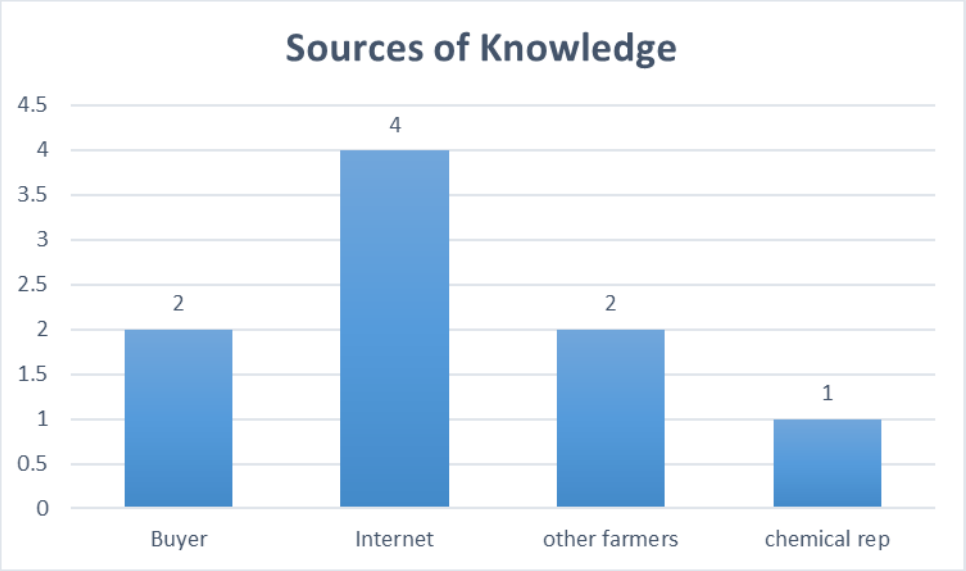


Figure 4 Sources of Knowledge used by Farmers (Teuchert, 2022)

According to figure 4, the most common sources of knowledge that farmers referred to in the Tulbagh and Malmesbury area were: the internet, the buyer or other farmers and chemical consultant. The farmers that made use of the internet mainly used it for weather forecasts, to check on market prices and input cost prices.

One of the interview questions inquired farmers about if they see any business value in ICT and digitalization. From the researcher's perspective only one farmer had an overall negative opinion about ICT and digitalization, but each farmer was in agreeance that ICT and digitalization increases business value. Reading further into this farmers attitude and the answers they provided, the farmer has a negative connotation about ICT and digitalization due to previous failed business ventures. The farmer was reluctant to discuss this topic in detail and the only reason provided for this failure was due to a lack of information. Farmers in the Tulbagh and Malmesbury use basic technologies like GPS, weather forecasts and the internet. After using basic Information Communication Technology and digitalization the farmers are in agreeance that it adds business value. This aligns with the statement made by (Abbasi, et al., 2022) that the Internet of Things can improve current farming methods and techniques.

4.2.3 Themes identified from Interviews:

The researcher identified two main themes during content analysis: Digitalization and Digital Agriculture Development Constraints (discussed further in section 4.3). This section focused on digitalization and its four sub-themes which link to the literature review performed in chapter 2.

The four themes that were identified in the data were:

- Internet of Things (IoT)
- Precision Farming
- Unmanned-Vehicles
- Digital Twin

The data that was collected during the interviews allowed the researcher to identify codes and group the information into the corresponding themes and categories. The researcher has linked the categories to the themes that were identified in the literature review. Each category has a corresponding code that contains information from the interviews to justify each category and code. The researcher created a themes or code diagram to gain a visual representation of the themes in the data. The researcher has broken down the main diagram into smaller diagrams to focus on each theme and to make it easier to read, refer to Appendix B: Figure 14 to view the main diagram.

Figure 5 illustrates the codes associated with the Internet of Things:

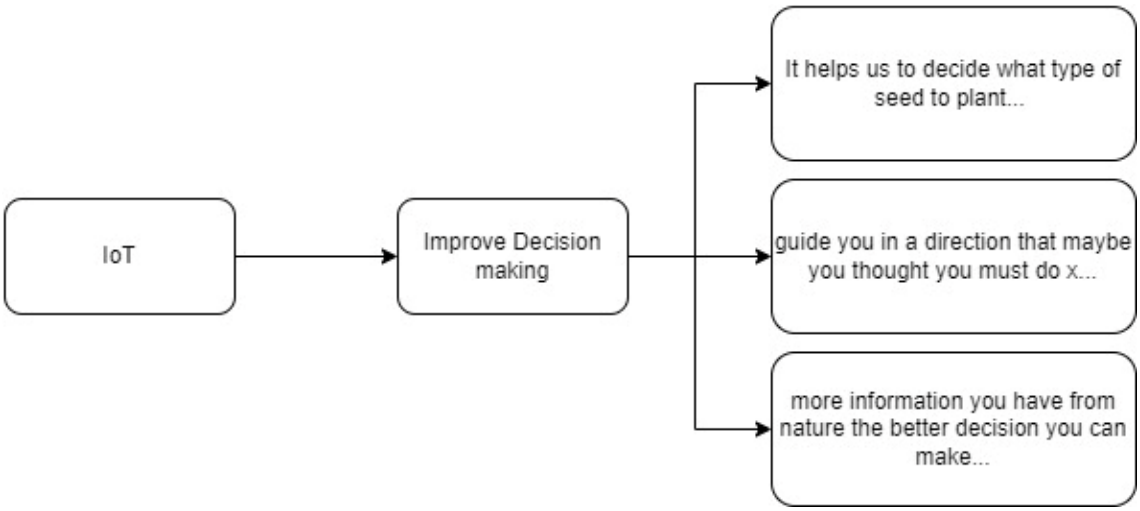


Figure 5 Internet of Things Code Diagram(Teuchert, 2022)

The internet of Things helped farmers to make more informed decisions therefore improving their crop-based decision making. This aligns with the statement made by (Himesh, 2018) that the Internet of Things aims to transform traditional skill-based agriculture into knowledge-based and technology-driven agriculture. Abbasi and co-authors discuss how the Internet of Things empowers farmers with decision tools which the researcher has also identified during analysis.

Figure 6 illustrates the codes associated with Precision farming:

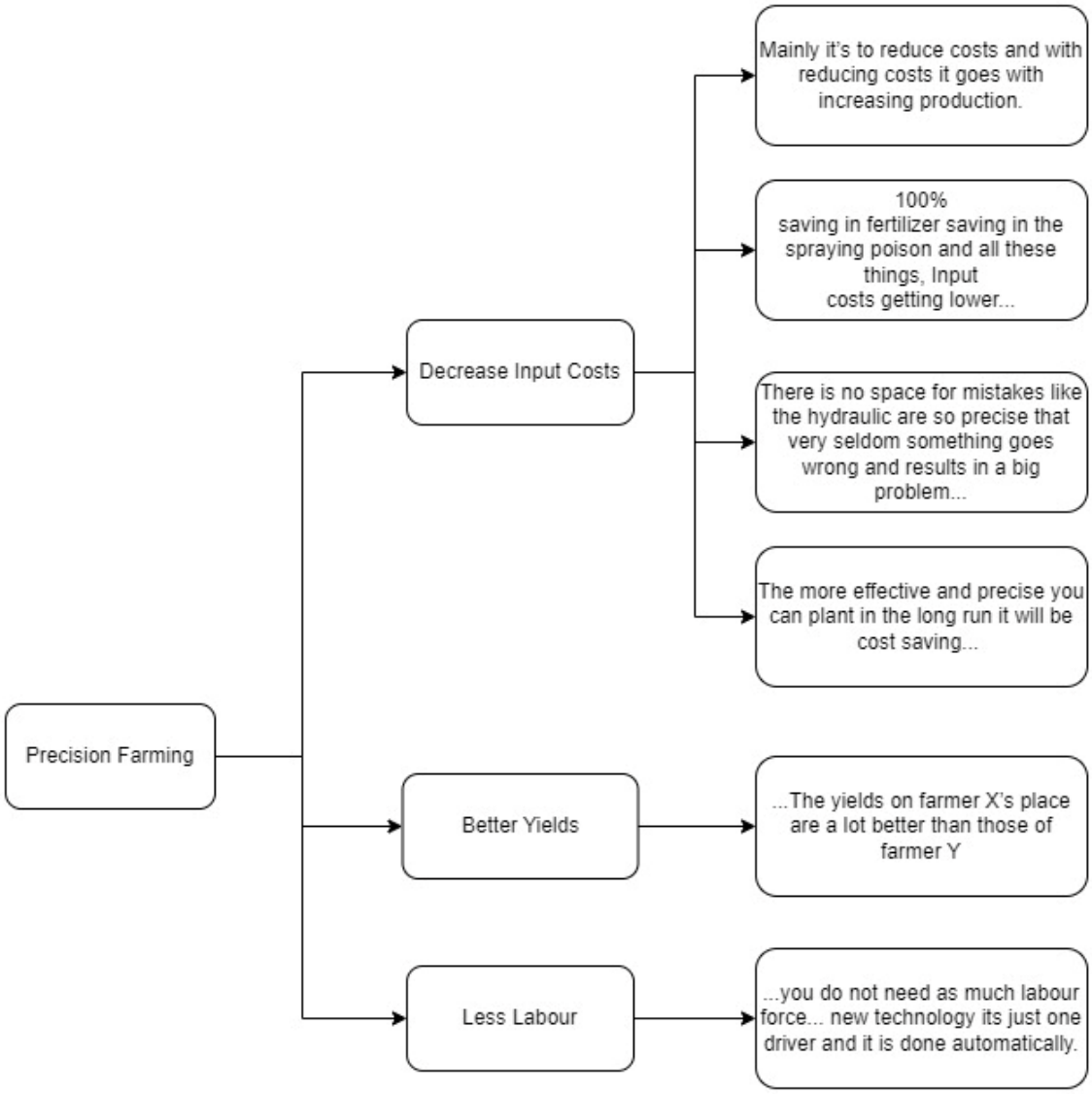


Figure 6 Precision Farming Code Diagram(Teuchert, 2022)

Precision farming reduces input costs, requires less labour, and increases yields during harvesting. (Menne, 2022) discusses how technology can be used to improve the process of fertilizing, this aligns with the concepts discussed during the interview with Farmer A, refer to Appendix A. Automating the planting process reduces the number of labourers required, as explained in the interview with farmer E in Appendix A, traditional planters require three or four labourers to ensure that the planter is running correctly while implementing sensor control and GPS only requires a tractor driver to check digital monitors. (Clapp & Ruder, 2020) state that digital farming enables decision-making

regarding seeds and chemical use based on soil conditions and weather patterns to maximize crop yields. This statement aligns the statement made by Farmer E about how digital technologies increase crop yields.

Figure 7 illustrates the codes associated with Unmanned-Vehicles:

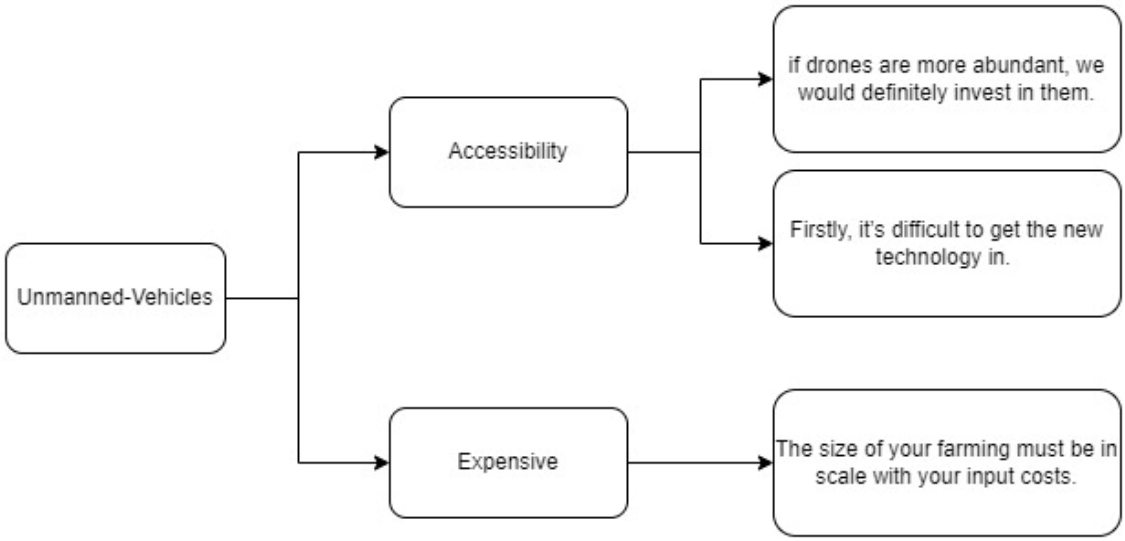


Figure 7 Unmanned-Vehicles Code Diagram(Teuchert, 2022)

Unmanned-Vehicles, specifically drones, were ideas farmers liked and they were aware of the benefits it may provide but due to being in South Africa drones are still too inaccessible and expensive for farmers to consider an option. During the interview with Farmer C, the farmer referred to a concept called spot spraying which can be performed using a drone. Spot spraying according to the farmer is the process of locating a specific location within a field that needs to be sprayed for fertilizer and only spraying that specific location rather than spraying the whole field. This aligns with the concepts explained by (Mironkina, et al., 2020) and (Saha, et al., 2018).

Figure 8 illustrates the codes associated with Digital Twins:

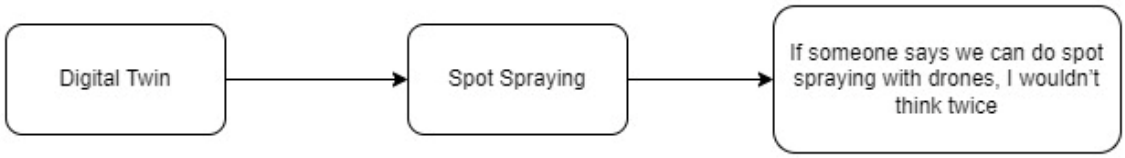


Figure 8 Digital Twin Code Diagram(Teuchert, 2022)

Farmers in the Tulbagh and Malmesbury area are inexperience with Digital Twins, but the researcher identified a possible task that can be performed using a digital twin. Spot spraying can be performed using drones to map a field and identify which areas of the field need to be sprayed or fertilized. (Angin, et al., 2020) discusses how digital twins can be used for plant disease and nutrient deficiency detection.

4.3 Digital Agriculture Development Constraints

This section outlines the factors that discourage farmers from investing into ICT and digitalization. It compromises of the following sections:

- General Constraints
- Constraints in South African Context

The researcher has identified two categories under the main theme of Digitalization constraints. The researcher has identified general constraints that might be experienced internationally and constraints that are specific to South Africa. The researcher has linked the categories to the themes that were identified in the literature review. Each category has a corresponding code that contains information from the interviews to justify each category and code. The researcher created a themes or code diagram to gain a visual representation of the themes in the data. The researcher has broken down the main diagram into smaller diagrams to focus on each theme and to make it easier to read, refer to Appendix B: Figure 15 to view the main diagram.

Figure 9 illustrates the codes associated with general constraints:

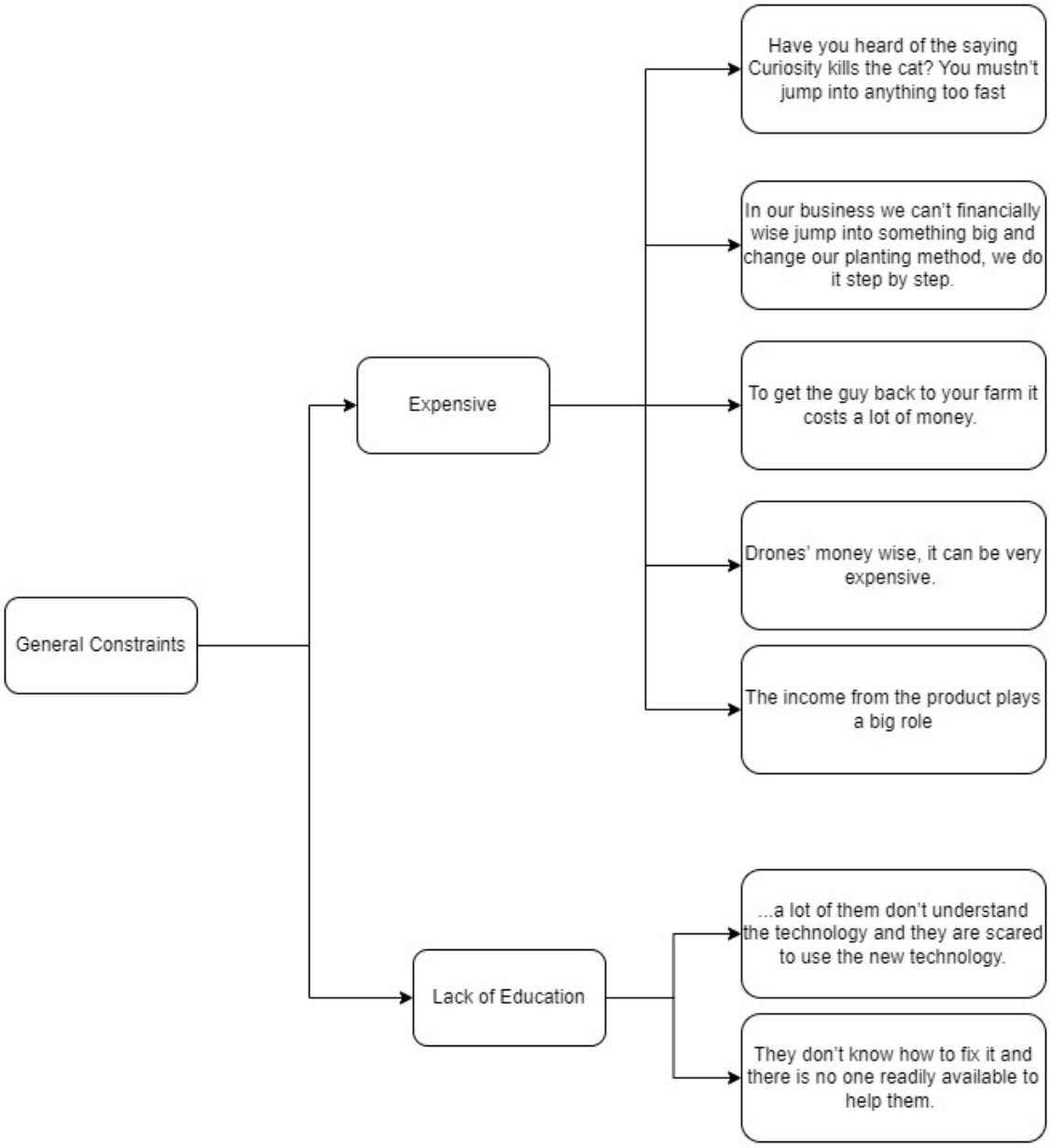


Figure 9 General Constrains Code Diagram (Teuchert, 2022)

The factors that may discourage farmers in general to invest into ICT and digitalization are investment cost being too high and a lack of education. Due to the nature of new technology, it can be very expensive to implement which discourages farmers to invest or they finically cannot support such business ventures. This concepts links to the fourth limiting factor for the development of digital agriculture discussed by (Mironkina, et al., 2020). A lack of education also played a role, when technology problems occurred

farmers couldn't solve it themselves which resulted in down time and farmers reverting to traditional methods. Rasputina (2022) discusses how the digital divide and a lack of trained personal in digital technologies can be a constraint towards the development of digital agriculture.

Figure 10 illustrates the codes associated with Constraints in South African context:

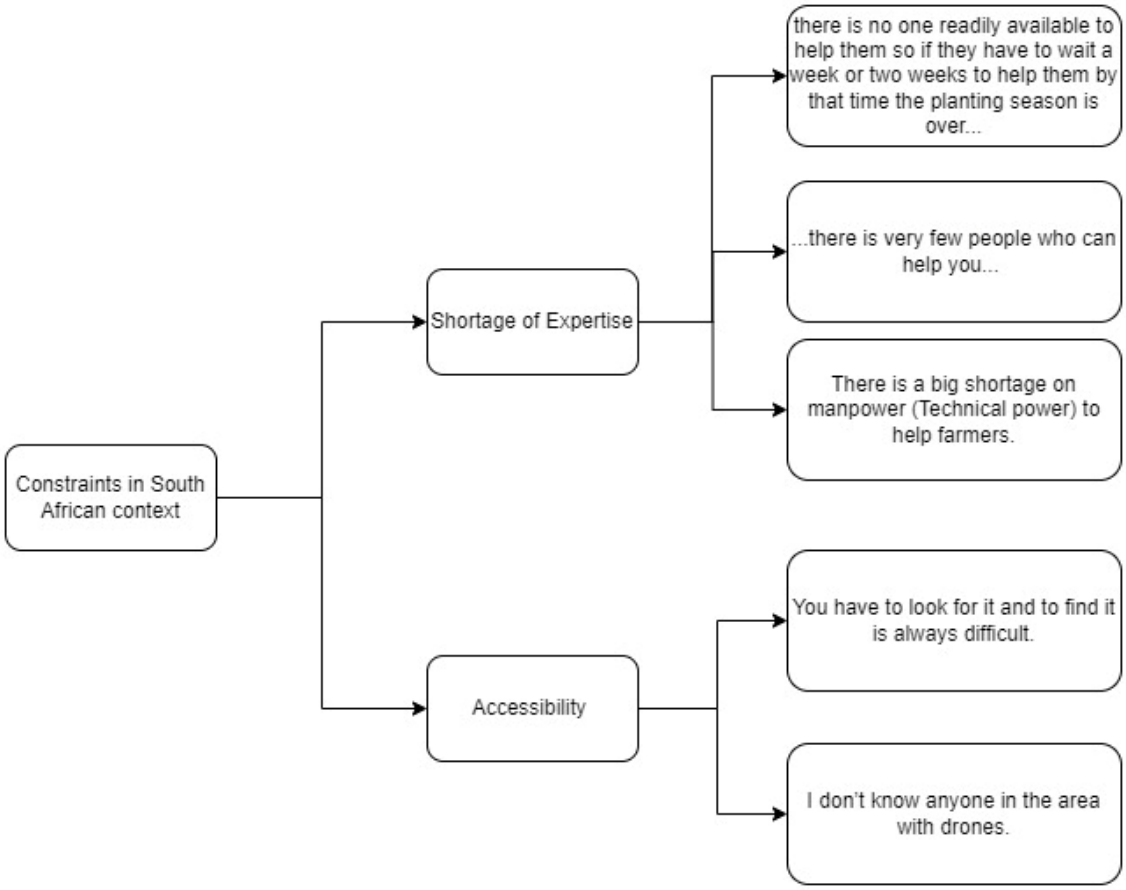


Figure 10 Constraint in a South African Context

The factors that may discourage South African Farmers to invest into ICT and digitalization are a shortage of expertise and lack of accessibility. A common theme that occurred during analysis is that farmers in South Africa do not understand how the technology they use during planting works. This has resulted in South African Farmers to outsource domain experts in that technology which is expensive, and these experts are hard to come by. Due to the shortage of expertise, farmers end up waiting for weeks or months to receive help and by the time they receive help the planting season is over.

Rasputina (2022) discusses how lacking trained personnel in digital technologies is a constraint for the development of digital agriculture, this is clearly an issue in South Africa. South African farmers lack the skills to use digital technologies and rely on third party expertise. There is a shortage of expertise in digital technologies for agriculture in South Africa which has resulted in farmers having to resort to traditional methods due to long waiting periods and the cost of hiring this expertise being too high. This aligns with the second limiting factor stated by (Mironkina, et al., 2020) that high input costs of modern technology make farmers reluctant to invest.

The second factor that may discourage South African Farmers to invest into ICT and digitalization is the lack of accessibility. According to the farmer B in Appendix A, new technology isn't easily sourced and cannot be found in a shop on the shelf in South Africa. This has resulted in farmers like Farmer D in Appendix A, that sourced their new technology via the internet from other countries. Rasputina (2022) states that a constraint to the development of digital agriculture is the digital divide. South African farmers experience this problem too. The farmers in the Tulbagh and Malmesbury areas have access to basic digital technologies, but it is hard to source more advance technologies. If a farmer wants to invest in the newest technology, it will most likely have to be imported which is very expensive and makes sourcing spares difficult.

Figure 11 illustrates the Digital Agriculture Constraints experienced by interviewees:

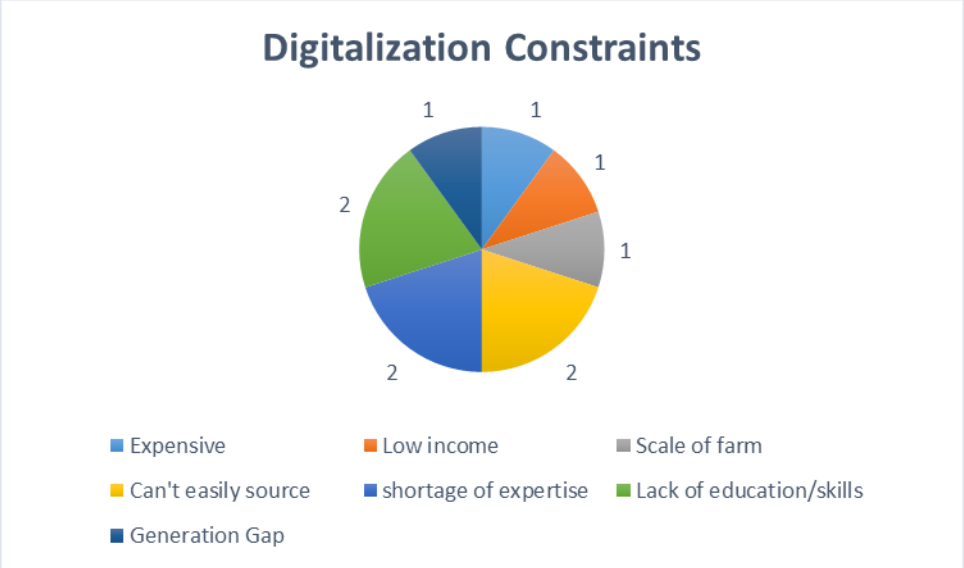


Figure 11 Digitalization Constraints Pie Chart

The pie chart above summarizes the digitalization constraints experienced by participants. The most common constraints experienced by participants is a lack of education or skills and shortage of expertise. Farmers in the Tulbagh and Malmesbury area are faced with similar limiting factors to international farmers when comparing the information gathered from the literature review e.g., expensive, low income, generation gap. There are certain limiting factors that mainly apply to South African Farmers or 3rd world country farmers e.g., lack of education/skills, shortage of expertise and can't easily source the equipment.

4.4 Links between Literature review and collected data

Table 1 illustrates the links between the themes in the literature review and collected data:

Table 1 Links between literature review and collected data

	Digitalization				Digital-Agriculture Development Constraints	
	Internet of Things	Precision Farming	Unmanned-Vehicles	Digital Twin	General Constraints	South African Contraints
Improve decision making	(Abbasi, et al., 2022)					
	(Kremer, 2020)					
	(Himesh, 2018)					
Decrease input costs		(Menne, 2022)				
Better yields		(Clapp & Ruder, 2020				
Less labour		(Abbasi, et al., 2022)				
Spot Spraying				(Angin, et al., 2020)		
				(Pyrandis, et al., 2021)		
Expensive			(Maddikunta, et al., 2021)		(Mironkina, et al., 2020)	
Accessibility			(Maddikunta, et al., 2021)			(Mironkina, et al., 2020)
Lack of education					(Rasputina, 2022)	
					(Mironkina, et al., 2020)	
Shortage of Expertise					(Rasputina, 2022)	(Rasputina, 2022)

The researcher has identified themes in the preliminary literatures review that was performed before collecting and analysing the data. Table 1 outlines how the themes identified in the preliminary literature review links to the themes identified in the collected data. The researcher has adopted triangulation as one of the research methods, Table 1 illustrates how the literature review backs up the results of the data analysis. Implementing this research method will increase the trustworthiness of the research project.

4.5 Content Analysis using Python:

This section outlines how Python can be used to perform content analysis. Additionally, it compromises of three concepts:

- Data pre-processing
- Wordcloud
- N-gram

The code for this section is included in Appendix D. The researcher has performed data pre-processing by removing punctuation and stop words using a pipeline to clean the data. A Wordcloud was used to create a visual representation of the frequency of words used during the interviews by interviewees.

Figure 12 illustrates the Wordcloud for question 2.1:

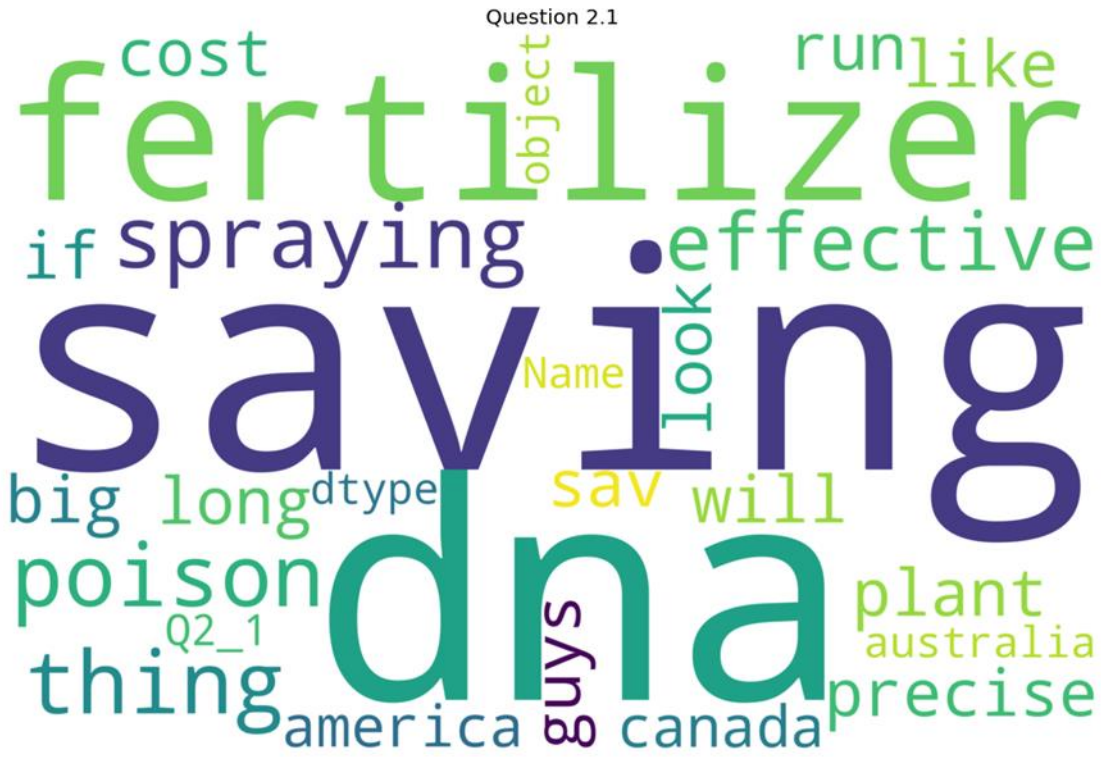


Figure 12 Question 2 Wordcloud

From Figure 12, we can interpret that the most used words were saving, fertilizer, effective, spraying and dna (refers to no answer). A problem with the Wordcloud method that the researcher identified is that when translating the interviews, it resulted in a large vocabulary and the same concepts were referred to but in different terms. The researcher has still included this method in the results to show that different approaches were adopted to produce the most accurate results.

The researcher also used N-grams to perform content analysis. An N-gram allows the researcher to identify what two-word phrases are most used during the interviews. The idea behind this concept is to help the researcher identify key terms like input costs or weather forecasts.

Figure 13 illustrates the N-grams for Question 2 and 3:

Figure 13 Question 2 and 3 N-grams

From figure 13, the researcher can get a visual representation of what two-word phrases were most used during the interview like weather forecast or using GPS. The N-gram method had the same problem as the Wordcloud method. The translated answers referred to similar concepts but were worded differently.

4.6 Conclusion

Farmers in the Tulbagh and Malmesbury area are accustomed to using digital technologies when planting crops and sourcing new knowledge. The limiting factors to adopting newer technology are high investment cost, accessibility, lack of education and lack of expertise

Chapter 5: Conclusions and recommendations

5.1 Introduction

As input costs continue to increase yearly, South African farmers are being forced to improve their farming practices and methods. Information Communication Technology (ICT) and digitalization has the potential to evolve the farming industry into a more profitable industry. South African farmers are also faced with limiting factors that prevent them from investing into the new digital era. This section compromises of the following sections: Conclusion, Recommendation, Future Work and Future Development.

5.2 Conclusion

The researcher aimed to answer the following research questions:

- How does Digitalization improve decision-making for planting crops?
- What are the limiting factors to adopting the digital era?

The research objectives were satisfied by the research questions and the researcher was able to explore ICT and Digitalization in the crop planting industry in the Western Cape. The researcher can conclude that Famers in South Africa may not know it but do use a form of ICT and digitalization. All the farmers that were interview agreed that digitalization adds business value, and the yearly input cost increase is a problem. The main benefits of ICT and Digitalization that were identified in this study are: improves decision-making, reduces input costs, increases crop yields, and reduces labour. This study identified four factors that may discourage farmers from investing into ICT and digitalization. A lack of education and high installation costs are two factors that discourage farmers in general while South African farmers are discouraged by a shortage of expertise and accessibility. South African Farmers do not understand how the digital technologies they use work and rely on third party expertise to solve digital technology related problems. This aligns with the theoretical approach selected by the researcher. According to the Technology Acceptance model (TAM), users' attitude towards technology depends on two categories: perceived usefulness and perceived ease of use.

This study has identified that South African farmers do understand the benefits of ICT and digitalization, but due to a lack of education in digital technologies they do not understand how to operate this technology. Therefore, the researcher has pinned this down to the farmers perceived ease of use being the result of a lack of investment into the digital era. The researcher was faced with the following three limitations: time, budget, availability of participants. The researcher overcame these limitations by using accepted methods to save time and used convenience sampling to ensure the study remained in budget and participants were available. A future direction for this study is to research into the factors that play a role in the perceived ease of use of technology by farmers.

5.3 Recommendations:

The crop planting industry is a seasonal industry which means there is a limited time to plant crops. Farmers will not invest into new technology if it reduces efficiency due to having to wait for someone to help them solve their problems. The researcher recommends to businesses that create new digital technologies for the planting industry to create a support system for these technologies. Farmers will not invest into technology if they do not understand it or have access to help when they are faced with a problem. An example of a support system would be providing farmers with training workshops which ensures the farmer has the correct set of skills to fix common problem. Another example of a support system is to create a maintenance team that's available for 24/7 callouts which guarantees a fast response time.

5.4 Future Work and Future Development:

During this research project the researcher identified the limiting factors that prevent South African farmers from investing into the new digital era. The researcher did not explore how to combat these limiting factors, therefore future research can still be explored on this topic. The researcher made two recommendations that were not explored in depth and alternative methods could be more appropriate. Exploring this topic further will contribute to the field of knowledge about digitalization in agriculture and uplift South African farmers to a more profitable future.

"Once a new technology rolls over you, if you're not part of the steamroller, you're part of the road."

Stewart Brand

References

Abbasi, R., Martinez, P. & Ahmad, R., 2022. The digitization of agricultural industry – a systematic literature review on agriculture 4.0. *Smart Agricultural Technology,* 2(2), p. 100042.

Angin, P. et al., 2020. AgriLoRa: A Digital Twin Framework for Smart Agriculture. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications,* 11(4), pp. 77-96.

Bezuidenhout, R., Davis, C. & Du Plooy-Cilliers, F., 2014. *Research Matters.* 1st ed. Claremont: Juta.

Clapp, J. & Ruder, S.-L., 2020. Precision Technologies for Agriculture: Digital Farming, Gene-Edited Crops, and the Politics of Sustainability.. *Global Environmental Politics,* 20(3), pp. 49-69.

Davis, F. D., 1989. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology.. *MIS Quaterly,* 13(3), pp. 319-340.

Himesh, S., 2018. Digital revolution and Big Data: a new revolution in agriculture. *CAB Reviews,* 13(21).

Kremer, M., 2020. *Weforum.* [Online]
Available at: https://www.weforum.org/agenda/2020/07/digital-agriculture-technology
[Accessed 20 04 2022].

Maddikunta, P. K. R. et al., 2021. Unmanned Aerial Vehicles in Smart Agriculture: Applications, Requirements, and Challenges. *IEEE Sensors Journal,* 21(16), pp. 17608-17619.

Menne, T., 2022. *CropLife.* [Online]
Available at: https://croplife.org/digital-revolution-takes-farming-to-another-level/#:~:text=What%20is%20digital%20farming%3F,while%20also%20increasing%20their%20efficiency.
[Accessed 20 04 2022].

Miles, M. B. & Huberman, A. M., 1994. *Qualitative Data Analysis: An Expanded Sourcebook.* 2nd ed. London: SAGE.

Mironkina, A., Kharitonov, S., Kuchumov, A. & Belokopytov, A., 2020. Digital technologies for efficient farming. *IOP Conference Series: Earth and Environmental Science,* 578(1).

OECD, 2019. *OECD Library.* [Online]
Available at: https://www.oecd-ilibrary.org/sites/367ac383-en/index.html?itemId=/content/component/367ac383-en
[Accessed 20 04 2022].

Pyliandis, C., Osinga, S. & Athanasiadis, I. N., 2021. Introducing digital twins to agriculture. *Computers and Electronics in Agriculture,* 184(1), p. 105942.

Rani, A. et al., 2019. HARIT DHARA 2(1) DRONE: THE GREEN TECHNOLOGY FOR FUTURE AGRICULTURE. *Harit Dhara Magazine,* 1(1).

Rasputina, A. V., 2022. Digitalization trends in the agricultural industry. *IOP Conference Series: Earth and Environmental Science,* Volume 979, p. 012009.

Saha, A. K. et al., 2018. IOT-based drone for improvement of crop quality in agricultural field. *IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC),* 1(1), pp. 612-615.

Saunders, M. N., Lewis, P. & Thornhill, A., 2019. *Research Methods For Business Students.* 8th ed. Harlow: Pearson.

Shekar, R., 2021. *Botany One.* [Online]
Available at: https://www.botany.one/2021/08/the-future-of-digital-technologies-for-crops/
[Accessed 20 04 2022].

Appendix A: Interviews

Farmer A

Q1: What is your farming background or experience?

I was born on the farm, grew up on the farm I was running school from the farm into town each and everything day on my bicycle, so I think that's a very good learning experience that you get there from living on a farm and growing up on a farm. Officially I've got no what you call it after school education or diploma nothing about that. that's what I see I learn from, my father my brother and by reading. Internet is very interesting and very good learning curve nowadays if you understand it, you will learn a lot from it. I'm on the farm now for 14 years back after 27 years but still a learning curve each and every day if I don't know you go and check on the Internet and you get good the learning curve.

Me talking…

Q2: What kind of technology do you use to plant crops?

Yeah, we are using GPS not directly for planting but for spraying because the planter has wheels you can work off. The sprayer is 10 m that side and 10m to the other side, so you don’t know where the line is so that’s where we use the GPS. You need to check weather forecasts like the last week two days we had wind where we can’t do any spraying , Thursday we could have done a little bit of spraying and Friday nothing, this morning again so yeah weather forecasts is something you need to be able to have access to and not one application but two or three, that one says so and that one says so , you can check on the average.

Me talking…

Q2.1: Do you use any form of ICT or Digitalization when planting crops?

100% saving in fertilizer saving in the spraying poison and all these things, Input costs getting lower, you not spraying double sewing fertilizers double. A meter wide doesn’t

sound very much but at the end of the day if you are doing 100 hectors and over spaying by a meter or two it's a lot of money.

Me talking…

Q3: Do you see any business value in investing in ICT and Digitalization?

Definitely because at the end of the day, you not saving money on one side you are also gaining money because if you spray correctly, you don't over spray, you don't get dead areas in your land, so everything is picking up at the end of the day and the value is going back into your pocket

Me talking…

Q4: Before making a crop-based decision, do you refer to any source of knowledge?

We always spoke to the buyer at the end of the day. If KaapAgri this year wants just oats, we will go 90% to the oats are, if they just want wheat, we will go to the wheat area. Each and every year they have got an information session and letters that they send out that says wheat prices this year is very good, oats prices is not that good, go that route or that route. We are making us of the, we call it "landbou kinderness" and they are helping us in the right directions. At the end of the day it's your own decision but they can only inform you of what the trends are

Me talking…

Q.4.1: How does this source improve your decision making?

It will guide you in a direction that maybe you thought you must do x but they are tell you they are doing Y and Z, you can incorporate all three of them and maybe get a good result but it all depends on the weather. It's a guidance or a guideline they are giving.

Me talking… set up next question by referring to drones

Q5: Are there any factors that are restricting you from involving ICT and Digitalization when planting crops?

Drones' money wise, it can be very expensive. Especially in SA where it's not widely used yet as far as I know. Some guys in Limpopo with citrus have crop observation and by

Citrusdal. The wheat is bringing in 10k a hector, citrus or peaches is bringing in 100k a hector, the machine costs you 250k. Is it worth to have a 250k machine on a 100 hector of wheat but on 100 hectors of citrus it's worth it. The income from the product plays a big role if you have a big long pocket it will work. Your farming area must be big enough to get a new fancy planter with all these new sensors on and all those kind of things and it costs you 2.5 mil and you only planted 100 hectares that sum is not going to work out in the end of the day. The size of your farming must be in scale with your input costs.

Me talking…

Farmer B

Q1: What is your farming background or experience?

I grew up in the Tulbagh area and went to Agricultural school. I have been involved in fruit farming, cattle, sheep, goats and planting different types of crops. I've been farming for 40 years.

Me talking…

Q2: What kind of technology do you use to plant crops?

At this stage we are only using GPS and the GPS calculates how many hectares I have planted.

Me talking…

Q2.1: Do you use any form of ICT or Digitalization when planting crops?

The more effective and precise you can plant in the long run it will be cost saving because you put in the correct amount of seed and fertilizer per hectare. Better technology you use, the less input costs you have. Because input costs have risen so much, you must be more precise using more or less depending on the type of soil.

Me talking…

Q3: Do you see any business value in investing in ICT and Digitalization?

Yes because the input costs have risen so high. You need to try and reduce your input costs because you can only get X amount per hectre. With the technology available you must try to use it as best as possible to cut costs.

Me talking…

Q4: Before making a crop-based decision, do you refer to any source of knowledge?

We use the internet for weather predictions and for the cost of wheat. The consultants tell us if the world price is going to go up or down. They will tell you if there is a shortage in the world market but in SA there is always a shortage so you must plant. I use the internet to find the costs of fertilizer locally vs importing. I use Yr and long Yr to get weather forecasts for the next three months to determine if you must plant earlier or later

Me talking…

Q.4.1: How does this source improve your decision making?

The more information you have from nature the better decision you can make, and all the information is not always correct especially the weather forecasts. You can also see if your fertilizer and poison costs are going up or down. If all the prices go up and the wheat price is low, then you are not going to plant wheat and look at alternative crops like canola. The decision is normally between canola or wheat.

Me talking… set up next question by referring to drones

Q5: Are there any factors that are restricting you from involving ICT and Digitalization when planting crops?

Firstly, it's difficult to get the new technology in. Where do you look for it? Its not always on a shelf that you can just go and pick it up. You have to look for it and to find it is always difficult. On the other hand, one of the big problems if you have got that equipment and it breaks there is very few people who can help you but like with the planter's seed feeding system breaks there are too few people who can fix it. There is a big shortage on manpower (Technical power) to help farmers. It can take up to two to three weeks before someone can come and repair it. The new equipment the people come and show you to use it one day and if you forget something in two to three weeks' time or if something goes wrong with the program or you put in the incorrect settings, to try and get someone to help you is just about impossible. To get the guy back to your farm it costs a lot of money and there are too few of them to help.

Farmer C

Q1: What is your farming background or experience?

I grew up on a farm. Through primary school and high school, I helped out on the farm. I studied for 4 years at university, not sure if I can say the university, then afterwards I worked in America for two years farming and this is my first-year farming full time in South Africa in the western cape. I did a BSC in agriculture degree in soil science.

Me talking…

Q2: What kind of technology do you use to plant crops?

We put GPSs in the tractor because it brought up productivity because you don't skip an area or double plant an area which keeps our costs low and increases efficiency. What we changed this year, normally we plant the field and afterwards we do the perimeter. This year we planted the perimeter first and then the middle section. We saved time and less double planting, and it brought down our costs. We changed the planter to it was a traditional mechanical driven planter we changed that to hydraulics so that we can apply variable rates if we wish to, or we can do a standard application.

Me talking…

Q 2.1 Do you use any form of ICT or Digitalization when planting crops?

If you look at the big guys like in Canada, America , Australia they got very big grain carts, they can run for instance, I'm not sure about the values but for like 150 hectares they can run consistently without filling up your productivity is so much than filling up every 20 or 30 mins , having technology like blockage monitors like we have on your planter so that you can just check everything from your cab that everything is running perfectly. You can't see what is happening behind you so you have to trust that the system is working correctly but if you have blockage monitors or systems in place that doesn't need you to climb out of the cab to check if everything is alright that can improve efficiency by a lot. The longer you can sit in the cab the more efficient you are.

Me talking…

Q3: Do you see any business value in investing in ICT and Digitalization?

Yes, I think business wise like with the fertilizer prices that are sky high at the moment. There is no space for mistakes like the hydraulic are so precise that very seldom something goes wrong and results in a big problem. If you plant incorrectly for a day, you create big problems and then you have to replant that whole area again which results in double costs. In business the lower your input costs the better your profit margin is. I get paid to run the tractor, the tractor hasn't got a union, its got diesel costs, wear and tear and all those factors results in your costs going up. Your costs go up because you have to fix your mistakes. You are not going to actually save costs by planting with variable rates because your soil differs so much, but you are definitely going to see better harvesting planting with variable rates which results in more profit.

Me talking…

Q4: Before making a crop-based decision, do you refer to any source of knowledge?

In this situation now, well my dad is 62 years old, he said that there was trial and error, that's the best thing and how you learn. You have made your mistakes in your past but now in this moment if you get the recipe right or let's say for instance you have more of an idea of how the recipe works you stick with what you got. We sometimes go on the internet and study some stuff if we have problems like our chemical rep can't help us with or we consult our neighboring farmers for extra help but like if we fertilizer a field mostly what we look at weather wise, how much rain we have got this season. That's the most important thing, how much rainfall we have gotten and that will determine how much fertilizer you are going to put in.

Me talking…

Q.4.1: How does this source improve your decision making?

Before we plant, we will check what seed variety we have planted last year, and we change what seed variety we plant this year on that specific field. There is only a limited amount of varieties we can use in this area. It helps us to decide what type of seed to plant in the area and how much fertilizer to use.

Me talking… set up next question by referring to drones

Q5: Are there any factors that are restricting you from involving ICT and Digitalization when planting crops?

We don't have much experience with Drones, I don't know anyone in the area with drones. Someone that sticks with traditional farming like technology is there to make your life easier for instance if drones are more abundant, we would definitely invest in them. If someone says we can do spot spraying with drones, I wouldn't think twice. The reason we don't use it is because we don't know about it, or we don't know about anyone using it at the moment. In our business we can't financially wise jump into something big and change our planting method, we do it step by step. Someone my age I'm 25 years old and I'm more familiar with technology so its easier for me to figure it out. Its easier for me to use the technology like I know how a cellphone and computer works but an older farmer that does not know about technology will not use it but younger farmers like me are more willing to. I would use the technology by trial and error and not change the whole farming practice at once and I would use it on a portion of the farm and if you see better results for one season then you change the whole farming method. As long as the technology is giving you a better profit, I would change the whole farming method.

Farmer D

Q1: What is your farming background or experience?

I have a bachelor's degree in agriculture and an honors degree in agronomy. I started farming in 1984 and I have 40 years farming experience.

Me talking…

Q2: What kind of technology do you use to plant crops?

For the last ten years, we have used GPS auto track. You don't want to know about harvesting and other stuff. What we do is yield mapping and a result of that we put in fertilizer and lime. About five years ago I tried variable rate planting, but the machine wouldn't do it, so I imported it from Canada, but we had problems with that machine then last year we fixed out machine to be able to do variable planting with sensor control. We are very skeptical about the variable rates when it comes to planting, but we are happy with it when it comes to lime spreading. We don't like to use it in planting because if you make a mistake, it can cost you a fortune. If your application isn't set correctly you can pick up big losses.

Me talking…

Q3: Do you see any business value in investing in ICT and Digitalization?

Mainly it's to reduce costs and with reducing costs it goes with increasing production. I can give you a long list but that's about it.

Me talking…

Q4: Before making a crop-based decision, do you refer to any source of knowledge?

It's a difficult question to answer so when we decide to plant, we have rotational crop and a rotational system and that we use. Then I adapt it to what the market tells me like at the moment with the war in Ukraine we are now planting more wheat. With the Ukraine war there is going to be a shortage of wheat and fertilizer price has increased. This has nothing to do with technology it is just good business practice. The big technology we use is the Internet to buy machinery from Canada, USA, and the UK. If

you look at the local suppliers and compare the prices on the internet to Canada, it's a lot cheaper on the internet. I get my crop prices from the internet, or I phone someone, I'm not interested searching the internet I'm lazy. I get emails that tell me what the world prices are in different countries.

Me talking… set up next question by referring to drones

Q5: Are there any factors that are restricting you from involving ICT and Digitalization when planting crops?

Have you heard of the saying Curiosity kills the cat? You mustn't jump into anything too fast. Now we are doing line planting and using less fertilizer and less seed before that we used to do board-based planting which is basically spreading out the seed and covering it up. I was one of the first people to try the no till method but failed. You must hold back, wait a bit, and get more information.

Farmer E

Q1: What is your farming background or experience?

I actually came from SA breweries. I worked as an automatic control technician for the past 30 or 40 years. My last job before I started farming was at SA breweries in Newlands where I was in charge of the utilities plant. What I learnt from SAB is if you can't measure it, you can't have it and then I started farming and I'm still a small-time farmer compared to a lot of big farmers in the area. I've spent the last 5 or 6 years learning as much as possible about farming.

Me talking…

Q2: What kind of technology do you use to plant crops?

At the moment I am in a 50/50 partnership with other farmers so basically, I do not own the machinery to plant the crops, but I get different farmers in the area to do the planting for me. Most of the farmers in this area have only got GPS and some have a little bit more advanced where they can actually measure the amount of seed or fertilizer being distributed from the planter. I think that's all they use in this area.

Me talking…

Q3: Do you see any business value in investing in ICT and Digitalization?

I'm not going to mention names alright but If I look at there's four or five farmers in the area. The one guy, Ill call him Farmer X, he basically has got GPS, satellite control, works out exactly how many tons per hectare and distributes it evenly and he can measure how much seed goes per planting area and basically, he measures everything that goes into the soil so that there is no bare areas and makes sure its evenly planted. If we look at another guy in the area, we will call him farmer Y and he basically just plants the old method where you put the seed and fertilizers in the planter either you have a laborer at the back to see if the seed goes out otherwise, they plant and three four weeks later they realize that that area doesn't have fertilizer or seed. If we look at the difference between the two and have a look at the yields. The yields on farmer X's place are a lot better than those of farmer Y. It has increased their yields plus you do not need as much labour force because the old technology uses three or four laborers to see if the seed or

fertilizer is coming out of the planter plus, they have laborers to show them were to plant with the new technology its just one driver and it is done automatically.

Me talking…

Q4: Before making a crop-based decision, do you refer to any source of knowledge?

If we go back, I did automation control which my job is to basically make people redundant because if you can automate it, you don't need the people. There is less chance of something going wrong so basically at the end of the day your profit should go up.

Me talking…

Q.4.1: How does this source improve your decision making?

First of all, before planting you need to have rain, so you look at your weather forecasts either by Ur or whatever weather station you are using for the next month or two. The next thing you do is soil analysis to check what the soil needs before you plant e.g., lime or other phosphates. Then you decide if the soil is correct for the type of crop you want to plant or if you need to put in more nutrients into the soil before you plant it. The internet is basically used to get your weather forecast and also sometimes to research on seeds. Your seeds change or improve all the time so you can always find out if there is a better seed you should be planting and also technology changes with fertilizers. You get organic and normal fertilizers, and you can take a look which seed is better suited for your ground

Me talking… set up next question by referring to drones

Q5: Are there any factors that are restricting you from involving ICT and Digitalization when planting crops?

Speaking to the farmers in the area not myself personally but a lot of them don't understand the technology and they are scared to use the new technology. Also, there problem is that once something goes wrong with the technology or program and they don't actually quite understand it. They don't know how to fix it and there is no one readily available to help them so if they have to wait a week or two weeks to help them by that time the planting season is over, so I think the main reason is basically a lack of education and skills to have backup and teach the people.

Appendix B: Content Analysis

Farmer A

I was born on the farm, grew up on the farm I was running school from the farm into town each and everything day on my bicycle, so I think that's a very good learning experience that you get there from living on a farm and growing up on a farm. Officially I've got no what you call it after school education or diploma nothing about that. that's what I see I learn from, my father my brother and by reading. Internet is very interesting and very good learning curve nowadays if you understand it, you will learn a lot from it. I'm on the farm now for 14 years back after 27 years but still a learning curve each and every day if I don't know you go and check on the Internet and you get good the learning curve.

Me talking…

Q2: What kind of technology do you use to plant crops?

Yeah, we are using GPS not directly for planting but for spraying because the planter has wheels you can work off. The sprayer is 10 m that side and 10m to the other side, so you don't know where the line is so that's where we use the GPS. You need to check weather forecasts like the last week two days we had wind where we can't do any spraying , Thursday we could have done a little bit of spraying and Friday nothing, this morning again so yeah weather forecasts is something you need to be able to have access to and not one application but two or three, that one says so and that one says so , you can check on the average.

Me talking…

Q2.1: Do you use any form of ICT or Digitalization when planting crops?

100% saving in fertilizer saving in the spraying poison and all these things, Input costs getting lower, you not spraying double sewing fertilizers double. A meter wide doesn't sound very much but at the end of the day if you are doing 100 hectors and over spaying by a meter or two it's a lot of money.

Me talking…

Q3: Do you see any business value in investing in ICT and Digitalization?

Definitely because at the end of the day, you not saving money on one side you are also gaining money because if you spray correctly, you don't over spray, you don't get dead areas in your land, so everything is picking up at the end of the day and the value is going back into your pocket

Me talking…

Q4: Before making a crop-based decision, do you refer to any source of knowledge?

We always spoke to the buyer at the end of the day. If KaapAgri this year wants just oats, we will go 90% to the oats are, if they just want wheat, we will go to the wheat area. Each and every year they have got an information session and letters that they send out that says wheat prices this year is very good, oats prices is not that good, go that route or that route. We are making us of the, we call it "landbou kinderness" and they are helping us in the right directions. At the end of the day it's your own decision but they can only inform you of what the trends are

Me talking…

Q.4.1: How does this source improve your decision making?

It will guide you in a direction that maybe you thought you must do x but they are tell you they are doing Y and Z, you can incorporate all three of them and maybe get a good result but it all depends on the weather. It's a guidance or a guideline they are giving.

Me talking… set up next question by referring to drones

Q5: Are there any factors that are restricting you from involving ICT and Digitalization when planting crops?

Drones' money wise, it can be very expensive. Especially in SA where it's not widely used yet as far as I know. Some guys in Limpopo with citrus have crop observation and by Citrusdal. The wheat is bringing in 10k a hector, citrus or peaches is bringing in 100k a hector, the machine costs you 250k. Is it worth to have a 250k machine on a 100 hector of wheat but on 100 hectors of citrus it's worth it. The income from the product plays a big role if you have a big long pocket it will work. Your farming area must be big enough

to get a new fancy planter with all these new sensors on and all those kind of things and it costs you 2.5 mil and you only planted 100 hectares that sum is not going to work out in the end of the day. The size of your farming must be in scale with your input costs.

Me talking…

Farmer B

Q1: What is your farming background or experience?

I grew up in the Tulbagh area and went to Agricultural school. I have been involved in fruit farming, cattle, sheep, goats and planting different types of crops. I've been farming for 40 years.

Me talking…

Q2: What kind of technology do you use to plant crops?

At this stage we are only using GPS and the GPS calculates how many hectares I have planted.

Me talking…

Q2.1

The more effective and precise you can plant in the long run it will be cost saving because you put in the correct amount of seed and fertilizer per hectare. Better technology you use, the less input costs you have. Because input costs have risen so much, you must be more precise using more or less depending on the type of soil.

Me talking…

Q3: Do you see any business value in investing in ICT and Digitalization?

Yes, because the input costs have risen so high. You need to try and reduce your input costs because you can only get X amount per hectre. With the technology available you must try to use it as best as possible to cut costs.

Me talking…

Q4: Before making a crop-based decision, do you refer to any source of knowledge?

We use the internet for weather predictions and for the cost of wheat. The consultants tell us if the world price is going to go up or down. They will tell you if there is a shortage in the world market but in SA there is always a shortage so you must plant. I use the internet to find the costs of fertilizer locally vs importing. I use Yr and long Yr to get

weather forecasts for the next three months to determine if you must plant earlier or later

Me talking…

Q.4.1: How does this source improve your decision making?

The more information you have from nature the better decision you can make, and all the information is not always correct especially the weather forecasts. You can also see if your fertilizer and poison costs are going up or down. If all the prices go up and the wheat price is low, then you are not going to plant wheat and look at alternative crops like canola. The decision is normally between canola or wheat.

Me talking… set up next question by referring to drones

Q5: Are there any factors that are restricting you from involving ICT and Digitalization when planting crops?

Firstly, it's difficult to get the new technology in. Where do you look for it? Its not always on a shelf that you can just go and pick it up. You have to look for it and to find it is always difficult. On the other hand, one of the big problems if you have got that equipment and it breaks there is very few people who can help you but like with the planter's seed feeding system breaks there are too few people who can fix it. There is a big shortage on manpower (Technical power) to help farmers. It can take up to two to three weeks before someone can come and repair it. The new equipment the people come and show you to use it one day and if you forget something in two to three weeks' time or if something goes wrong with the program or you put in the incorrect settings, to try and get someone to help you is just about impossible. To get the guy back to your farm it costs a lot of money and there are too few of them to help.

Farmer C

Q1: What is your farming background or experience?

I grew up on a farm. Through primary school and high school, I helped out on the farm. I studied for 4 years at university, not sure if I can say the university, then afterwards I worked in America for two years farming and this is my first-year farming full time in South Africa in the western cape. I did a BSC in agriculture degree in soil science.

Me talking…

Q2: What kind of technology do you use to plant crops?

We put GPSs in the tractor because it brought up productivity because you don't skip an area or double plant an area which keeps our costs low and increases efficiency. What we changed this year, normally we plant the field and afterwards we do the perimeter. This year we planted the perimeter first and then the middle section. We saved time and less double planting, and it brought down our costs. We changed the planter to it was a traditional mechanical driven planter we changed that to hydraulics so that we can apply variable rates if we wish to, or we can do a standard application.

Me talking…

Q 2.1

If you look at the big guys like in Canada, America , Australia they got very big grain carts, they can run for instance, I'm not sure about the values but for like 150 hectares they can run consistently without filling up your productivity is so much than filling up every 20 or 30 mins , having technology like blockage monitors like we have on your planter so that you can just check everything from your cab that everything is running perfectly. You can't see what is happening behind you so you have to trust that the system is working correctly but if you have blockage monitors or systems in place that doesn't need you to climb out of the cab to check if everything is alright that can improve efficiency by a lot. The longer you can sit in the cab the more efficient you are.

Me talking…

Q3: Do you see any business value in investing in ICT and Digitalization?

Yes, I think business wise like with the fertilizer prices that are sky high at the moment. There is no space for mistakes like the hydraulic are so precise that very seldom something goes wrong and results in a big problem. If you plant incorrectly for a day, you create big problems and then you have to replant that whole area again which results in double costs. In business the lower your input costs the better your profit margin is. I get paid to run the tractor, the tractor hasn't got a union, its got diesel costs, wear and tear and all those factors results in your costs going up. Your costs go up because you have to fix your mistakes. You are not going to actually save costs by planting with variable rates because your soil differs so much, but you are definitely going to see better harvesting planting with variable rates which results in more profit.

Me talking…

Q4: Before making a crop-based decision, do you refer to any source of knowledge?

In this situation now, well my dad is 62 years old, he said that there was trial and error, that's the best thing and how you learn. You have made your mistakes in your past but now in this moment if you get the recipe right or let's say for instance you have more of an idea of how the recipe works you stick with what you got. We sometimes go on the internet and study some stuff if we have problems like our chemical rep can't help us with or we consult our neighboring farmers for extra help but like if we fertilizer a field mostly what we look at weather wise, how much rain we have got this season. That's the most important thing, how much rainfall we have gotten and that will determine how much fertilizer you are going to put in.

Me talking…

Q.4.1: How does this source improve your decision making?

Before we plant, we will check what seed variety we have planted last year, and we change what seed variety we plant this year on that specific field. There is only a limited amount of varieties we can use in this area. It helps us to decide what type of seed to plant in the area and how much fertilizer to use.

Me talking… set up next question by referring to drones

Q5: Are there any factors that are restricting you from involving ICT and Digitalization when planting crops?

We don't have much experience with Drones, I don't know anyone in the area with drones. Someone that sticks with traditional farming like technology is there to make your life easier for instance if drones are more abundant, we would definitely invest in them. If someone says we can do spot spraying with drones, I wouldn't think twice. The reason we don't use it is because we don't know about it, or we don't know about anyone using it at the moment. In our business we can't financially wise jump into something big and change our planting method, we do it step by step. Someone my age I'm 25 years old and I'm more familiar with technology so its easier for me to figure it out. Its easier for me to use the technology like I know how a cellphone and computer works but an older farmer that does not know about technology will not use it but younger farmers like me are more willing to. I would use the technology by trial and error and not change the whole farming practice at once and I would use it on a portion of the farm and if you see better results for one season then you change the whole farming method. As long as the technology is giving you a better profit, I would change the whole farming method.

Farmer D

Q1: What is your farming background or experience?

I have a bachelor's degree in agriculture and an honors degree in agronomy. I started farming in 1984 and I have 40 years farming experience.

Me talking…

Q2: What kind of technology do you use to plant crops?

For the last ten years, we have used GPS auto track. You don't want to know about harvesting and other stuff. What we do is yield mapping and a result of that we put in fertilizer and lime. About five years ago I tried variable rate planting, but the machine wouldn't do it, so I imported it from Canada, but we had problems with that machine then last year we fixed out machine to be able to do variable planting with sensor control. We are very skeptical about the variable rates when it comes to planting, but we are happy with it when it comes to lime spreading. We don't like to use it in planting because if you make a mistake, it can cost you a fortune. If your application isn't set correctly you can pick up big losses.

Me talking…

Q3: Do you see any business value in investing in ICT and Digitalization?

Mainly it's to reduce costs and with reducing costs it goes with increasing production. I can give you a long list but that's about it.

Me talking…

Q4: Before making a crop-based decision, do you refer to any source of knowledge?

It's a difficult question to answer so when we decide to plant, we have rotational crop and a rotational system and that we use. Then I adapt it to what the market tells me like at the moment with the war in Ukraine we are now planting more wheat. With the Ukraine war there is going to be a shortage of wheat and fertilizer price has increased. This has nothing to do with technology it is just good business practice. The big technology we use is the Internet to buy machinery from Canada, USA, and the UK. If

you look at the local suppliers and compare the prices on the internet to Canada, it's a lot cheaper on the internet. I get my crop prices from the internet, or I phone someone, I'm not interested searching the internet I'm lazy. I get emails that tell me what the world prices are in different countries.

Me talking…

Q.4.1: How does this source improve your decision making?

Me talking… set up next question by referring to drones

Q5: Are there any factors that are restricting you from involving ICT and Digitalization when planting crops?

Have you heard of the saying Curiosity kills the cat? You mustn't jump into anything too fast. Now we are doing line planting and using less fertilizer and less seed before that we used to do board-based planting which is basically spreading out the seed and covering it up. I was one of the first people to try the no till method but failed. You must hold back, wait a bit, and get more information.

Farmer E

Q1: What is your farming background or experience?

I actually came from SA breweries. I worked as an automatic control technician for the past 30 or 40 years. My last job before I started farming was at SA breweries in Newlands where I was in charge of the utilities plant. What I learnt from SAB is if you can't measure it, you can't have it and then I started farming and I'm still a small-time farmer compared to a lot of big farmers in the area. I've spent the last 5 or 6 years learning as much as possible about farming.

Me talking…

Q2: What kind of technology do you use to plant crops?

At the moment I am in a 50/50 partnership with other farmers so basically, I do not own the machinery to plant the crops, but I get different farmers in the area to do the planting for me. Most of the farmers in this area have only got GPS and some have a little bit more advanced where they can actually measure the amount of seed or fertilizer being distributed from the planter. I think that's all they use in this area.

Me talking…

Q3: Do you see any business value in investing in ICT and Digitalization?

I'm not going to mention names alright but If I look at there's four or five farmers in the area. The one guy, I'll call him Farmer X, he basically has got GPS, satellite control, works out exactly how many tons per hectare and distributes it evenly and he can measure how much seed goes per planting area and basically, he measures everything that goes into the soil so that there is no bare areas and makes sure its evenly planted. If we look at another guy in the area, we will call him farmer Y and he basically just plants the old method where you put the seed and fertilizers in the planter either you have a laborer at the back to see if the seed goes out otherwise, they plant and three four weeks later they realize that that area doesn't have fertilizer or seed. If we look at the difference between the two and have a look at the yields. The yields on farmer X's place are a lot better than those of farmer Y. It has increased their yields plus you do not need as much labour force because the old technology uses three or four laborers to see if the seed or

fertilizer is coming out of the planter plus, they have laborers to show them were to plant with the new technology it's just one driver and it is done automatically.

Me talking…

Q4: Before making a crop-based decision, do you refer to any source of knowledge?

First of all, before planting you need to have rain, so you look at your weather forecasts either by Ur or whatever weather station you are using for the next month or two. The next thing you do is soil analysis to check what the soil needs before you plant e.g., lime or other phosphates. Then you decide if the soil is correct for the type of crop you want to plant or if you need to put in more nutrients into the soil before you plant it. The internet is basically used to get your weather forecast and also sometimes to research on seeds. Your seeds change or improve all the time so you can always find out if there is a better seed you should be planting and also technology changes with fertilizers. You get organic and normal fertilizers, and you can take a look which seed is better suited for your ground

Me talking…

Q.4.1: How does this source improve your decision making?

If we go back, I did automation control which my job is to basically make people redundant because if you can automate it, you don't need the people. There is less chance of something going wrong so basically at the end of the day your profit should go up.

Me talking… set up next question by referring to drones

Q5: Are there any factors that are restricting you from involving ICT and Digitalization when planting crops?

Speaking to the farmers in the area not myself personally but a lot of them don't understand the technology and they are scared to use the new technology. Also, there problem is that once something goes wrong with the technology or program and they don't actually quite understand it. They don't know how to fix it and there is no one readily available to help them so if they have to wait a week or two weeks to help them by that time the planting season is over, so I think the main reason is basically a lack of education and skills to have backup and teach the people.

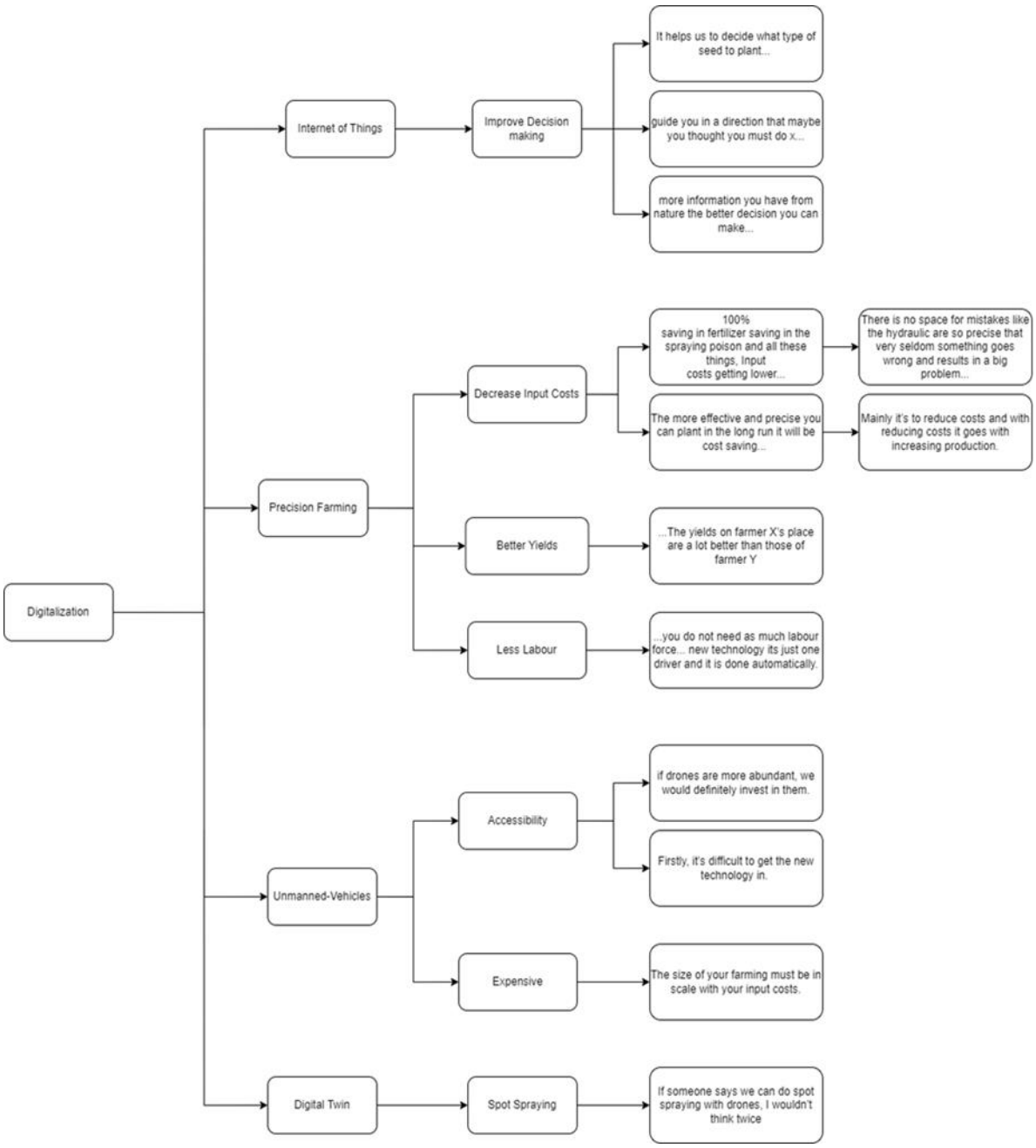


Figure 14 Digitalization Codes Diagram

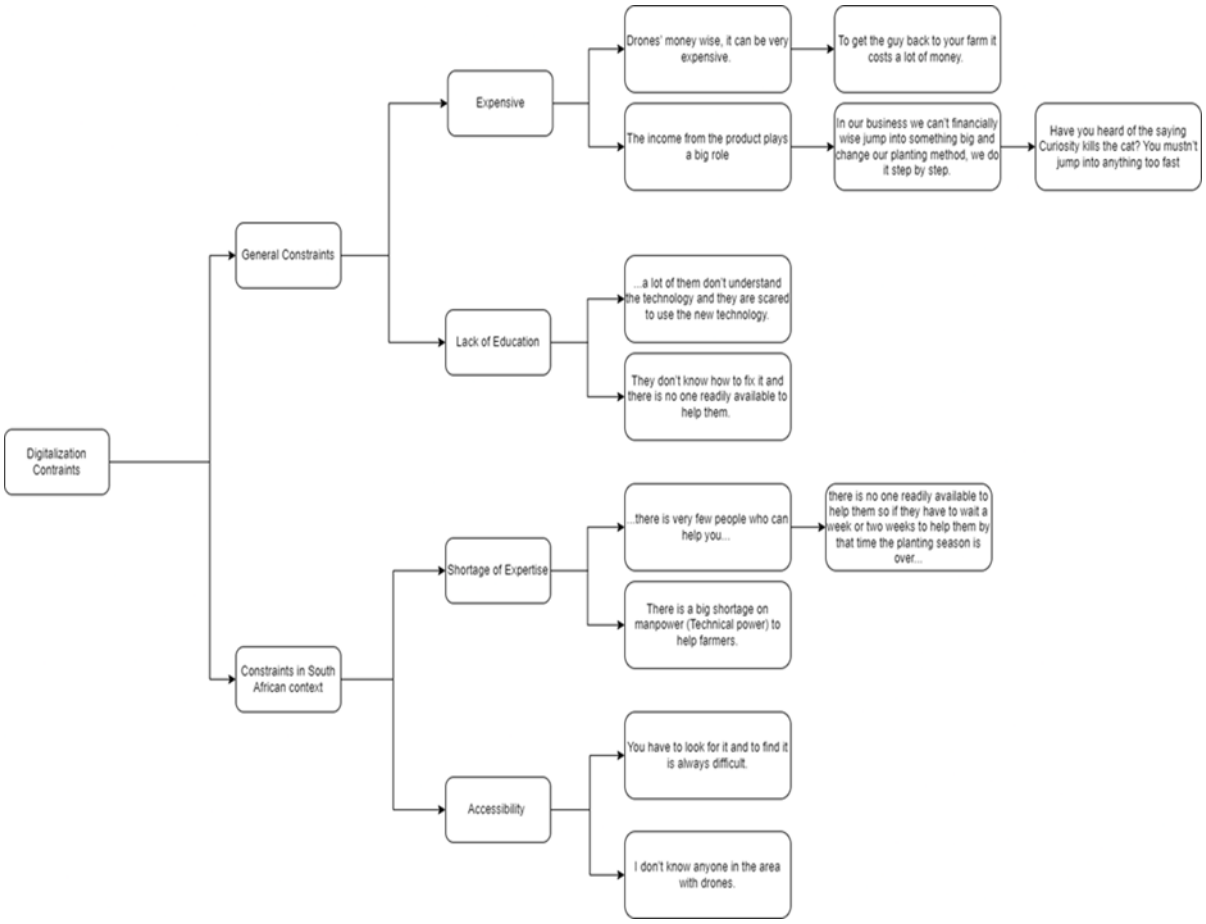
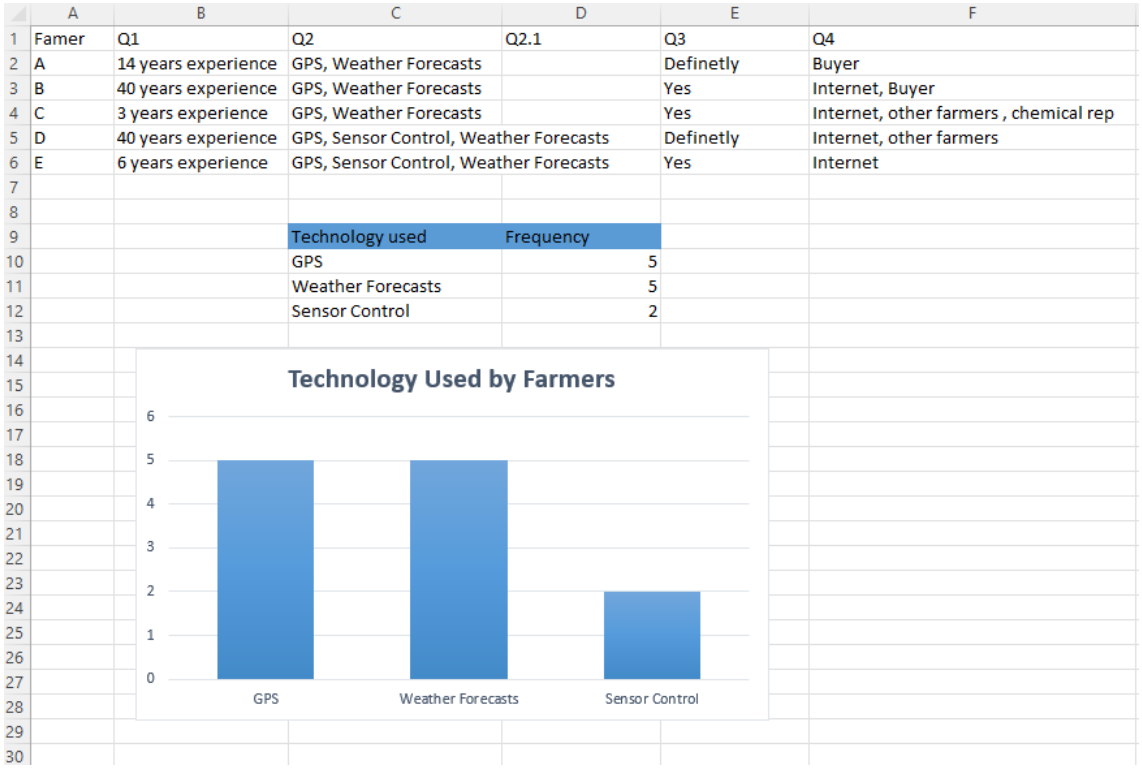


Figure 15 Digitalization Constraints Codes Diagram

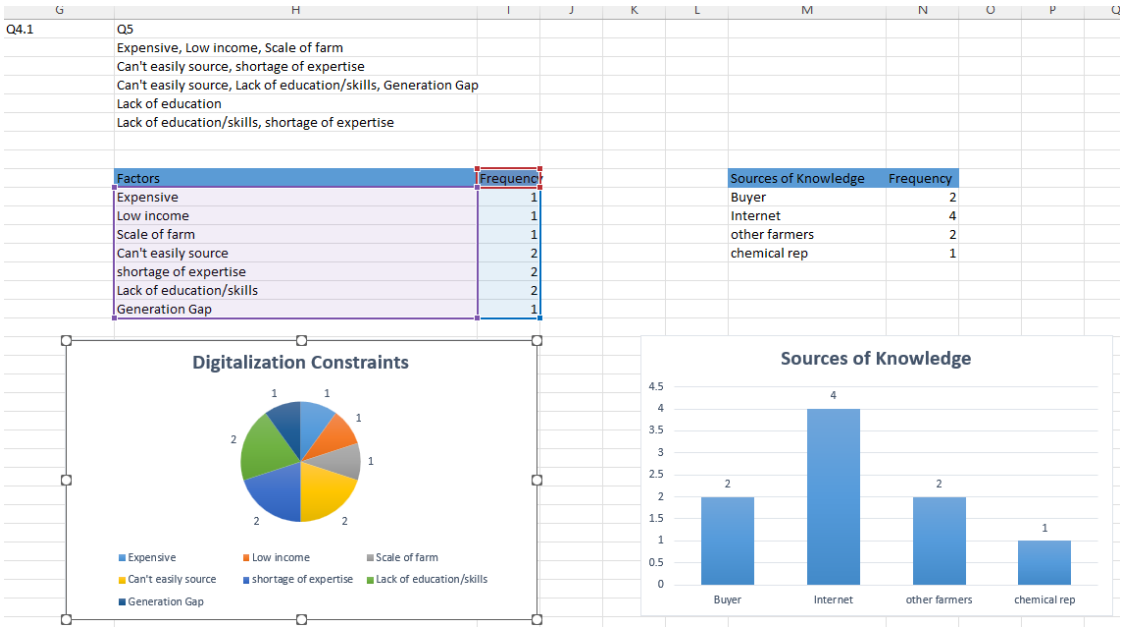
Appendix C: Qualitative Content Analysis using Excel



Famer	Q1	Q2	Q2.1	Q3	Q4
A	14 years experience	GPS, Weather Forecasts		Definetly	Buyer
B	40 years experience	GPS, Weather Forecasts		Yes	Internet, Buyer
C	3 years experience	GPS, Weather Forecasts		Yes	Internet, other farmers , chemical rep
D	40 years experience	GPS, Sensor Control, Weather Forecasts		Definetly	Internet, other farmers
E	6 years experience	GPS, Sensor Control, Weather Forecasts		Yes	Internet

Technology used	Frequency
GPS	5
Weather Forecasts	5
Sensor Control	2

Figure 16 Screenshot of analysis process using Excel



Q4.1	Q5
	Expensive, Low income, Scale of farm
	Can't easily source, shortage of expertise
	Can't easily source, Lack of education/skills, Generation Gap
	Lack of education
	Lack of education/skills, shortage of expertise

Factors	Frequency
Expensive	1
Low income	1
Scale of farm	1
Can't easily source	2
shortage of expertise	2
Lack of education/skills	2
Generation Gap	1

Sources of Knowledge	Frequency
Buyer	2
Internet	4
other farmers	2
chemical rep	1

Figure 17 Screenshot 2 of analysis process using Excel

Appendix D: Python Code

```
# Pipeline methods for cleaning the data and feature engineering

# Specified stop words to be removed from the data
# Certain words may change the meaning of the next word therefore the stop words are specified

stop_words= ['yourselves', 'between', 'whom', 'itself', 'is', "she's", 'up', 'herself', 'here', 'your', 'each',
             'we', 'he', 'my', "you've", 'having', 'in', 'both', 'for', 'themselves', 'are', 'them', 'other',
             'and', 'an', 'during', 'their', 'can', 'yourself', 'she', 'until', 'so', 'these', 'ours', 'above',
             'what', 'while', 'have', 're', 'more', 'only', "needn't", 'when', 'just', 'that', 'were', "don't",
             'very', 'should', 'any', 'y', 'isn', 'who',  'a', 'they', 'to', 'too', "should've", 'has', 'before',
             'into', 'yours', "it's", 'do', 'against', 'on',  'now', 'her', 've', 'd', 'by', 'am', 'from',
             'about', 'further', "that'll", "you'd", 'you', 'as', 'how', 'been', 'the', 'or', 'doing', 'such',
             'his', 'himself', 'ourselves',  'was', 'through', 'out', 'below', 'own', 'myself', 'theirs',
             'me', 'why', 'once',  'him', 'than', 'be', 'most', "you'll", 'same', 'some', 'with', 'few', 'it',
             'at', 'after', 'its', 'which', 'there','our', 'this', 'hers', 'being', 'did', 'of', 'had', 'under',
             'over','again', 'where', 'those', 'then', "you're", 'i', 'because', 'does', 'all']

review_features = ['review_text']
```

Figure 18 Screenshot of python code

```
# Function to remove punctuation
class RemovePunctuation(BaseEstimator, TransformerMixin):
    def fit(self, X, y=None):
        return self

    def transform(self, X):
        # remove punctuation
        def clean_review(text):
            text = str(text).lower()
            text = text.replace("’","'")
            text = re.sub('\[.*?\]', '', text)
            text = re.sub('https?://\S+|www\.\S+', '', text)
            text = re.sub('<.*?>+', '', text)
            text = re.sub('[%s]' % re.escape(string.punctuation), '', text)
            text = re.sub('\n', '', text)
            text = re.sub('\w*\d\w*', '', text)
            return text

        X['Q1'] = X['Q1'].apply(lambda x: clean_review(x))
        X['Q2'] = X['Q2'].apply(lambda x: clean_review(x))
        X['Q2_1'] = X['Q2_1'].apply(lambda x: clean_review(x))
        X['Q3'] = X['Q3'].apply(lambda x: clean_review(x))
        X['Q4'] = X['Q4'].apply(lambda x: clean_review(x))
        X['Q4_1'] = X['Q4_1'].apply(lambda x: clean_review(x))
        X['Q5'] = X['Q5'].apply(lambda x: clean_review(x))
        return X

# Function to remove stop words
class StopWords (BaseEstimator, TransformerMixin):
    def fit(self, X, y=None):
        return self

    def transform(self, X):
        # remove stop words
        X['Q1'] = X['Q1'].apply(lambda x: ' '.join([word for word in x.split() if word not in (stop_words)]))
        X['Q2'] = X['Q2'].apply(lambda x: ' '.join([word for word in x.split() if word not in (stop_words)]))
        X['Q2_1'] = X['Q2_1'].apply(lambda x: ' '.join([word for word in x.split() if word not in (stop_words)]))
        X['Q3'] = X['Q3'].apply(lambda x: ' '.join([word for word in x.split() if word not in (stop_words)]))
        X['Q4'] = X['Q4'].apply(lambda x: ' '.join([word for word in x.split() if word not in (stop_words)]))
        X['Q4_1'] = X['Q4_1'].apply(lambda x: ' '.join([word for word in x.split() if word not in (stop_words)]))
        X['Q5'] = X['Q5'].apply(lambda x: ' '.join([word for word in x.split() if word not in (stop_words)]))
        return X

✓ 0.3s
```

Figure 19 Screenshot 2 of python code

```
# Calls the pipeline methods
pipe = Pipeline([
    ('removepunctuation', RemovePunctuation()),
    ('stopword', StopWords())

])

# Apply the pipeline to the dataframe
data = pipe.fit_transform(data)
data.head()
```

[8] ✓ 0.3s

	Farmer	Q1	Q2
0	A	born farm grew farm running school farm town e...	yeah using gps not directly planting but spray...
1	B	grew tulbagh area went agricultural school inv...	stage using gps gps calculates many hectares p...
2	C	grew farm primary school high school helped fa...	put gpss tractor brought productivity dont ski...
3	D	bachelors degree agriculture honors degree agr...	last ten years used gps auto track dont want k...
4	E	actually came sa breweries worked automatic co...	moment partnership farmers basically not machi...

Figure 20: Screenshot 3 of python code

```
# method generates a wordcloud for positive reviews
text = Question2
wordcloud = WordCloud(
    width = 3000,
    height = 2000,
    background_color='white',
    stopwords = stop_words).generate(str(text))

fig = plt.figure(figsize = (16, 10))
plt.imshow(wordcloud, interpolation = 'bilinear')
plt.axis('off')
plt.title('Question 2', fontsize = 20)
plt.rcParams['figure.facecolor'] = 'white'
plt.tight_layout(pad=0)
plt.show()
```

[ ] ✓ 2.9s

Figure 21 Screenshot 4 of python code

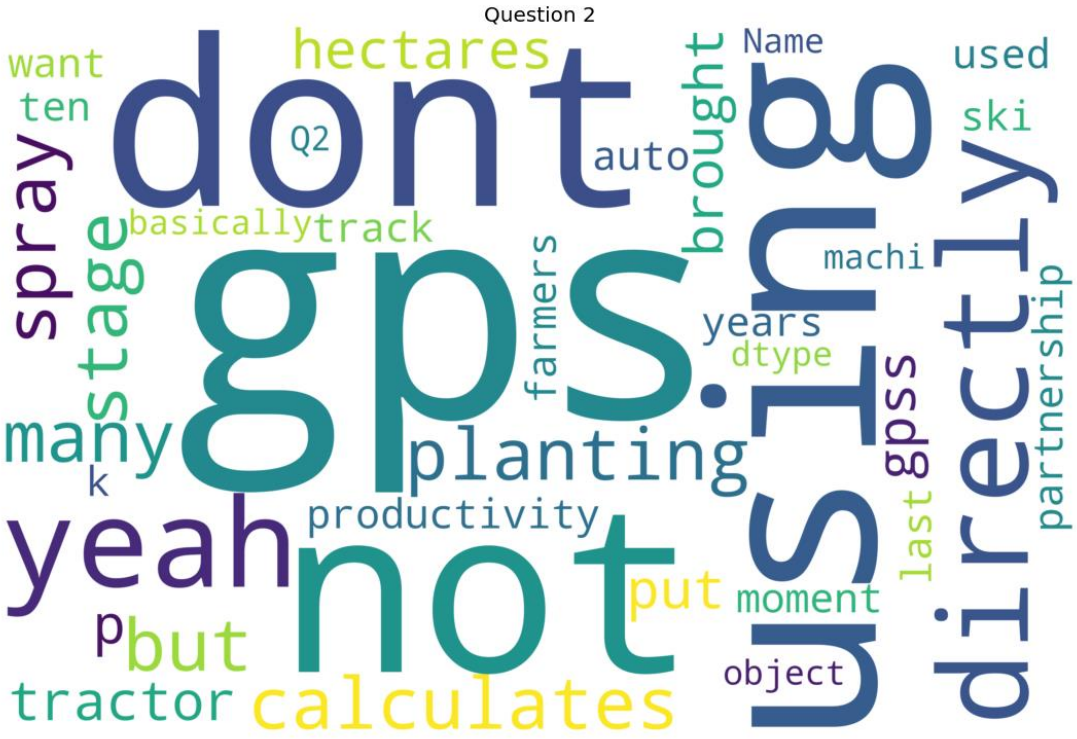


Figure 22 Screenshot of Question 2 Wordcloud

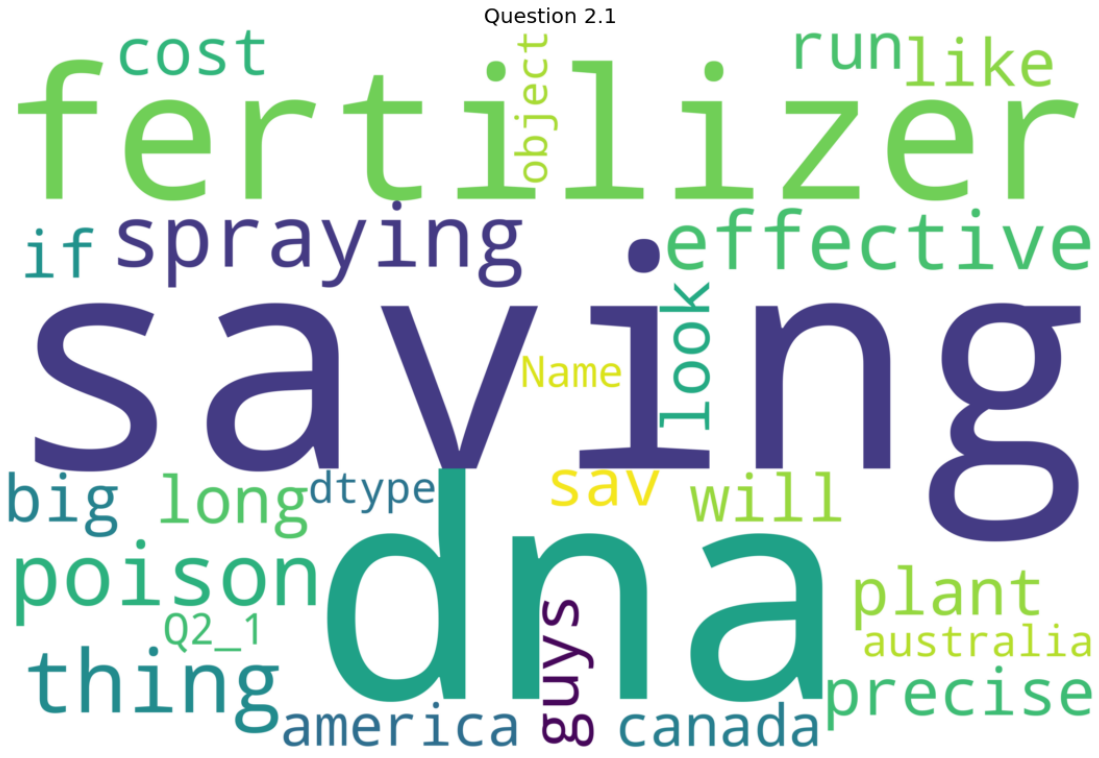


Figure 23 Screenshot of Question 2.1 Wordcloud

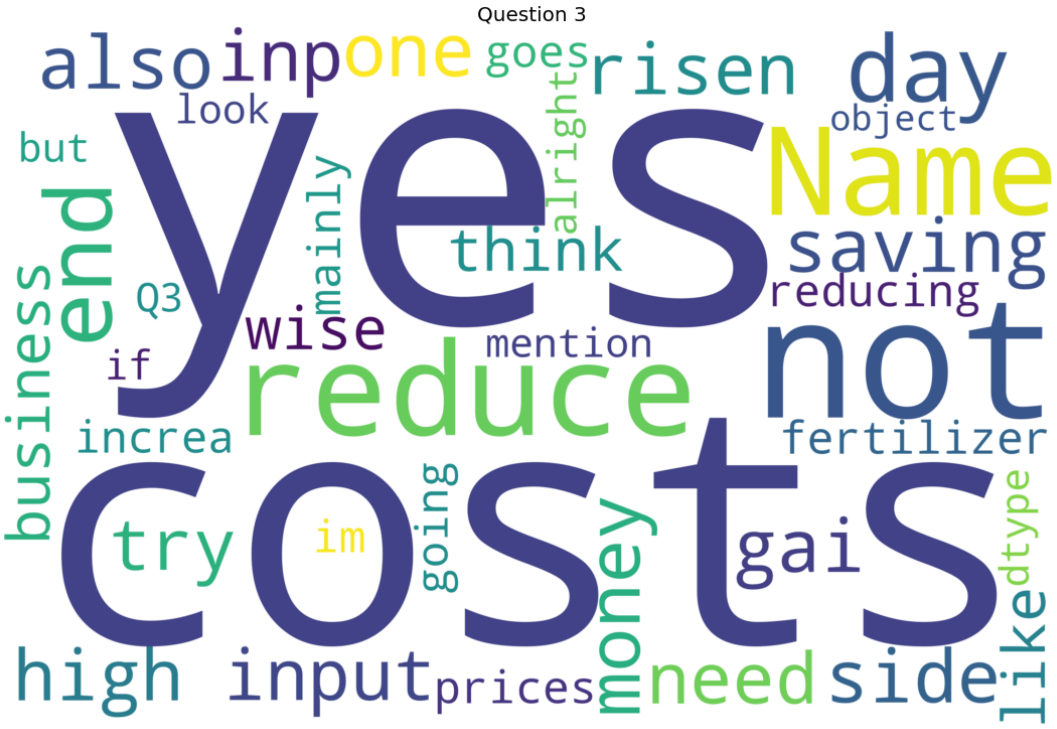


Figure 24 Screenshot of Question 3 Wordcloud

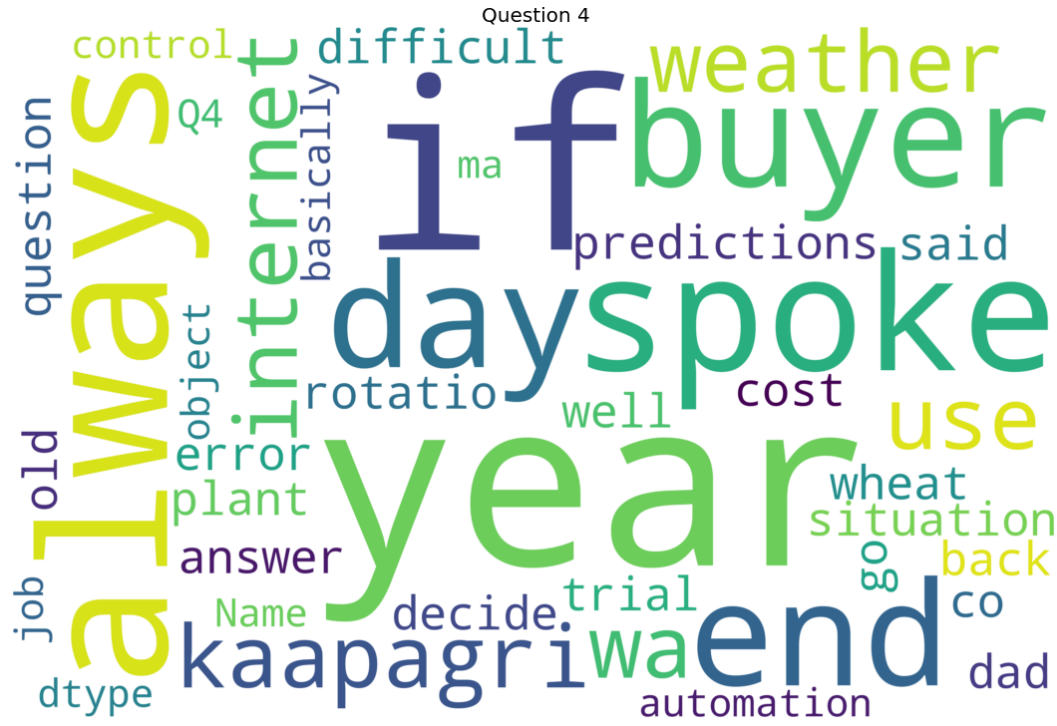


Figure 25 Screenshot of Question 4 Wordcloud

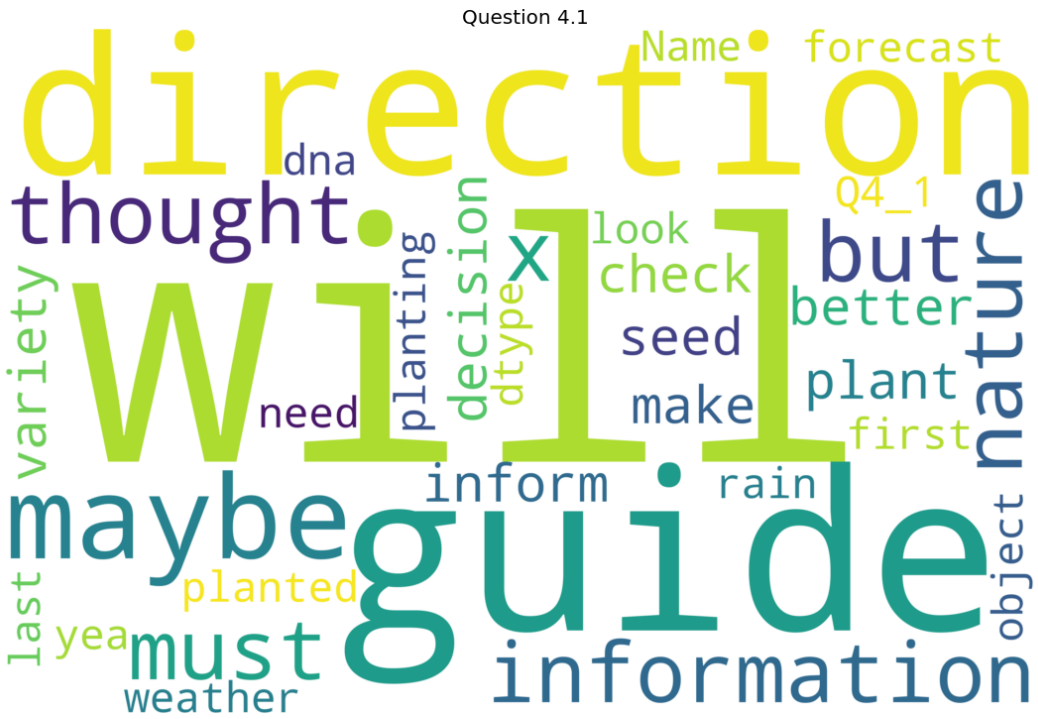


Figure 26 Screenshot of Question 4.1 Wordcloud

Figure 27 Screenshot of Question 5 Wordcloud

```
## custom function for ngram generation
def generate_ngrams(text, n_gram = 1):
    token = [token for token in text.lower().split(" ") if token != "" if token not in stop_words]
    ngrams = zip(*[token[i:] for i in range(n_gram)])
    return [" ".join(ngram) for ngram in ngrams]
```

✓ 0.3s

```
# custom function for horizontal bar chart
def horizontal_bar_chart(df, color):
    trace = go.Bar(
        y =df["word"].values[::-1],
        x = df["wordcount"].values[::-1],
        showlegend = False,
        orientation = 'h',
        marker = dict(
            color = color,
        ),
    )
    return trace
```

✓ 0.3s

Figure 28 Screenshot 5 of python code

```
## Get the bar chart from Question 2
freq_dict = defaultdict(int)
for sent in Question2:
    for word in generate_ngrams(sent, 2):
        freq_dict[word] += 1
fd_sorted = pd.DataFrame(sorted(freq_dict.items(), key=lambda x: x[1])[::-1])
fd_sorted.columns = ["word", "wordcount"]
trace2 = horizontal_bar_chart(fd_sorted.head(20), 'yellow')

## Get the bar chart from Question 3
freq_dict = defaultdict(int)
for sent in Question3:
    for word in generate_ngrams(sent, 2):
        freq_dict[word] += 1
fd_sorted = pd.DataFrame(sorted(freq_dict.items(), key=lambda x: x[1])[::-1])
fd_sorted.columns = ["word", "wordcount"]
trace3 = horizontal_bar_chart(fd_sorted.head(20), 'green')

## Get the bar chart from Question 4
freq_dict = defaultdict(int)
for sent in Question4:
    for word in generate_ngrams(sent, 2):
        freq_dict[word] += 1
fd_sorted = pd.DataFrame(sorted(freq_dict.items(), key=lambda x: x[1])[::-1])
fd_sorted.columns = ["word", "wordcount"]
trace4 = horizontal_bar_chart(fd_sorted.head(20), 'orange')

## Get the bar chart from Question 5
freq_dict = defaultdict(int)
for sent in Question5:
    for word in generate_ngrams(sent, 2):
        freq_dict[word] += 1
fd_sorted = pd.DataFrame(sorted(freq_dict.items(), key=lambda x: x[1])[::-1])
fd_sorted.columns = ["word", "wordcount"]
trace5 = horizontal_bar_chart(fd_sorted.head(20), 'brown')
```

20] ✓ 0.7s

Figure 29 Screenshot 6 of python code

```
# Creating two subplots
fig = tools.make_subplots(rows=4, cols=1, vertical_spacing = 0.03, subplot_titles=["Frequent words of Question 2",
                                                                                   "Frequent words of Question 3","Frequent words of Question 4",
                                                                                   "Frequent words of Question 5",])

fig.append_trace(trace2, 1, 1)
fig.append_trace(trace3, 2, 1)
fig.append_trace(trace4, 3, 1)
fig.append_trace(trace5, 4, 1)



fig['layout'].update(height=1600, width=1080, paper_bgcolor='rgb(255,255,255)', title="Word Count Plots")
iplot(fig, filename='word-plots')
```

✓ 2.2s

Figure 30 Screenshot 7 of python code

Figure 31 Screenshot of Question 2 & 3 N-gram

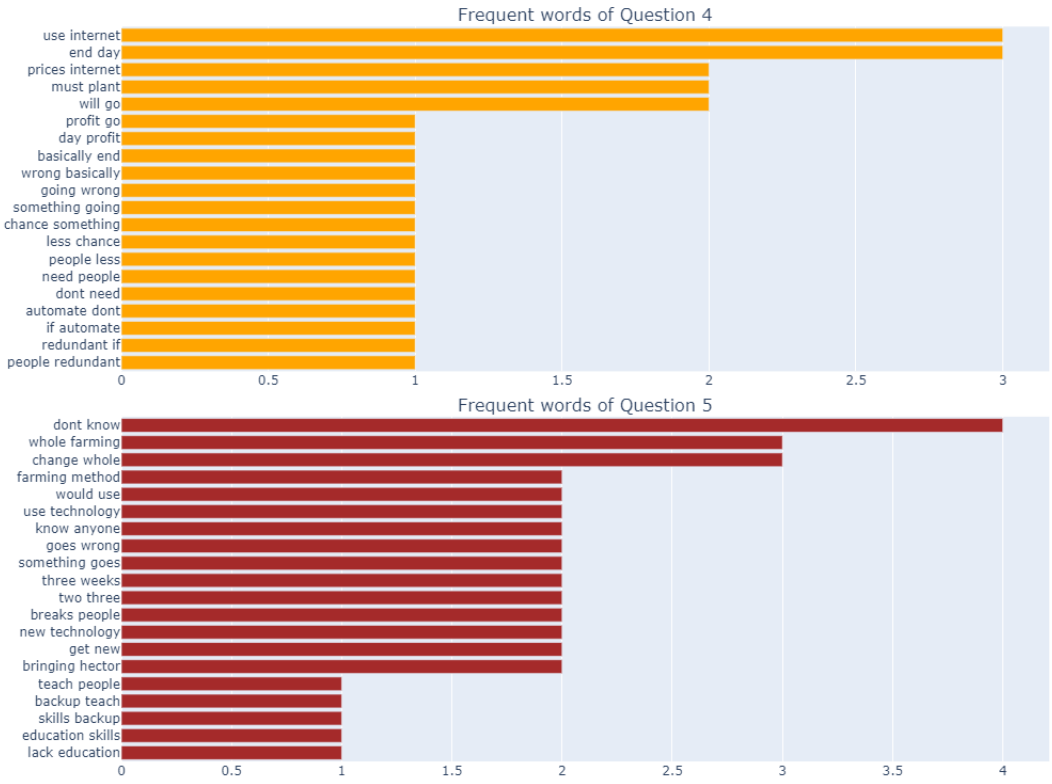


Figure 32 Screenshot of Question 4 & 5 N-gram