

UNIVERSITY OF BAHRAIN

College of Science



**Environment and Sustainable
Development Program**

AGROBAHRAIN: TOWARDS SMART FARMING IN THE BIG DATA ERA

A Thesis Submitted in Partial Fulfilment of the Requirements for the
Doctor of Philosophy in Environment and Sustainable Development

Prepared by

Mohamed Essam Hussain Ghanim
ID: 19982899

Supervisor

Dr. Jaflah Al Ammary
(Associate Professor)
University of Bahrain

Kingdom of Bahrain

May 2024

Defense Committee's Approval

University of Bahrain
Deanship of Graduate Studies
And scientific Research



جامعة البحرين
عمادة الدراسات العليا والبحث العلمي

Thesis Discussion Committee Decision

Student Name: Mohammad Essam Hussien Ghanim	Student ID: 19982899
Department: Choose an item.	College: Science
Program: PhD in Environment and Sustainable Development	
Thesis Title: AGROBAHRAIN : Towards Smart Farming in the Big Data Era	
Supervisor Name: Dr. Jafiah Al Ammari	Student GPA: 3.81

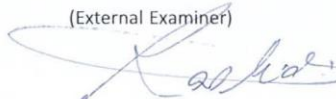
The Discussion Committee met on 27-Jun-24 In the College of Science – Room S41-0002. Upon discussing the student whose details are stated above in public, the discussion committee recommends the following:

The student is awarded a grade A in the thesis subject.

- ☐ Without making any changes to the thesis.
- ☒ After making the modifications attached with this decision.

Discussion Committee Members Approval:

Name	Signature
1. Dr. Jafiah Al Ammari (Advisor)	Jafiah Alammary
2. Dr. Haroun Abdulla Alryalat (Internal Examiner)	HAROUN ALRYALAT
3. Prof. Abdul Rashid Bin Mohamed Shariff here to enter text.	(External Examiner) Click


12/9/24

1 Approved at the University Council 6th meeting, 2013
In accordance with Decision No. 736/2013 dated 14/3/2013
Modified by University Council Decision No. 2043/2013 dated 20/11/2013.

Abstract

The agriculture sector is on the edge of a revolutionary era driven by advancing technologies. These technologies are revolutionizing agriculture, providing unprecedented opportunities for enhanced efficiency, sustainability, and resilience. To successfully integrate these technologies, governments, researchers, and farmers must work together to overcome obstacles and ensure a fair, ethical, and inclusive change in global agriculture.

The Kingdom of Bahrain, a small island in the Arabian Gulf, faces intrinsic agricultural challenges such as arid soil and climatic conditions and a lack of water supply. Furthermore, the agriculture industry in Bahrain has yet to embrace the opportunities presented by the digital revolution ultimately. On the other hand, due to the lack of available explicit agricultural data, Bahraini farmers rely heavily on traditional methods and existing knowledge.

This study introduced a framework for national agriculture Big Data in Bahrain with four main objectives. Firstly, it focused on exploring the level of awareness and digital readiness among farmers in the Kingdom of Bahrain. This investigation is crucial for enhancing the farmers' digital capacity and skills. Secondly, the study investigated agriculture data needs supply and demand dynamics. It involved identifying available data sources, understanding the specific requirements of each stakeholder, and addressing the challenges associated with acquiring agricultural data. Additionally, the study emphasized the importance of such data to establish a user/owner structure for managing agriculture Big Data in the Kingdom of Bahrain.

Building upon the insights gained from the first two objectives, the third objective was to develop a conceptual framework for agriculture Big Data in Bahrain. Integrating Big Data tools and cloud services facilitate multi-access levels of the stored data for different stakeholders. Lastly, the fourth objective involved the development of a simulation model based on the framework developed in the previous objective. This model assesses the feasibility and simulate different processes and functionalities of a smart agriculture platform.

Given the multifaceted nature of the research, a diverse set of research approaches were implemented. The methodologies encompassed a closed-ended questionnaire distributed to over 100 farmers and farm tenants. On-site farm visits were conducted to observe and understand the available farming operations and how they ran. Semi-

structured interviews were conducted with different agriculture stakeholders, and an extensive review of pertinent literature was conducted, all framed within a conceptual framework. The tools and technological resources utilized included wireless soil sensors, a multispectral camera, cloud services (Google Cloud Platform), a remote sensing platform, and different other tools as the primary development environment. The study revealed notable obstacles encountered by agriculture data applications in the Kingdom of Bahrain, such as the lack of sufficient techniques and tools for data gathering, inadequate data access interfaces, and restrictions on data sharing due to different limitations. Moreover, there is a lack in Big Data frameworks that can efficiently manage Bahrain's agricultural environment. However, the findings demonstrated a conspicuous and evident enthusiasm for agricultural technology among individuals; nonetheless, agriculture stakeholders showed a limited understanding and adoption of Information and Communication Technology tools to run their daily farming operations.

Based on the approaches conducted to explore the agriculture environment in Bahrain, a framework was built to cater specifically to the distinctive attributes while considering obstacles such as small-scale agricultural practices and farmers' low proficiency in Information and Communication Technology. The framework consisted of two stages: an initial stage focused on enhancing Information and Communication Technology capabilities among agriculture stakeholders and a subsequent step demonstrating the semi-autonomous gathering, storage, and analysis of agriculture data. Finally, a simulation model was developed in the cloud using Google Cloud services, incorporating customized Application Programming Interfaces, a Progressive Web App, and smart agricultural equipment and technologies.

Achieving agriculture sustainability in Bahrain necessitates an action plan with collaboration and support from agriculture stakeholders and policymakers. Initially, this collaboration should raise awareness and usage towards adopting Information and Communication Technology in agriculture and transform the awareness into actual use. To encourage more usage, Information and Communication Technology services for agricultural production should be developed to be accessible to all users, considering all age groups. Moreover, Information and Communication Technology for agriculture diffusion using emergent technologies such as Internet of Things and Big Data is a priority. This initiative marked the first step in building a conceptual

process for collecting, storing, and transforming agriculture Big Data in Bahrain. In line with the agricultural policy formulated in the early 2010s, this study set the stage for furthering research in several areas relevant to agriculture.

Publications

- M. E. Ghanim and J. Alammmary, "AgroBahrain: A Conceptual Framework for Agriculture Big Data for Bahrain," 2021 International Conference on Electronic Information Technology and Smart Agriculture (ICEITSA), Huaihua, China, 2021, pp. 545-550, doi: 10.1109/ICEITSA54226.2021.00108.
- Ghanim, M., Alammmary, J. (2023). Cloud-Based Simulation Model for Agriculture Big Data in the Kingdom of Bahrain. In: Yang, XS., Sherratt, R.S., Dey, N., Joshi, A. (eds) Proceedings of Eighth International Congress on Information and Communication Technology. ICICT 2023. Lecture Notes in Networks and Systems, vol 693. Springer, Singapore. https://doi.org/10.1007/978-981-99-3243-6_59.
- Al-Ammmary, J.H. and Ghanem, M.E. (2023), "Information and communication technology in agriculture: awareness, readiness and adoption in the Kingdom of Bahrain", Arab Gulf Journal of Scientific Research. <https://doi.org/10.1108/AGJSR-07-2022-0113>

Dedication

To the solid pillars of my life—my loving mother, my late father, and my beloved family,

In sincere gratitude and admiration for the unlimited love, enduring encouragement, and unwavering support that have lit every step of my arduous academic journey. To my mother, whose unrelenting strength and dedication have served as my guiding light, and to my late father, whose memory continues to inspire and echo inside the corridors of my heart.

This PhD journey was more than just a solitary goal; it was a collaborative effort fuelled by the collective warmth of my family. To my siblings, wife, and children, whose laughter cut through the difficulties.

With immense love and gratitude

Acknowledgements

I would like to express my sincere gratitude to the National Initiative for Agriculture Development (NIAD) for their invaluable support throughout my doctoral journey. The material and moral support provided by NIAD have played a crucial role in the successful completion of my PhD.

I am also deeply grateful to my supervisor, Dr. Jaflah, for her unwavering guidance, mentorship, and dedication. Her expertise, encouragement, and constructive feedback have been instrumental in shaping the course of my research and academic development.

Finally, I extend my heartfelt thanks to the interviewees along the farms' tenants for their genuine material and moral support, without which this accomplishment would not have been possible.

Table of contents

<i>Defense Committee's Approval</i>	II
<i>Abstract</i>	III
<i>Publications</i>	VI
<i>Dedication</i>	VII
<i>Acknowledgements</i>	VIII
<i>Table of contents</i>	IX
<i>List of Tables</i>	XII
<i>List of Figures</i>	XIV
<i>List of Abbreviations</i>	XVI
<i>List of Appendices</i>	XVIII
Chapter One	19
1.1. Overview	19
1.2. Background	19
1.3. Research Problem	23
1.4. Research Questions	26
1.5. Objectives	27
1.6. The Model of the Study	28
1.7. Thesis Outline	30
1.8. Summary	30
Chapter Two	32
2.1. Overview	32
2.2. ICT in Agriculture	32
2.2.1. ICT Literacy and Readiness	35
2.2.2. ICT and Extension Services in Agriculture	38
2.2.3. Smart Agriculture	40
2.2.4. Precision Agriculture	43
2.2.5. ICT in Agriculture and Sustainability	44
2.3. Big Data	47
2.3.1. Big Data Formats	50
2.3.2. Big Data Storage and Processing	51
2.4. Big Data in Agriculture	57

2.4.1. Challenges and Opportunities	65
2.5. Agriculture in Bahrain	67
2.6 Summary	74
Chapter Three	75
3.1. Overview	75
3.2. Research Methods	75
3.2. Phase I: ICT Awareness, Readiness, & Adoption	77
3.3. Phase II: Agriculture Data Lifecycle	80
3.4. Phase III: Conceptual Framework	82
3.5. Phase IV: The Simulation Model	84
3.5.1. AgroBahrain Progressive Web App (PWA)	87
3.5.2. Pycno Soil Sensor	88
3.5.3. Satellite Remote Sensing	89
3.5.4. Multi Spectral Camera	91
3.5.5. Google Cloud Platform (GCP)	92
3.5.6. Implementation Environment	93
3.6. Summary	95
Chapter Four	96
4.1. Overview	96
4.2. ICT Awareness, Readiness, & Adoption	96
4.3. Agriculture Data Lifecycle	106
4.4. SWOT-TOWS Analysis for strategic options for adopting ICT in agriculture in Kingdom of	
4.5. Summary	129
Chapter Five	130
5.1. Overview	130
5.2. The Framework	131
5.2.1. Preparation Stage	133
5.2.2. Data Lifecycle: Overview	138
5.2.3. Implementation Scheme	142
5.3. Model Simulation	153
5.3.1. Registering New Stakeholders	159
5.3.2. Farm owner features	161
5.3.3. Indirect Data Sources	164
5.3.4. Crop Production	169

5.3.5. Agriculture Enquiries Service (AES)	173
5.3.6. Web Scrapping	176
5.3.7. Data Extraction and Presentation	177
5.4. Summary	182
Chapter Six	184
6.1. Conclusion	184
6.2. Contribution	188
6.3. Recommendations	191
6.4. Future Work	195
6.5. Limitations	196
References	199
Appendix A	212
Appendix B	220
Appendix C	223
الملخص	228

List of Tables

Table (2.1): SDGs Interlinked to Agriculture.	45
Table (2.2): Datawarehouse Vs. Data Lake.	56
Table (2.3): Cloud Storage Services.	57
Table (2.4): Agricultural productivity.....	68
Table (3.1): Tools and Technologies Used for the Simulation.....	85
Table (3.2): Technical Details – Pycno Soil Sensor	89
Table (3.3): Technical Details – EOS Crop Monitoring Platform.....	90
Table (3.4): Technical Details - Parrot SEQUOIA+ Multispectral camera.....	91
Table (3.5): GCP services utilized for the simulation.....	92
Table (3.6): Technical Details – Development Environment	94
Table (4.1): Description of the farmers and agriculture business	97
Table (4.2): Digital readiness of the farmers	98
Table (4.3): Usage of ICT in Agriculture by farmers.	100
Table (4.4): Perception of the difficulties in adopting ICT in agriculture.	102
Table (4.5): The impact of adopting ICT in agriculture.	102
Table (4.6): Field visits analysis	104
Table (4.7): Interviews and the Interviewees details	107
Table (4.8): Interview Questions (Availability of data).....	109
Table (4.9): Interview Questions (ICTs used to exchange agriculture data between stakeholders).	111
Table (4.10): Interview Questions (Collection of Agriculture Data).....	115
Table (4.11): Interview Questions (Storage, Access, and Challenges).....	119
Table (4.12): Agriculture data owners/users’ matrix for the Kingdom of Bahrain. ..	123
Table (4.13): SWOT analysis for the readiness for adopting ICT in agriculture in Kingdom of Bahrain	126
Table (4.14): TOWS matrix for the strategic directions for adopting ICT in agriculture.	128
Table (5.1): Agriculture data categories and sample sources	142
Table (5.2): Data used for the simulation model.	155
Table (6.1): sample workshops and training programs.....	194
Table (B.1): Interview Card (questions and interview meta data).....	220
Table (C.1): List of Cloud Functions and their features.	223

Table (C.2): Cloud Function “agro_init_auto” used to initialise automated sources.	224
Table (C.3): Cloud Function “agro_auto_extract” used to automatically collect data every 6 hours from farms with active sensors.	225

List of Figures

Figure (1.1): The Model of the Study	29
Figure (2.1): The Big Data characteristics (7 Vs).....	49
Figure (3.1): Agriculture Stakeholders	81
Figure (3.2): The conceptual agriculture Big Data management approach.	87
Figure (4.1): Digital readiness of the farmers.....	98
Figure (4.2): Awareness of farmers on the ICT used in agriculture.	99
Figure (5.1): The AgroBahrain Framework.....	132
Figure (5.2): The PWA Main Page	143
Figure (5.3): Agriculture stakeholder registration scheme	145
Figure (5.4): The Pycno soil sensor positioned in the selected farm.	156
Figure (5.5): Multispectral images of tomato. Bands from left (RGB, Green, Red, Red-edge, Near Infrared).	156
Figure (5.6): sample of data collected. On the left sample JSON data from Pycno soil sensor, and on the right sample JSON of EOS satellite data.	157
Figure (5.7): Satellite remote sensing images captured through the EOS crop monitoring platform. Left image represents the NDVI index, and the right image represents the RECI index.	158
Figure (5.8): Adding new stakeholder data scheme.....	159
Figure (5.9): Adding new stakeholder pipeline using GCP tools	160
Figure (5.10): New user registration form	161
Figure (5.11): Farm Owner Interlinked Operations	162
Figure (5.12): Simulated farm owner features. Emphasized farms features to include adding and updating crop production.	163
Figure (5.13): Adding Farms and crops.....	164
Figure (5.14): Sample of rejected/accepted CSV file.	165
Figure (5.15): The user interface for uploading data files manually.	165
Figure (5.16): Uploading data files manually pipeline using GCP tools.....	166
Figure (5.17): Sample storage bucket output.....	166
Figure (5.18): The add data file (manual upload) data scheme.	167
Figure (5.19): Adding automated data sources - interfaces sequence.	167
Figure (5.20): Adding crop production form.	170
Figure (5.21): Crops main page. Active production and completed.....	171

Figure (5.22): Basic data visualization for active production.....	171
Figure (5.23): Crop production and activating automated data source data scheme .	172
Figure (5.24): Pipelines for crop production with active data sources.	173
Figure (5.25): Enquiries main page (list of enquiries).	174
Figure (5.26): Add new enquiry form.....	175
Figure (5.27): Sample of stored enquiry as a NoSQL document on Google Firestore (key value pairs).....	176
Figure (5.28): The target webpage used in web scrapping.	176
Figure (5.29): Agriculture data interface, displays statistical data along the ability to extract custom data sets.	178
Figure (5.30): The queries used to extract the data based on the provided filters.	180
Figure (5.31): The data entities required for autonomus sources.	180
Figure (5.32): Retrieved data frame prior transformation.	181
Figure (5.33): Transformed data set.....	181
Figure (6.1): A conceptual cooperation map to improve ICT readiness and adoption.	195

List of Abbreviations

AES	Agriculture Enquiries Service
AGU	Arabian Gulf University
AI	Artificial Intelligence
API	Application Programing Interface
AWS	Amazon Web Services
CGIAR	The Consultative Group on International Agricultural Research
CSV	Comma-Separated Values
ELT	Extract, Load, Transform
EOS	Earth Observation System
ETL	Extract, Transform, Load
FAO	Food and Agriculture Organization
GCF	Google Cloud Function
GCP	Google Cloud Platform
GDP	Gross Domestic Product
GFS	Google File System
GHGs	Greenhouse Gases
GIS	Geographic Information System
GNSS	Global Navigation Satellite System
GPS	The Global Positioning System
HDFS	Hadoop Distributed File System
ICT	Information and Communications Technology
ICARDA	International Centre for Agriculture Research in the Dry Areas
IGA	Information and eGovernment Authority
IoT	Internet of Things
ITU	International Telecommunications Union
JSON	JavaScript Object Notation
MDGs	Millennium Development Goals
MMAA	Ministry of Municipalities Affairs and Agriculture

NGOs	Non-Governmental Organizations
NIAD	National Initiative for Agricultural Development
NoSQL	Not only SQL
NSSA	National Space Science Agency
PA	Precision Agriculture
PATs	Precision Agriculture Technologies
PWA	Progressive Web App
SDGs	Sustainable Development Goals
SIDS	Small Islands Development States
SQL	Structured Query Language
UAV	Unmanned Ariel Vehicle
UN	United Nations
URL	Uniform Resource Locator
WSNs	Wireless Sensor Networks
XML	Extensible Markup Language

List of Appendices

- Appendix A Presents the Google form's version of the online questionnaire
- Appendix B The semi-structured interview questions' card template.
- Appendix C List of Google Cloud Functions and sample code of 'agro_init_auto'
and 'agro_auto_extract' functions.

Chapter One

Introduction

1.1. Overview

The first chapter will introduce the research topic by establishing a brief background about ICT in agriculture and its present status. Then, it will shed light on the challenges in the Kingdom of Bahrain that led to the development of the research concept. The rest of the chapter will establish the basis of the research through a set of questions and objectives, the study model, and the final section will introduce concisely the following chapters.

1.2. Background

The exponential proliferation of information and communications technology (ICT) in recent decades has brought about a profound transformation of the world into an information society. Since the advent of ICT infrastructure, such as the Internet and mobile phones, governments, businesses, and individuals have significantly improved their access to information, expertise, and knowledge (Bahrini & Qaffas, 2019). The adoption of ICT has substantially increased the efficiency of resource allocation, decreased production costs by a significant margin, and stimulated demand and growth in every sector of the economy (Kostoska & Kocarev, 2019).

For example, the widespread adoption of Information and Communication Technology (ICT) platforms orchestrating transactions within the sharing economy has induced a profound transformation in stakeholder interactions in recent years (Gössling & Hall, 2019). These platforms have ascended as globally influential intermediaries in their

own right, attributed to their extensive market reach, significant control over transactional volumes, and intricate affiliations with other transnational corporations and capital.

Moreover, Information and Communication Technology (ICT) has played a pivotal role in enhancing educational and learning approaches. Unwin (2019) articulated that ICT served as the catalyst for establishing the Virtual University for Small States for the Commonwealth in 2000.

On the other hand, ICT is one of the biggest drivers of agricultural innovation. This phenomenon is propelled by a swift surge in the utilization of Big Data to enhance prevailing smart farming technologies, including automated tractors. Additionally, there is a growing prevalence of mobile applications and software designed to integrate data related to agricultural production processes, such as input quantity and timing, with information pertaining to on-farm work procedures and quality management (Knierim et al., 2019). Agriculture experts believe that smart farming technologies can significantly improve agricultural sustainability and economic growth by increasing the precision of inputs according to site-specific requirements and connecting these elements to farm management platforms (Annosi, Brunetta, Capo, & Heideveld, 2020).

The agricultural revolution is believed to have begun around 10000 BC, as people harnessed their strength, as well as that of animals, and utilized basic tools to cultivate land and crops. The second industrial revolution in the 19th century marked the advent of Agriculture 2.0. This era was marked by the utilization of machinery and chemical inputs in substitution of human labor to enhance agricultural output. During the latter part of the 20th century, there was a rise in the adoption of ICT across several industries. This led to the emergence of the third agricultural revolution, known as Agriculture 3.0. The proliferation of data gathered from diverse sources, coupled with the advent of

technologies like Big Data, Cloud Computing, and Artificial Intelligence (AI), has given rise to the approaching agricultural revolution known as Agriculture 4.0 (Zhai, Martínez, Beltran, & Martínez, 2020).

Even though the agriculture industry is considered the major food supplier, it is also responsible for many challenges economically and environmentally. The agriculture sector faces three main challenges: insufficient production, crop quality, and the environmental impacts and the exhaustion of natural resources (Trivelli et al., 2019), such as global deforestation, soil degradation, and the excessive consumption of water resources (Clercq, Vats, & Biel, 2018). Cooperation between policymakers and agriculture stakeholders and the integration of agriculture technologies are required to overcome such challenges.

By 2050, the world population will reach approximately 9.1 billion, and the demand for food will increase by 70%, which increases the pressure on food production (FAO, How to Feed the World in 2050, 2009a). The global food crisis that occurred in 2008 triggered severe concerns regarding food security in most regions globally (Independent Evaluation Group, 2013). Despite governments' efforts to reduce food insecurity, this challenge still affects many people worldwide. The Food and Agriculture Organization (FAO) reported in 2021 that between 702 and 828 million people worldwide were suffering from hunger (FAO, IFAD, UNICEF, WFP, & WHO, 2022). This rise in food shortage clearly warns countries to take considerable action towards the achievement of the first Sustainable Development Goal (SDG 1), which aims to 'end poverty in all its forms everywhere', and the second goal (SDG 2), which aims to 'end hunger, achieve food security, improve nutrition, and promote sustainable agriculture', by 2030 (Gil et al., 2019). According to Murray W. Scown (2020), to achieve the SDGs, food production must be harmonized with environmental, socioeconomic, and rural development objectives, and

based on the United Nations, the means to achieve SDGs are science and technology, knowledge sharing, and capacity building (Mohieldin, 2018).

Agriculture technologies work as an enabler with the potential to positively affect the three pillars of sustainability. **Socially**, by providing higher accessibility to information and better data security and privacy; economically, by minimizing process costs and improving productivity and quality; and **environmentally**, by minimizing resource consumption (Annosi et al., 2020).

Agriculture technologies can be roughly divided into ICTs presented by hardware/software and biological-related technologies such as (genetic modifications and biotic and abiotic factors). ICTs in agriculture were considered immature and were recognized internationally in 2003 during the World Summit on the Information Society (WSIS) through the term e-Agriculture (FAO, 2017). ICTs include a wide range of technologies, starting from the radio and television to broadcast agriculture extension programs, computers and the Internet, smartphones, and recently reached advances in technologies where the collection of real-time data remotely, analysis of the data to produce meaningful insights for more precise decisions (Ali et al., 2019). Their effects differ significantly depending on the technology employed and the stakeholders' literacy level.

Digital technologies like AI, robotics, and the Internet of Things (IoT) are anticipated to be game changers. Real-time, in-depth data analysis is possible because of digital technologies. Technologies such as smart sensors, Unmanned Aerial Vehicles (UAVs), and remote sensing are all used to gather data continuously. AI algorithms analyze this large volume of data (Big Data) and provide information. Therefore, farmers may use both their in-field expertise and data analysis to inform their decisions (Bertoglio, Corbo, Renga, & Matteucci, 2021).

Big Data is a concept that permeates many economic sectors in the contemporary technology-based era, and it is essential to build an automatic method to extract operational insights from bulk data due to the increasing volume of data accessible for field management (Saiz-Rubio & Rovira-Más, 2020). As per Magnin (2016), Big Data will positively impact all food-related processes (farm to fork). Therefore, Agriculture Big Data frameworks and applications have greatly interested researchers. A significant number of studies discussed applications for crops (Majumdar, Naraseeyappa, & Ankalaki, 2017), weed control (Evert et al., 2017), precision agriculture (Mekonnen, Namuduri, Burton, Sarwat, & Bhansali, 2020), and many other related topics. Most of the frameworks presented focus on analytics, machine learning, and the internet of things; however, few studies focus on agriculture Big Data engineering operations (data collection, processing, and storage), and up to the authors' knowledge, agriculture Big Data engineering implementation on Google Cloud Platform (GCP) has not been discussed by previous research.

1.3. Research Problem

The Kingdom of Bahrain is an archipelago with several small islands connected and suffers from many challenges, including lack of fresh water, desert climate, unsuitable soil for cultivation, and the annual shrinking of the agricultural area (FAO, Irrigation in the Middle East region in figures, 2009b). Furthermore, Bahrain has been classified by the United Nations among the Small Islands Development States (SIDS) (Naser, 2022) due to the similarity of the main challenges and the ability to find solutions to achieve sustainable development goals, especially combating climate change.

The term (SIDS) appeared in 1992 during a scientific conference in the Maldives to discuss sea-level rise (Wong, 2011). Countries in this category face many vulnerabilities, such as limited areas and small populations with high density, which cause increasing pressure on the limited land resources and institutional capacities (Wong, 2011).

Agriculture is the primary source of food, and due to the increase in world population, food security is becoming a very challenging target that requires the creation of sustainable agriculture practices to ensure an increase in production while maintaining world-limited resources. Although the Kingdom of Bahrain was previously famous for cultivating palm trees and producing different types of dates, it suffers from a noticeable absence in the production of many crops (Bani, 2020). On the other hand, farmers rely heavily on managing their farms traditionally and based on their experience (albiladpress, 2020). They did not taste the advantages of using digital tools to provide much better decisions.

Many developing countries, including Bahrain, experience a significant yield shortfall compared to their prospective yield. This gap is due to many reasons, including the lack of the adoption of ICTs in the agriculture sector and poor research capabilities due to data-related challenges. Transforming the agriculture system in Bahrain requires socio-technical reform. This can be achieved through emerging technologies such as Big Data, IoT, and cloud computing (Mekonnen et al., 2020), along with building ICT capacity among different agriculture stakeholders.

The Kingdom of Bahrain is one of the small islands in the Middle East that will benefit from agricultural advancements to mitigate the problematic factors that hinder agricultural production, such as difficult soil conditions, relatively hot weather, and lack of rain. To use digital technologies and implement smart agriculture initiatives in Bahrain, this

study intends to focus on two significant challenges. First, there is a lack of scientific research and studies related to the usage and adoption of digital technologies in the agriculture sector in Bahrain, and this study is considered the first. Furthermore, the lack of such research establishes a substantial knowledge deficit that necessitates initial efforts to investigate and comprehend the state of the sector regarding ICTs in agriculture.

On the other hand, data is a crucial factor in finding solutions to the challenges faced by the agricultural field, and it is necessary first to collect data about the agricultural environment and the factors surrounding it and then analyze this data to achieve optimal solutions and action plans. However, despite the initiatives undertaken by many governmental and non-governmental institutions to advance the agricultural sector in the Kingdom of Bahrain, it severely lacks agriculture data. Data is either scattered, outdated, inaccessible, or mostly not available. Accordingly, there needs to be a framework/mechanism to take advantage of the technological advancements to collect, store, and analyze agrarian data. Furthermore, the available initiatives, such as (AgroBH) project, focus only on structured data, which creates a huge gap in data usability.

As per Wolfert et al. (2017), Big Data applications in cultivation are not only about primary production but also assume a significant job in improving the production network's proficiency and easing food security concerns. A sustainable agriculture system relies heavily on data to create knowledge that leads to effective decisions. Big Data analysis has proven effective in many fields, including agriculture. However, building and maintaining a Big Data infrastructure is costly, time-consuming, and resource consuming. A significant research gap emerges when comparing the utilization of agricultural big data frameworks between agricultural and non-agricultural countries. While numerous frameworks have been

developed to accommodate the needs of agricultural nations, focusing on individual farms and supporting stakeholders proficient in ICT, a notable deficiency exists in addressing the requirements of non-agricultural countries with limited ICT knowledge and adoption. The lack of studies targeting these regions at a national scale underscores a crucial gap in current research efforts, highlighting the pressing need for tailored agricultural big data frameworks that can bridge the technological differences between agricultural and non-agricultural contexts. There is a need for a national smart agriculture ecosystem that allows the smooth collection of data from different sources, where data is processed, stored, and efficiently accessible. The system should be suitable and adaptable for the Bahraini agricultural environment concerning the simplicity of usage, inclusion, and resilient infrastructure.

1.4. Research Questions

The use of digital technologies related to agriculture, such as remote sensing applications, wireless soil sensors, or even the fundamental input of data into an electronic form, demands a certain level of ICT literacy and understanding. On the other hand, the collection of data in agriculture is very challenging due to the heterogeneity of the data and the vast number of sources that are available. This demands the existence of an effective framework that enables the collection, processing, and storage of the data in an efficient manner while also providing access to all of the various stakeholders. The feasibility of implementing an agriculture Big Data framework and a simulation model requires finding an answer to the following questions:

1. To what extent are the agriculture stakeholders in Bahrain digitally ready, aware, and adopt ICTs in the sector?

2. Is there any agriculture data being collected in Bahrain? Who collects it? What type of data is collected? What are the tools and methods used for the collection and storage processes? Who benefits from the collected data, and how is it accessed?
3. What is the best approach to building an agriculture Big Data framework that suits the Kingdom of Bahrain to enable seamless collection and exploitation of agriculture data?
4. How to assess the feasibility of the developed framework?

1.5. Objectives

The term "agriculture" should be defined in a comprehensive manner to encompass all forms of biological production systems, such as arable crops, perennial forages, forestry, orchards, vineyards, horticulture, cattle husbandry, and aquaculture. Nevertheless, this study specifically concentrates on the management and production of crops. The aim of this study is to build a smart agricultural framework utilizing Big Data and cloud computing technologies in order to enhance the sustainability of crop production in the Kingdom of Bahrain. Multiple objectives were established to address the research inquiries and effectively accomplish the primary goal.

1. Exploring the awareness, digital readiness, and adoption of ICT among farmers in the Kingdom of Bahrain. This will be done toward building digital capacity among the farmers.
2. Investigating the supply/demand side of the agriculture data needs by identifying the available data, the different needs of each stakeholder, the difficulties and challenges of acquiring the agriculture data, and the importance of the data. This will be done toward building a user/owner structure for the Kingdom of Bahrain agriculture Big Data.

3. Based on the first two objectives, the third objective is to build a conceptual agriculture Big Data framework for the Kingdom of Bahrain to visualize and identify different approaches and technologies of data collection, transformation, storage, and presentation integrating Big Data tools and cloud services. The developed framework will enable multi-access levels of the stored data for the different stakeholders and must be flexible and resilient for future customization.
4. The fourth objective is to develop a simulation model based on the developed framework to assess its feasibility and imitate a smart agriculture platform's different processes and functionalities.

1.6. The Model of the Study

The current research is based on the flow chart illustrated in Figure 1.1. The study is divided into three interconnected phases according to the model. Assessing the level of ICT awareness and adoption is essential for determining the appropriate scale of the framework design (whether it is at the farm or national level), understanding the complexity of the system and its interface, and identifying the users and their preparedness for the system. Conversely, comprehending the agriculture data lifecycle facilitates the classification of data to respective owners and users among many stakeholders, enables an overview of data requirements, and establishes access levels. The third phase involves constructing the framework by utilizing the fundamental understanding of the agriculture stakeholders and their correlation with the agriculture data. The last stage will replicate the stakeholders' data lifecycle and showcase the many features of the framework.

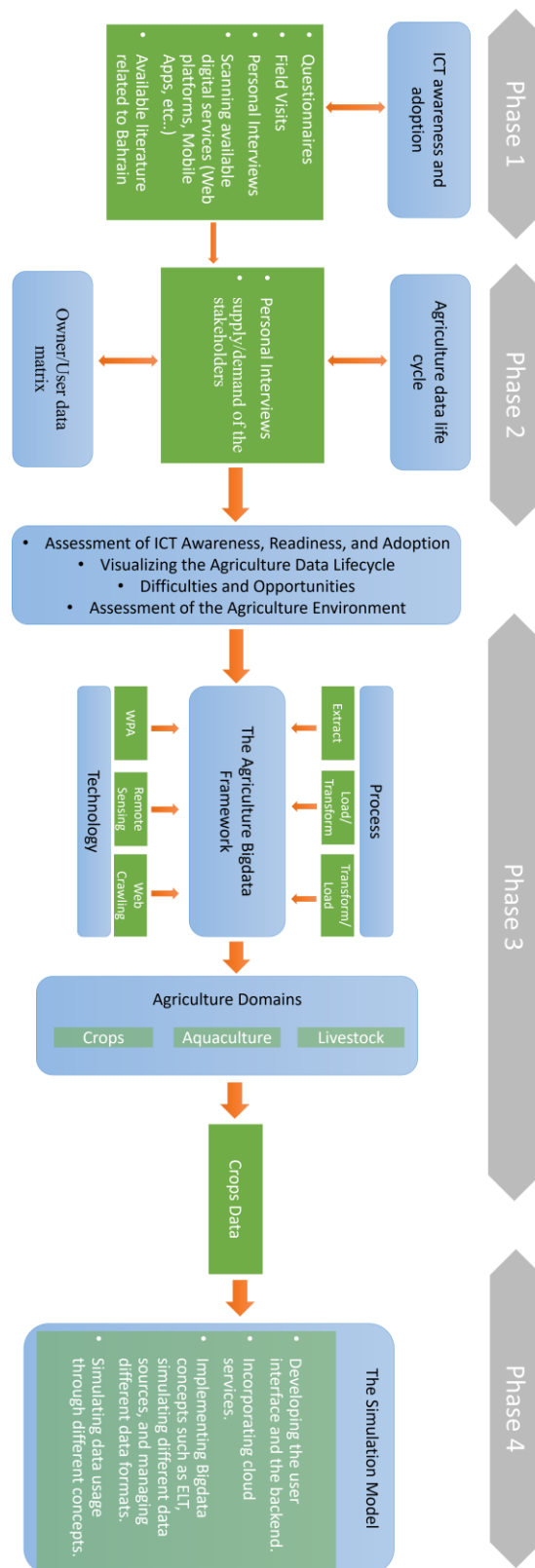


Figure (1.1): The Model of the Study

1.7. Thesis Outline

The study is divided into six chapters. In Chapter 1, an overview of the study and the problem led to the research idea and discussion of the technology's importance in achieving the SDGs. Chapter 2 reviews the importance of ICTs in general and their impact on the agriculture sector and then concentrates on agriculture Big Data implementations and the advantages and significant challenges. Moreover, the chapter will discuss the initiatives taken by Bahrain to improve the agriculture sector as well as the available practices that depend on ICTs. Finally, it will emphasize the role of agriculture sustainability in achieving the SDGs. Chapter 3 discusses the different approaches used to understand the agriculture ecosystem in Bahrain, and then it will explain the methodology used to build the Big Data framework. Finally, the chapter will list all the GCP tools, development languages, and environments used to implement the simulation model. Quantitative and qualitative results, a SWOT/TOWS analysis, and the agriculture data owners/users matrix will be presented and discussed in Chapter 4. The framework design and the simulation model (data schemes, user interface, functionality) will be presented and discussed in Chapter 5. Finally, Chapter 6 presents the concluded remarks, contribution of the study, limitations, and future works.

1.8. Summary

This chapter delivers a comprehensive outline of the thesis subject matter, centring on the integration of ICT in agriculture and the incorporation of emerging technologies, notably Big Data and cloud computing. Additionally, it explores the historical context of technological progress in the agricultural field and traces its evolution over time. This chapter focuses on clarifying the current research gaps and highlighting the issues faced by the agricultural sector in Bahrain.

In addition, the chapter presents the fundamental research questions required to achieve the main goals of the study, along with outlining the core objectives. Finally, the chapter provides a concise summary of the research methods, and the model adopted in this thesis.

Chapter Two

Literature Review

2.1. Overview

This chapter's main objective is to review the literature concerning ICT in agriculture, Big Data in Agriculture, and agriculture in the Kingdom of Bahrain. The literature review would help set a background on recent publications related to ICT in agriculture and identify the direction of the research.

The chapter highlights the influence of ICT on the agriculture sector in developing and developed countries. Furthermore, it depicts some factors that play a role in the success and failure of smart agriculture. On the other hand, this review will discuss Big Data applications in the agriculture sector and how it affects its sustainability.

2.2. ICT in Agriculture

There is no exact definition of Information and Communication Technologies (ICTs); however, it can be explained as data collection, analysis, and transfer through the combination of computing devices, communication capabilities, and software (Bilali & Allahyari, 2018). The acronym ICT or ICTs can be used interchangeably to refer to the utilization of devices and infrastructures to enable the digital exchange of information (Zuppo, 2012). ICT covers a wide range of aspects of our lives, such as education, the economy, governmental and private services, and even our lifestyle. Hence, it is considered as a vital component for the enduring progress of individuals and nations and is acknowledged by global governmental and non-governmental entities such as the World Health Organisation (WHO), the Organisation for Economic Co-operation and Development (OECD), and the European Union (EU)

(Zuppo, 2012). ICT was identified in 2002 by the OECD according to four products-based categories named (Telecommunications, Consumer Electronics, Computers and office machinery, and Other ICT, and their inadequacy to encompass technological trajectories or paradigm shifts). However, this definition is no longer compatible with current technological advancement and diversification. In 2017, ICT was redefined according to technology and not products by classifying 13 technology-related areas, including (High-speed network, Mobile communication, High-speed computing, Sensor and device network, Large-capacity and high-speed storage, and Large-capacity information analysis) (Inaba & Squicciarini, 2017).

Before the emergence of ICT and its usage in agricultural development, people primarily depended on traditional farming methods and would resort to others to provide services. This is called 'Traditional Agriculture' (Singh & Singh, 2017).

Agriculture is one of the oldest crafts in the world. As per Thompson (1968), 1815 was the year when the first revolution matured, and the second evolution was born. This era was distinguished by the use of machines such as (tractors) and chemical inputs instead of human and animal power to increase crop production and profit. In the second half of the 20th century, with the dissemination of ICTs, came the third agriculture revolution, Agriculture 3.0, distinguished by the transition from traditional farming to business-based farming (Bertoglio et al., 2021). Agriculture 3.0 is estimated to have existed ten years after the third industrial revolution, during the period of (1980 – 2010) and was characterized by the new terminology (Smart Agriculture) (Mazzetto, Gallo, & Sacco, 2020). Moreover, despite being a far old concept, Precision Agriculture (PA) improved substantially during this era. The spread of PA techniques has been primarily facilitated by the

technological advancement of positioning systems such as (Global Navigation Satellite System (GNSS) devices) (Saiz-Rubio & Rovira-Más, 2020).

The tremendous increase in data sources and data amount was the main driver of the rise of the next agriculture revolution, Agriculture 4.0 (Mazzetto et al., 2020). This data-driven era is characterized by emerging technologies such as Big Data, Cloud Computing, and AI. It led to a significant leap in the agriculture sector that enables future predictions and precise real-time decisions to be managed intelligently (Zhai et al., 2020).

Agriculture requires the effective utilization of innovation and technology to accelerate production and enhance employment opportunities. Multiple types of innovations are developed to enhance efficiency and productivity in the agriculture industry. According to Parke (2015), the introduction of new ideas and logical approaches has had a significant impact on the agricultural industry in recent years. Milovanović (2014) states that ICT can be effectively employed in the agricultural sector to enhance farming productivity by directly contributing to its efficiency. Furthermore, in 2017, the declaration of the group of twenty (G20) agriculture ministers in Berlin recommended and focused on empowering ICT in agriculture to boost the efficiency and sustainability of the food sector (Bilali & Allahyari, 2018).

In the agricultural context, the potential of ICT can be estimated broadly under two leads: either as a tool for direct contribution to agrarian productivity or as an indirect tool for empowering farmers to make more informed and quality decisions, which will have a positive result on the way agriculture and allied activities are conducted (Patel & Iliyas, 2014). Furthermore, using ICT in different business fields has demonstrated great efficiency in cost and productivity. Therefore, ICT is being adopted

in the agriculture business through concepts such as smart farming and e-agriculture to overcome the agriculture sector's challenges.

ICTs in agriculture were introduced a long time ago in many countries worldwide. They paved the road towards sustainable farming by allowing real-time information communication even for rural areas. They demonstrated the ability to improve crop productivity and quality, minimize resource consumption, and reduce environmental impact. Some innovative ICTs in agriculture are autopilot tractors, crop sensors, variable rate technology (VRT) and Swath Control Technology, Monitoring and controlling crop Irrigation systems, Field documentation, IoT, and Big Data (Ruder, 2019).

The evolution of ICT from a simple radio program to machine learning and advanced analytics of agriculture data demonstrated the power of ICT to agriculture business processes globally.

2.2.1. ICT Literacy and Readiness

ICT has emerged as a game-changer in agriculture, significantly impacting every aspect of farming, from crop management to marketing and distribution. The integration of ICT in agriculture has the potential to enhance productivity, increase efficiency, and promote sustainability in food production (Bilali, Bottalico, Palmisano, & Capone, 2020). However, ICT literacy and readiness among farmers are essential to realize these benefits. ICT literacy refers to the ability of farmers to effectively use and navigate digital tools and communication technologies (Alant & Bakare, 2021). Achieving ICT literacy in agriculture requires training and education initiatives tailored to farmers' specific needs and contexts.

Governments, Non-Governmental Organizations (NGOs), and private organizations should develop and implement training programs that teach

farmers how to use ICT tools effectively. These programs should cover basic digital literacy and specialized skills related to agricultural technology (Aldosari, Shunaifi, Ullah, Muddassir, & Noor, 2019).

Ensuring that farmers have access to affordable smartphones or tablets and reliable internet connectivity is crucial. Efforts should be made to bridge the digital divide in rural areas. As a step towards enabling farmers in the rural areas and minimizing the digital gap, the United Nations (UN) sponsored the WSIS in Geneva in 2003 and in Tunisia in 2005 (ITU, 2003). The summit declared many goals and targets, including creating an inclusive information society, establishing resilient ICT infrastructure, and reducing the digital gap between developing and developed countries. Moreover, in 2015, the Third International Conference on Financing for Development recommended establishing a collaborative framework by the UN called Technology Facilitation Mechanism (TFM) (UN, 2015). The framework goal is to enhance ICT access and minimize the digital divide through the collaboration and information sharing between scientific and technology stakeholders from the civil society, private sector, UN members, and the scientific community (UN, 2015). These initiatives resulted in noticeable progress; according to the International Telecommunications Union (ITU), the share of internet users in the Arab states increased from 8.5% in 2005 to 66% in 2021. At least 70% of the population in 2021 will be covered by LTE (4G) mobile networks, compared to 19% in 2015 (ITU DataHub, 2022). As for Bahrain, the share of internet users reached 100% by 2019, and by 2020, almost all citizens and residents own a smartphone. On the other hand, ownership of computer devices (desktop, laptop, or tablet) per household increased from 32% in 2001 to 95% in 2021 (ITU DataHub, 2022).

On the other hand, Localized training materials and digital resources should be made accessible to meet every location's unique demand and

problem (Nallusamy, Balasubramaniam, & Chellappan, 2015). To assist farmers in overcoming any difficulties they may experience while employing ICT technologies, ongoing technical support and mentorship are crucial. This might involve providing updates, troubleshooting, and responding to inquiries.

ICT readiness goes beyond individual literacy; it encompasses the broader environment and infrastructure required for successful ICT adoption in agriculture (Raksha, Rao, Meera, & Kumar, 2015). It is imperative for governments and stakeholders to allocate resources towards the development of essential ICT infrastructure in rural areas, encompassing broadband access, mobile networks, and power supplies. Furthermore, digital platforms serve as efficient instruments for providing farmers with access to agricultural knowledge and training materials (Nallusamy et al., 2015), enabling them to acquire information on optimal methods, pest control, and climate-adaptive strategies. However, policymakers should prioritise laws and regulations that encourage the adoption of ICT in agriculture. This can be achieved by measures like providing tax incentives for technology investments and safeguarding the digital rights of farmers (Purnomoi & Lee, 2010). In addition, scientists and developers must consistently conduct research and develop innovative agricultural technology to create cost-effective and user-friendly ICT solutions for farmers.

Public-private partnership is crucial to help facilitate the adoption of ICT in agriculture by providing resources, expertise, and digital services (Moreddu, 2016). Mobile banking and digital payment systems facilitate financial transactions for rural farmers, making accessing credit, saving money, and engaging in financial planning easier.

Accordingly, proficiency in ICT and preparedness are pivotal elements for the effective implementation of digital technology in agriculture. Such

proficiency contributes to heightened productivity, sustainability, and increased income for farmers (Subashini & Fernando, 2017). Empowering farmers with digital competencies, coupled with providing them access to required tools and infrastructure, serves to improve the digital disparity prevalent in rural agriculture. This, in turn, establishes a foundation for a more robust and prosperous farming sector. It is imperative to recognize that the integration of ICT in agriculture transcends technological progress, serving as a means to address food security, alleviate poverty, and foster sustainable development within rural communities on a global scale.

2.2.2. ICT and Extension Services in Agriculture

Agricultural extension is typically defined as the provision of services that aim to spread technology among farmers, guide, instruct, and advise farmers, and develop human capital (Spielman, Lecoutere, Makhija, & Campenhout, 2021).

The increasing global interest in using ICT and bridging the gap between developed and developing countries has led to a significant development in the methods of providing agricultural extension services and the speed of their delivery, as well as increasing the number of beneficiaries from them.

The Food and Agriculture Organization (FAO) has outlined seven essential factors contributing to the efficacy of ensuring the availability and accessibility of ICTs for farmers and rural communities, as elucidated in FAO's 2018 report (FAO, 2018). Firstly, content customization to align with local needs and contexts is imperative. Secondly, capacity development is essential, encompassing access to related information, ICT literacy, and skills enhancement. The third factor underscores the significance of gender and diversity considerations, aiming to provide

equitable access for women, youth, older farmers, and individuals residing in rural areas. The fourth aspect emphasizes the importance of ensuring widespread access and active participation, thereby minimizing the digital divide.

Furthermore, fostering partnerships, mainly through public-private collaborations, is identified as a crucial factor. The sixth factor recognizes the challenge of identifying an optimal mix of technologies tailored to local needs. Lastly, economic, social, and environmental sustainability considerations are integral to the success of ICT integration in agriculture and rural development initiatives. These seven interrelated factors collectively form a comprehensive framework for fostering the effective deployment of ICTs in agricultural contexts, thereby contributing to sustainable development.

Adopting ICT tools will advance access to extension services by reducing the costs of agents' physical visits and improving agents' availability and accountability. Extension services can be either through agriculture agents that provide training, knowledge transfer, and expert advice to farmers, or they can be facilitation to connect farmers with different agriculture stakeholders, such as the market and the government, through mobile apps or web platforms (FAO, 2017). It can be implemented through different paradigms, including online platforms, mobile applications, and radio and TV specialized programs.

One of the main contributors to the socioeconomic improvement of rural farmers' lives is the provision of extension services based on ICT. According to Ravindra, Sehgal, and Narasimhamurthy (2020), ICTs like TVs and smartphones have been crucial in giving Karnataka farmers useful information regarding agricultural inputs. Nonetheless, the study concluded that the primary obstacles to ICT adoption include low ICT literacy, a lack of knowledge about the advantages of ICTs, and a lack of

programmes and organizations for vocational training that emphasize ICT use in agriculture.

In Pakistan, farmers firmly believe that utilizing ICTs such as mobile phones and the Internet will enable them to get accurate and timely information, establishing a direct link with agriculture extension service providers. The study's findings indicated that farmers with higher levels of education exhibit a greater willingness to embrace new technologies. Consequently, it is advised to provide training to extension service providers in the use of ICTs (Aldosari et al., 2019).

Fabregas, Kremer, & Schilbach (2019) broaden the ICT focus to incorporate GPS-enabled cellphones and subscription-based mobile services, emphasizing their potential to provide farmers with personalized information and motivate them to share information with the system.

2.2.3. Smart Agriculture

The rapid penetration of technologies in the agricultural field, such as the internet, AI, Blockchain technologies, IoT, and Big Data, has made a quantum leap in farming practices. It has been dominated by names such as e-agriculture, smart agriculture, and digital agriculture (Bilali et al., 2020).

Smart agriculture indicates the deep integration of modern information technologies such as the Internet, IoT, Big Data, cloud computing, and AI. This integration advances decision-making, precision farming, and information perception and enables personalized services such as water management, soil and crop monitoring, field management, etc. (Yang et al., 2021). The adoption of smart agriculture practices has been expanding, covering almost every agriculture domain, including crops, livestock, soil, machines, pests, fertilizers, and human power (Alreshidi, 2019).

According to Yang et al. (2021), IoT is a key technology in smart agriculture applications, and the authors identified seven IoT applications for smart agriculture, including field management, greenhouse automation, livestock monitoring, plant factory, aquaculture, photovoltaic agriculture, and solar insecticidal lamps. IoT refers to the billions of remote devices connected wirelessly to complex networks and forms of intelligent architecture, such as smart buildings, smart factories, and smart schools. It is a concept based on features that combine physical and virtual interconnected devices to communicate and exchange information (Sorri, Mustafee, & Seppänen, 2022). These devices include sensors and actuators, smartphones, Radio-frequency identification (RFID) tags, and smart vehicles with embedded systems enabling smart automation of daily activities such as irrigation, livestock status, and soil and crop management (Tariq et al., 2019).

Smart irrigation systems have proven efficiency in water consumption, where countries that adopted these technologies have improved water consumption by 10-50% and noticed an increase per unit yield of land and water by 20-100% (Ali et al., 2019). Water management systems use IoT technologies such as weather stations to send real-time weather forecasts to farmers to control water consumption and minimize irrigation costs. Moreover, sensor nodes connected to cloud services are utilized to monitor water quality by automatically sending live measurements of pH, temperature, conductivity, and oxygen levels (Farooq, Riaz, Abid, Abid, & Naeem, 2019).

ICT applications for soil and crop monitoring are becoming essential tools for farmers and other agriculture stakeholders. Crop management is considered an exhausting and time-consuming operation to be carried out by farmers physically. It includes various tasks such as crop spraying and crop scouting; however, the advances in sensing technologies have made

such tasks effortless (Clercq et al., 2018; Mekala and Viswanathan, 2017). Nutrient Manager is an app that farmers can download on their phones or computers, answer a series of questions, and then the answers will be uploaded to the cloud, where analytics will run on the answers and reply with specialized pieces of advice for fertilizing practices either by SMS messaging or as an image (Ali et al., 2019).

China is one of the world's top rice producers; however, their resources are diminishing due to the growing population, and they must cope with the idea of producing more with less. 5G+ MEC is a project that employs the latest AI tools and technologies to ensure smart real-time rice and vegetable cultivation decisions. The project combines applications of 5G, AI, a deep learning toolbox (DeepField), and a customized database. It has reported a cut in field operating costs, reduced carbon emissions, and improved quality and production (ITU, 2022).

On the other hand, machines and technology serve as the primary instruments for embracing this change. Advanced GNSS-enabled tractors and sprayer machines have the capability to autonomously navigate around fields with utmost precision, eliminating the need for human drivers. An essential benefit of utilizing auto-steering machines is the reduction of human exertion without sacrificing mechanical efficiency. These machines are operated by real-time kinematic GNSS sensors, which improve the precision of lateral movement to less than five centimetres when moving in straight paths. (Lange & Peake, 2020).

IoT has been introduced in greenhouses to enable smart remote control of the cooling and ventilation of the environment based on data collected from CO₂, soil, temp, and light sensors (Pallavi, Mallapur, & Bendigeri, 2017). The mixture of IoT and AI has been included in many agriculture tasks, such as (farm management systems, crop scouting, and livestock

monitoring) to enhance business process efficiency, improve crop production, and promote environmental protection (Alreshidi, 2019).

2.2.4. Precision Agriculture

PA is a concept encompassing different farming aspects such as crop production, soil, and plant genetics (Bucci, Bentivoglio, & Finco, 2019). Achieving PA requires collecting and analysing data using different ICT approaches. Based on the National Research Council, PA is defined as "a management strategy that uses information technology to bring data from multiple sources to bear on decisions associated with crop production" (Adrian, Norwood, & Mask, 2005). It depends on different vital factors including but not limited to field data, amount of inputs, governmental and non-governmental support, and farmers' knowledge.

ICTs in PA have been used for more than 20 years in developed countries, confirming the digital gap between developed and developing countries. As per the ICT in Agriculture report by FAO (2017), it has been utilised in many large, developed countries such as Australia and the United Kingdom to minimise the machinery and inputs cost, live monitoring of animals' growth, diseases, and production (milk, eggs) at Southwest England. Moreover, ICTs feed the Big Data platforms and applications such as the Farmers Business Network in the United States of America, which improves PA practices. ICTs provide PA with a set of tools that enable in-farm management efficiently and can be called Precision Agriculture Technologies (PATs) (Aubert, Schroeder, & Grimaudo, 2012; Lowenberg-DeBoer & Erickson, 2019). This toolkit includes technologies such as auto guidance technology (auto tractors and sprayers) that depends on the GNSS, data collection technologies including (handheld sensors, wireless sensors, and remote sensing), and variable rate technologies (e.g., in irrigation and nutrients) (Barnes et al.,

2019). Adopting PATs is increasing rapidly, specifically in the US market, where farms are vast, and PATs over the cost will be more profitable. According to MarketsandMarkets (2023), the rapid uptake of technological advancement is expected to grow the global PA market from USD 9.7 billion to USD 21.9 billion by 2031. However, countries with small farms and fewer ICT-educated farmers tend to have a prolonged adoption rate of PA technologies (Bucci et al., 2019).

2.2.5. ICT in Agriculture and Sustainability

The 2030 agenda confirmed that ICTs are significant in achieving the SDGs. In 2015, all United Nations member states agreed that by 2030, a plan of 17 goals and a total of 169 targets are to be achieved (Wu, Guo, Huang, Liu, & Xiang, 2018).

The attainment of the SDGs represents a significant step towards eradicating poverty and hunger, fostering good health, and facilitating quality education on a global scale. These goals also encompass universal access to clean water and the provision of affordable and clean energy. Moreover, the SDGs are intricately linked to fostering economic growth and facilitating progress in various industries. By addressing these interlinked objectives, the international community aims to mitigate the challenges posed by climate change, safeguard biodiversity on both terrestrial and marine fronts, and work towards the establishment of a sustainable world. This endeavour underscores the significance of adopting responsible production and consumption practices to ensure the long-term viability of our planet (Wu et al., 2018).

Agriculture sustainability is a significant objective for attaining the SDGs; it affects and is affected by many goals, either directly or indirectly. Agriculture has been connected mainly to two goals: SDG1 (End poverty in all its forms everywhere) and SDG2 (End hunger,

achieve food security and improved nutrition, and promote sustainable agriculture). However, in many contexts, it has been linked to almost all other goals (Scown, Brady, & Nicholas, 2020; Nhemachena et al., 2018; Priyadarshini & Abhilash, 2020). Table 2.1 presents goals and targets related to agriculture.

Table (2.1): SDGs Interlinked to Agriculture.

Source: (UN, Global indicator framework for the Sustainable Development Goals and targets of the 2030 Agenda for Sustainable Development, 2022)

Goal	Target	Indicator
SDG 1	1.4	1.4.2 Proportion of total adult population with secure tenure rights to land, (a) with legally recognized documentation, and (b) who perceive their rights to land as secure, by sex and by type of tenure
SDG 2	All	All
SDG 5	5.a	5.a.1 (a) Proportion of total agricultural population with ownership or secure rights over agricultural land, by sex; and (b) share of women among owners or rights-bearers of agricultural land, by type of tenure
		5.a.2 Proportion of countries where the legal framework (including customary law) guarantees women's equal rights to land ownership and/or control
SDG 12	12.3	12.3.1 (a) Food loss index and (b) food waste index
SDG 13	13.2	13.2.2 Total greenhouse gas emissions per year
SDG 14	14.4	14.4.1 Proportion of fish stocks within biologically sustainable levels
	14.6	14.6.1 Degree of implementation of international instruments aiming to combat illegal, unreported and unregulated fishing
	14.7	14.7.1 Sustainable fisheries as a proportion of GDP in small island developing States, least developed countries and all countries
	14.b	14.b.1 Degree of application of a legal/regulatory/policy/institutional framework which recognizes and protects access rights for small-scale fisheries

SDG 15	15.3	15.3.1 Proportion of land that is degraded over total land area
--------	------	---

Digitalization has become a substantial tool for achieving agriculture sustainability. Its significance has been recognized since the Millennium Development Goals (MDGs) through the first goal (MDG1: Eradicating Extreme Poverty and Hunger) and the eighth goal (MDG 8, target F: In cooperation with the private sector, make available benefits of new technologies, especially information and communications) (Clarke, Wylie, & Zomer, 2013).

With advances in technology and the rise of emergent technologies such as Big Data and IoT, the impact of ICT on agriculture sustainability is emphasized and has led to increased production, minimized natural resources consumption and management costs, and promoted accountability (Bilali et al., 2020).

Initiatives and projects using technological advancement pave the road towards sustainable production and consumption in the agriculture sector. Water consumption can be minimized by precisely controlling it through smart irrigation systems that depend on IoT devices (Keswani et al., 2020). When combining IoT and AI tools, seamless control of greenhouses without human intervention becomes possible; this will minimize input errors and achieve higher efficiency in processes (Mekala & Viswanathan, 2017). UAVs have proven effective in spraying and monitoring crop health remotely (Alreshidi, 2019). Moreover, farm management information systems are handy tools that provide precise decisions and actions based on the analysis of data collected by sensors, drones, weather stations, etc. (Knierim et al., 2019).

On the other hand, agriculture practices are responsible for 19-29% of greenhouse gas (GHG) emissions, which are among the main contributors

to the climate dilemma. In the first climate action plan, the World Bank promised to collaborate with countries to deliver a new approach called Climate-Smart Agriculture (CSA) (World Bank, Climate-Smart Agriculture, 2021b). It is a concept that aims to incorporate ICTs in agriculture to increase productivity, enhance resilience, and reduce emissions through the management of agriculture practices such as minimizing tillage, precision input management, and using high-yielding and stress-tolerant breeds (Annosi et al., 2020; Aryal, Sapkota, Rahut, & Jat, 2020). This new term is now rising and replacing the known (smart agriculture) concept by focusing on fighting climate change by efficiently managing agriculture practices through technology.

Accordingly, smart or climate-smart agriculture concepts depend on many technologies to collect, process, and store data for analytics and decision-making. The data collection process will include different data formats, exponential increase in the size of collected data, and variable collection speed (real-time and batch). Traditional solutions such as relational databases will not be enough to handle the heterogeneity in the data, so emergent technologies such as Big Data are essential.

2.3. Big Data

Every day, the number of devices generating data from/to the cloud increases massively. In 2017, CISCO reported a global mobile data traffic of 11.2 exabytes per month, which is expected to double 13 times within five years (Amalina et al., 2019). According to the National Security Agency, the amount of data processed by the Internet each day is 1826 petabytes (PB) (Hariri, Fredericks, & Bowers, 2019). Moreover, in 2018, 2.5 quintillion bytes were produced daily. Data is an essential tool to extract insights and predictions to improve all aspects of every sector. It is generated from different sources such as mobile phones, cars, and

sensors, and these generated data have different formats that would not be compatible with traditional relational database systems. This enormous amount of generated data is so-called “Big Data” and is characterized by its massive size, diversity, and high generation rate (Hashem et al., 2015). According to Mauro, Greco, & Grimaldi (2016), “Big Data is the Information asset characterized by such a high Volume, Velocity and Variety to require specific Technology and Analytical Methods for its transformation into Value.”. The authors suggested that Big Data has been identified in related literature based on four domains: data attributes, technological needs, size, and social impact. On the other hand, it has been classified based on five categories: data sources (smartphones, IoT), content format (images, documents, video), data stores (document-based, key/value pairs), data staging (ELT/ETL), and data processing (batch, real-time) (Hashem et al., 2015).

Big Data was first identified in 2001 by the three characteristics (Volume, Velocity, and Variety); later, in 2011 and 2012, additional characteristics (Value and Veracity) were added. Although Big Data has many V characteristics (Validity, Viability, Virility, Viscosity, Vitality, Volatility, Vagueness, and Venue) (Niculescu, 2020), the focus is primarily on the primary three Vs (Volume, Velocity, and Variety). However, in many cases, an additional two to four characteristics are discussed, as depicted in Figure 2.1, including Veracity and Value (Ceravolo et al., 2018), Variability (Saranya & Asha, 2019), and Visibility (Niculescu, 2020).

Volume	•Refers to the size of data. •The introduction of different data sources and different data types in the daily computation processes resulted in an enormous amount of generated data. •It is expected that the global data sphere will grow from 45 Zetta Bytes to 175 Zetta Bytes by 2025
Velocity	•The speed at which data is generated, captured, and processed. •The advances in telecommunication technologies has improved data throughput, and as a result data rate has evolved from batch to realtime and live streaming
Variety	•Due to the increase in data sources, heterogeneity of data is increase as data now classified as structured, semi-structured, and unstructured. •The introduction of technologies such as smart phones, IoT devices, and remote sensing transformed how data are presented. •Data are presented in different formats including images, text, videos, and files.
Value	•Extracting insights and predictions to improve the social, economic, and environment factors. •Value is achieved through revealing hidden patterns and discovering new relations and associations among the data. •Advanced analytical technologies such as data mining, deep learning, and neural networks are the main tools to extract value
Veracity	•Reliability of the collected data •One of the major challenges in the big data field is to ensure the integrity of the generated data. •MetaData is one of the main techniques that mitigates the data bias and abnormality.
Visibility	•Refers to data accessibility •Blocking data can result in an incomplete analysis which may produce misleading knowledge. •On the other hand it is related to privacy. Data are either public such as weather forecast or private like bank accounts.
Variability	•The uncertainty in data generation rate. •Continues change in format, structure, and the amount of data requires increase in computing power and storage, and hence increase in cost.

Figure (2.1): The Big Data characteristics (7 Vs)

Big Data has been an enabler to many AI paradigms, including machine learning, data mining, natural language processing, and computational intelligence. The intelligent algorithms adopted by these technologies are designed to extract knowledge, discover relations, and reveal concealed patterns (Hariri et al., 2019).

Big Data employs a wide range of established and emerging technologies, including data lakes, data warehouses, cloud computing, Extract – Transform - Load processes (ETL), and data pipelines. The applications of Big Data follow traditional data management approaches, starting with the stages of data collection, processing, storage, analytics, and visualisation. However, the unique features of Big Data, which allow it to handle many forms of data, such as structured, semi-structured, and unstructured data, have caused a change in specific old methods. This advancement is a direct result of the distinct issues presented by the unique characteristics of Big Data. It necessitates a departure from traditional paradigms in order to ensure efficient data management.

2.3.1. Big Data Formats

The boom in telecommunications and the rapid digital progress enabled the generation of vast data types such as images, videos, sensor data, and text. These types are classified into three categories: (i) structured (predefined format), (ii) semi-structured (implicit or irregular format), and (iii) unstructured (random, disorganized format) (Sambrekar, Rajpurohit, & Joshi, 2018).

Structured data has predefined data models where it is stored and processed using Relational Database Management Systems (RDBMS) such as Microsoft Access and MySQL, and it is processed using Structured Query Language (SQL) (Södergård, Mildorf, Berre, Tsalgatidou, & Charvát, 2021). Traditional databases are based on schema-on-write data management, where the schema is created before data is collected (organizational transactions, objects) (Panwar, 2020). It tends to be limited in size and structure compared to other categories; however, analyzing, visualizing, and presenting data is much easier. Despite the relative ease of dealing with structured data, the presence of

Big Data applications caused a heterogeneity of data sources in one field, leading to increased requirements and complexity.

According to Khine & Wang (2018), around 70 – 80% of data generated in any sector is either semi-structured or unstructured. Semi-structured and unstructured data have no schema; the structure is either implicit, such as in Extensible Markup Language (XML) data or no structure, like in a tweet message. These types of data are called schema-less, and they are challenging to analyze directly, which, in most situations, are converted to structured data (Sambrekar et al., 2018). The data management schema used to handle such data types is called schema-on-read, and it depends on technologies such as distributed file systems (Hadoop) and NoSQL.

2.3.2. Big Data Storage and Processing

ETL is the traditional approach used to collect data from sources. It runs different transformations and cleans the data to store it in conventional repositories (schema-on-write) such as data warehouses and data marts (Panwar, 2020). Data extraction is considered an intensive step that requires preparation and modelling of the data sources and is divided into initial extraction and incremental extraction (El-Sappagh, Hendawi, & Bastawissy, 2011). Initial extraction happens only once when connecting to the data source and extracting its data for the first time; after that, it will occur when data changes (new or updated transactions). Transformation is a significant operation that determines and processes the extracted data to become valid for storage in a data repository. It involves many tasks, including 1- cleaning and unifying the records, 2- filtering the data (selecting specific columns or records), 3- merging/splitting the data, and 4- producing new data using aggregation and calculations (Mali & Bojewar, 2015).

The traditional ETL approach can handle a few data sources with a predefined format; however, the increase in data sources, format heterogeneity, and generation speed will require revisiting and modifying the approach. Researchers either modulate the approach or propose new concepts to improve its performance. Therefore, Extract – Load – Transform (ELT), a new process, emerged to enable the collection and storage of heterogeneous data directly in repositories called data lakes, where data is stored in their raw format (images, text, files, etc.) (Zouari, Kabachi, Boukadi, & Ghedira-Guegan, 2021). Moreover, Marín-Ortegaa, Dmitriyevb, Abilovb, & Gómezb (2014) suggested an extension to the approach where analysis is included in the process for business intelligence (BI) solutions as follows: Extract, Load, Transform, and Analyze (ELTA). On the other hand, Mallek, Ghozzi, & Gargouri (2020) proposed a solution called (BigDimETL) that depends on HBase to create a column-based database out of the extracted data, then utilize MapReduce algorithms to transform the data and finally load it into a multidimensional warehouse. Further, a semantic ETL framework to classify related datasets before storage by generating a Resource Description Framework (RDF) was suggested by Bansal & Kagemann (2015).

The remarkable diversity and abundance of data sources, which led to the massive increase in data types and the speed of their generation, requires new technologies to keep pace with this progress. Storage technologies improved and varied significantly during the past few years, and the advancement in data repositories derived mainly from the increased complexity and heterogeneity of data generated (Strohbach, Daubert, Ravkin, & Lischka, 2016).

Digital data storage technologies and tools go back to the 1950s when IBM released the first magnetic storage disk in 1953. It was enormous

(physical size), costly, and limited in storage size. Later, from 1970 through 1980, IBM launched the flexible (floppy) disks that were smaller in size, more prominent in capacity, and much cheaper (Bhushan, 2018). Then, the evolution continued, and the 1990s were the golden period for CDs, DVDs, and HDDs until 2010 when the revolutionary SSD storage technology rose. On the other hand, the 1970s was also the time when the relational database concept was introduced, and specialized storage systems were developed to support concepts such as (Database Machines) (Corbellini, Mateos, Zunino, Godoy, & Schiaffino, 2017; M.H.Padgavankar & S.R.Gupta, 2014). It is a monoprocessor computer dedicated only to database management; however, the demand for more storage and processing power for more extensive databases introduced new concepts. In the 1980s, the idea of sharing data using multiprocessor computer was a solution to overcome the data storage and processing demand and was called a parallel database management system (Valduriez, 2009). According to Morris & Truskowski (2003), the period of the 1980s and 1990s was critical when new advanced models for storage and processing, such as Client/Server and network storage concepts, were introduced.

Storage technologies are divided into Direct Attached Storage (DAS) and network storage, which is further divided into Network Attached Storage (NAS) and Storage Area Network (SAN) (Chen, Mao, Zhang, & Leung, 2014). In DAS systems, disc drives such as (HDD) are connected directly to the servers and managed by RAID (redundant array of independent disks) technologies. However, due to physical limitations, this system doesn't support the rapid increase in data sizes and challenging upgradeability features. On the other hand, NAS is a set of file access protocols such as NFS, HTTP, and FTP that enable flexible access to storage over networks, while SAN is the system used to manage storage

blocks on networks. SAN structure paved the way for improved storage technologies such as block virtualization and the SAN file system, enabling efficient data sharing and disaster recovery (Morris & Truskowski, 2003). The increase of online services in the late 1990s led to the inability of available technologies to keep up with the rapid increase in the volume and diversity of data. In 2000, the CAP (Consistency, Availability, Partition Tolerance) model was introduced to build distributed storage systems; however, the proposed techniques could not meet the CAP requirements combined. Yet, the proposed systems comply with 2 out of 3 requirements in the following combinations: CA system (Consistency, Availability), CP system (Consistency, Partition Tolerance), and AP system (Availability, Partition Tolerance). The CP-based solutions, such as Bigtable (based on the Google file system) and HBase (Chen et al., 2014), were widely used when data accuracy and scalable storage are much more crucial than real-time availability.

In 2003, google introduced a distributed storage system called Google File System (GFS) as a solution for data-intensive applications. The system was designed to store file chunks of a maximum of 64 KB per block and the ability to replicate these chunks to ensure data reliability (Siddiqua, Karim, & Gani, 2017). Nevertheless, the exponential increase in data size and heterogeneity (structured and unstructured) led to the emergence of more flexible and scalable solutions such as the Bigtable cloud service. It was launched in 2005 by Google as a NoSQL wide-column database that can accommodate petabytes of applications using key/value mapping (Google, Cloud Bigtable, n.d.). The term NoSQL was first used in 1998 to refer to (not having an SQL interface) and then redefined to become (not only SQL) (Corbellini et al., 2017). NoSQL is a schema-less database that has more flexibility than relational databases in

handling Big Data size and different formats and has four database categories: 1- key-value, 2- wide-column, 3- document, and 4- graph databases (Davoudian, Chen, & Liu, 2018).

In 2006, Apache launched its open-source project, Hadoop, to handle the massive increase in data through a set of developed tools under that project (Siddiqa et al., 2017). The Hadoop project became a standard for many Big Data solutions with two main backbone models, the Hadoop Distributed File System (HDFS) and the MapReduce paradigm. HDFS is a distributed file system based on GFS that can handle massive data files, and due to its sheer number of components, it is highly available, fault-tolerant, and supports automatic recovery. MapReduce is a set of algorithms that facilitate parallel processing of large batch files by splitting the files across different logical clusters (Hashem et al., 2015). It is also based on GFS and has become a highly adopted model for cloud service providers. MapReduce consists of two operations, Map and Reduce, where Map divides and transforms the file into a set of key/value pairs, and then the output is converted into smaller data sets ready for operations using the Reduce functions (Oussous, Benjelloun, Lahcen, & Belfkih, 2018).

Data storage is now far from being a simple task; it is cloud-based with auto-scaling and the availability of enormous capabilities at the user's fingertips. Establishing an efficient and reliable data warehouse is essential to perform business intelligence analysis on data; however, a data lake (Table 2.2 presents the difference between Datawarehouse and Data Lake (Aissi et al., 2022; Coursera, 2022; Khine & Wang, 2018)) must be used as a pre-processing step to handle raw data (Neves & Cruvinel, 2020).

The data lake is a new storage concept introduced in 2010 by the CEO of Pentaho as a solution to the increased complexity and diversity of data

generated (Khine & Wang, 2018). Data lakes provide an unconditional repository for data, and it can custom-tailor parts of these data to form a valuable piece of information based on user requests.

Table (2.2): Datawarehouse Vs. Data Lake.

Criteria	Data Warehouse	Data Lake
Data Format	Structured	All (Structured, Semi-structured, Unstructured)
Data Schema	Pre-defined and fixed schema definition (schema on write)	No Schema (schema on read)
Storage	Low agility, complex	High agility and flexible
Process Scheme	Extract, Transform, Load (ETL)	Extract, Load, Transform (ELT)
Use Cases	Business Intelligence, Data visualization, data analytics.	Predictive analytics, machine learning, data visualization
Users	Business Professionals	Data scientists
Cost	Relatively high	Relatively low
Maintenance	Time consuming due to its complex nature	Less time required for management

The impressive advance in cloud services has led to a multiplicity of data storage services and tools, whether for individuals or enterprises, without the need to build special data centres, which saves effort and technical expertise and enables cost flexibility (Januzaj, Ajdari, & Selimi, 2015). Platforms such as AWS, GCP, and Microsoft Azure provide multiple storage tools for all scenarios, including data warehouses, data lakes, relational databases, and NoSQL. Table 2.3 lists the storage services according to the use case for the platforms Azure, AWS, and GCP (AWS, n.d.; Azure, n.d.; Google, Google Cloud databases, n.d.).

Table (2.3): Cloud Storage Services.

Platform	Relational DB	NoSQL	Data Warehouse	Data Lake
GCP	Cloud SQL (MySQL, PostgreSQL, SQL Server)	Cloud Bigtable	Google BigQuery	Google Cloud Storage
AWS	Amazon Aurora (MySQL, PostgreSQL), Amazon Relational Database Service (RDS)	Amazon DynamoDB	Amazon Redshift	Amazon S3
Microsoft Azure	Azure SQL Database, Azure Database for PostgreSQL, Azure Database for MySQL, Azure Database for MariaDB	Azure Cosmos DB	Azure Synapse	Azure Data Lake Store

Although cloud storage is highly scalable, available, fault-tolerance, and has low latency, it also faces many security and privacy concerns (Strohbach et al., 2016).

2.4. Big Data in Agriculture

A sustainable agriculture system relies heavily on data to create knowledge that will eventually lead to effective decisions. Analysis of Big Data has proven effective in many fields, including agriculture. According to the World Bank, data points per farm per day will increase from 190,000 to 4.1 million by 2050, which will create a massive amount of data to be analyzed daily (WorldBank, 2021a).

Agriculture data collection and analysis initiatives using ICTs are increasing rapidly. Research and studies related to Big Data in agriculture can be roughly divided into five categories as follows:

1. Scholarly reviews on Big Data in Agriculture (Yang et al., 2021; Zhang, Jakku, Llewellyn, & Baker, 2018). These studies, which are distinguished by their comprehensive approach, examine the different aspects of Big Data applications in the agricultural sector. Academics involved in these studies conduct thorough evaluations across a range of elements, including the practical deployments, advantages, and disadvantages linked to the integration of substantial volumes of data. Moreover, the analyses encompass a comprehensive discussion of the environmental, social, and economic consequences that arise from the implementation of Big Data technologies within the agricultural industry. Furthermore, the academic conversation delves into an examination of pertinent policies and regulations that regulate the application and utilization of Big Data in the agricultural sector.
2. Big Data in agriculture frameworks. They serve as foundational structures for developing comprehensive Big Data applications. These frameworks essentially represent the initial stages or essential components necessary for the construction of fully qualified Big Data applications tailored to the agricultural domain (Gill, Chana, & Buyya, 2017; Neves & Cruvinel, 2020). The utilization of such frameworks becomes imperative as they provide a structured approach, laying the groundwork for the subsequent development and implementation of sophisticated Big Data applications aimed at addressing diverse challenges and opportunities within the agricultural landscape.
3. The technologies implemented in big data solutions for agriculture. The integration of Big Data into agricultural technologies signifies the combination of tangible technologies, such as remote sensing, IoT devices, and robotics, with the vast capabilities of Big Data,

which is commonly stored on cloud platforms (Ray, 2017; Huang, Zhong-xin, Tao, Xiang-zhi, & Xing-fa, 2018). A pertinent collection of agricultural data is obtained in real-time by employing IoT devices, robotics, and remote sensing technologies. Following this, the collected datasets are transferred to Big Data repositories hosted in the cloud, where advanced analytics are utilized to extract significant insights. Subsequently, the acquired knowledge is utilized to provide insights and direction for agricultural decision-making processes, encompassing resource allocation and precision farming approaches.

4. Big Data Analytics (BDA) for agriculture constitutes the strategic application of advanced computational techniques and analytical methodologies to harness the vast reservoirs of data generated within agricultural systems. This paradigm leverages the fundamental capabilities of Big Data by employing sophisticated methodologies such as machine learning, data mining, and image processing systems (Evert et al., 2017; Majumdar et al., 2017). By harnessing the power of these cutting-edge technologies, practitioners in the agricultural sector can derive actionable intelligence, enhance productivity, optimize resource allocation, and address the challenges that characterize contemporary agricultural landscapes.
5. Finally, the Big Data engineering processes in agriculture. The complex nature of Big Data engineering processes results in a comprehensive exploration of approaches related to data collection, transmission, transformation, and storage. Within this context, scholarly researches primarily focus on the technical methods pertaining to the collection and manipulation of data, the transformative processes such as cleaning and aggregation, and the

establishment of robust data storage frameworks, including data lakes, warehouses, and data marts. (Ngo, Le-Khac, & Kechadi, 2019; Nguyen, Ngoc, & Kim, 2017).

Review studies on Big Data in agriculture addressed different factors such as data domains, data privacy and legal concerns, and the social impact. Agriculture Big Data is a result of the collection of individual farms' data to form one large data repository which can be utilized by agriculture stakeholders and the country's economy to their advantage. However, Coble, Mishra, Ferrell, & Griffin (2018) argue the need for legal protection for agricultural data similar to other sectors (health and financial) so data owners can trust sharing. Applications of Big Data are data-domain dependent. Sarker, Islam, Murmu, & Rozario (2020) identified five data domain applications: weather forecast, crop yield prediction, selection of best crops for a specific area based on analysis of climate, detection of pests and diseases, and market information systems. Minet et al. (2017) analyzed agriculture crowdsourcing initiatives by studying the applications, data inputs, data quality, and involved parties. The authors recognized four crowdsourcing applications (analysis tasks, visual observations, sensors generated data, and knowledge platforms) and eight data categories, including land use and cover, soil data, weather data, crop rotation calendar, pests and diseases, crop yield, market data (prices), and general knowledge. Moreover, Alreshidi (2019) identified eleven interconnected data domains for a smart agriculture system encompassing human power, crops, soil, weather, fields, machines, pests, fertilization, irrigation, livestock, and market data.

In Australia, a cross-industry survey was conducted to identify agriculture stakeholders' needs and the benefits and barriers to adopting Big Data in the agriculture sector (Zhang et al., 2018). The authors conducted the study on 1000 stakeholders in the period of 7 March – 18 April 2017, and

the results revealed the following: the current telecommunication infrastructure is not reliable to manage data movements wirelessly, there is a need for more data collection to achieve better results, and despite the willingness to share data, some participants raise privacy concerns and limitation to share with specific parties.

Agriculture as a Service (AaaS) is an IoT and cloud-based Big Data framework designed to modernize and enlighten the Indian agricultural sector. The framework is composed of two subsystems: the first is the user subsystem, which illustrates the diverse methods by which agriculture users interact with the system through multiple interfaces such as web platforms and mobile applications, and the second is the cloud subsystem, which utilizes the Platform as a Service ANEKA to collect, store, and process agricultural data (Gill et al., 2017). Huang et al. (2018) proposed the FLTL framework, which consists of twelve levels and four layers. This framework is intended to manage and analyze vast quantities of remote sensing Big Data for field-specific precision agriculture. It concentrates on high-resolution imaging technologies, including UAV-obtained low-altitude remote sensing (LARS) images. The framework under consideration is a modified iteration of the fifteen-level, five-layer (FLFL) framework that was initially designed by the Chinese Academy of Sciences (CAS), Institute of Remote Sensing and Digital Earth. Adopting ETL/ELT models in framework design becomes necessary in developing efficient Big Data applications. Neves & Cruvinel (2020) proposed a theoretical semantic-based structuring framework of Big Data for agriculture risk management. The framework will extract different agriculture data (structured, semi-structured, and unstructured) from multiple sources and load it directly into a data lake. Then, the data will be distributed into various data marts based on subject (soil, plant, weather). The next step is to move the data to a central warehouse where

different transformations are applied using data mining based on machine learning algorithms to extract knowledge. Finally, additional filters are applied to the mined data to enhance its quality before loading them as data vectors.

Remote sensing through satellites is one of the earliest methods that enabled monitoring the green coverage on the earth. Technological advancement and its applications have paved the way towards more straightforward and lower-cost tools for establishing geospatial agriculture Big Data. WebGIS is an online platform connecting individual GIS applications and is called "A System of Systems" (Delgado, Short, Roberts, & Vandenberg, 2019). The platform has been adopted by many collaborative initiatives, such as the Agricultural Collaborative Research Outcome System (AgCROS), a platform for communicating agricultural discoveries and techniques in the fields of soil health, greenhouse emissions, and crops. IoT technologies are used to connect sensors with other devices to build a wireless network that enables seamless data movement, and it has been applied within many Big Data domains, including agriculture, smart cities, healthcare, energy, and many others (Ge, Bangui, & Buhnova, 2018). IoT and Big Data can minimize water consumption by controlling irrigation valves by creating a soil moisture content distribution map. The system uses ZigBee wireless sensors to collect soil moisture data, air temperature, humidity level, CO₂, and sunlight intensity. The data is then loaded into a Hadoop cluster using a Raspberry Pi gateway and analyzed using two methods (Deep Neural Networks and Random Forest models) (Keswani et al., 2020).

Data engineering studies related to agriculture Big Data focusing on data ingestion, transformation, and storage are not sufficiently presented. Available studies either demonstrate the complete ETL/ELT processes or focus on only one of the processes, such as the storage or transformation

of data. Data transformation is where data is cleaned and converted to a suitable format for analysis. Sambrekar et al. (2018) used Couchbase (NoSQL tool) to transform unstructured agriculture Big Data into semi-structured or even structured data. The framework suggests the collection of agriculture data from various unstructured sources such as (text and XML) and then applying the MapReduce (Distributed Data Processing Algorithm) algorithm on Couchbase to convert it into structured data. Ngo et al. (2019) designed a data integration module for agriculture Big Data by combining Apache Hive (a data warehouse running OLAPs), MongoDB (for document storage), and Apache Cassandra as a data lake for raw data. Full implementation of data engineering processes in agriculture Big Data was found on Nguyen et al. (2017). The study is based on a platform built and deployed on five machines connected through TCP/IP. Apache Flume and Sqoop were used to collect real-time data and import data from Comma-Separated Values (CSV) and Excel files. For large files, MapReduce algorithms were adopted to manage them in parallel. Subsequently, the collected data are stored either on HDFS for archival or Hbase and Hive for random access and real-time analytics.

On the other hand, agriculture Big Data analytics, which involves extracting the value of agriculture data sets using machine learning, neural networks, and data mining, has been studied heavily. According to the DLG (2018) report, Big Data analytics applications in Germany reached around 60% of digital implementations in the agriculture sector in 2017. Nevertheless, the report also affirms that the successful integration of Big Data in agriculture necessitates a robust telecommunications infrastructure. Additionally, farmers should possess ownership of their farm data, deriving tangible benefits from it. Moreover, it is imperative to expand data security and privacy policies to

encompass agricultural data, enhance transparency, and facilitate public access to data without constraints (Coble et al., 2018).

The concept behind Big Data paved the way for many agriculture-related applications and frameworks, many of which are related to analytics to extract value for better decisions. Predictive analysis is one of the leading research methods on agriculture Big Data. Neural Networks (NN), Support Vector Machine (SVM), and Graphical Models are the most used machine learning algorithms for agriculture data analysis (Coble et al., 2018). Evert et al. (2017) suggested a conceptual framework using SVM and NN algorithms to identify crops' weeds and potential diseases by labelling and classifying field images collected using UAVs in the Netherlands. GreenLink is a mobile app that analyzes the data collected using six Wireless Sensor Networks (WSNs) in a small test farm. The sensors collect data about water, weather, energy consumption, and crops, then store the data on the Azure cloud platform for analysis using deep neural networks and regression trees algorithms (Mekonnen et al., 2020). Data mining algorithms such as Partitioning Around Medoids (PAM), Clustering Large Applications (CLARA), and Density-based Spatial Clustering of Applications with Noise (DBSCAN) can provide wheat farmers in India with valuable insights to minimize inputs (resources) and maximize production (Majumdar et al., 2017). Li & Niu (2020) used the K-means data mining algorithm with deep learning to detect the optimal growing environment for crops inside greenhouses based on clustering temperature and relative humidity data. The implementation suggested a five-layer IoT-based Big Data framework. The first layer is where data is collected through sensing devices only. The second layer is for data transportation. The third layer is where business insights are developed. The fourth and fifth layers present the analytics and visualization to different user levels using multiple applications and platforms.

2.4.1. Challenges and Opportunities

Agriculture stands at the intersection of tradition and innovation in today's rapidly evolving world. Traditional farming practices have coexisted with modern technology, resulting in a constantly changing sector. One of the most significant game-changers in recent years has been the integration of Big Data into agriculture. Big Data offers unprecedented opportunities to enhance productivity, sustainability, and agricultural decision-making. However, it also comes with its fair share of challenges that must be overcome to unlock its full potential.

One of the primary challenges in harnessing the potential of Big Data in agriculture is the sheer volume and diversity of data sources (Osinga, Paudel, Mouzakitis, & Athanasiadis, 2022). Farms are equipped with various sensors, drones, and satellite systems that generate vast amounts of data. Integrating and ensuring the quality and consistency of these disparate data sets can be an immense challenge. Furthermore, in the agricultural sector, the accuracy and reliability of data are critical. Erroneous decisions may result from inaccurate data, which may negatively impact farm operations and crop yields (Sarker et al., 2020). Constantly ensuring data quality presents a challenge. Privacy concerns pose a significant obstacle for applications that utilize Big Data. As the agricultural sector increasingly relies on data collection and sharing, data privacy and security concerns become more prevalent. Farmers and stakeholders must address issues related to data ownership, access control, and protection against cyber threats (P.S., Gopal, Chintala, & Renta, 2020).

Access to sophisticated technology and data analytics tools may be an obstacle for those with fewer resources and smaller-scale producers. Bridging the digital divide and ensuring that all farmers, regardless of their scale or resources, can benefit from Big Data solutions is a pressing

challenge (Osinga et al., 2022). The agricultural workforce must acquire new data analysis and technology operation skills to fully leverage Big Data. Training and education programs are essential to equip farmers and agricultural professionals with the necessary expertise.

Although many challenges surround Big Data, the technology solves many agricultural obstacles. Big Data analytics enables precision agriculture, where farmers can make data-driven decisions at a granular level. Resource allocation optimization, such as water, fertilizers, and pesticides, increases yields and reduces environmental impact (Bhat & Huang, 2021). Advanced algorithms can process historical and real-time data to predict crop diseases, pest infestations, and weather patterns. This allows farmers to take proactive measures and mitigate risks, ultimately increasing crop resilience and reducing losses.

Analyzing market trends and consumer preferences using Big Data can help farmers make informed decisions about what crops to grow and when to sell. This can lead to better pricing and increased profitability for agricultural businesses (Coble et al., 2018).

Big Data facilitates transparency and traceability throughout the agricultural supply chain. This helps reduce food waste, ensure food safety, and improve the distribution process's overall efficiency. As per Bhat & Huang (2021), incorporating emergent technologies such as Big Data and Blockchain in the food supply chain improves transparency and reliability of operations. They are pushing forward the progression of technologies like ML and IoT, which utilize Big Data in their operations, to continuously improve sustainability within agriculture. Given the limited expansion of arable land, declining groundwater levels, and the degradation in soil quality, these technologies have the capacity to ensure the efficient use of arable land, water, and other resources, simultaneously meeting global demands while preserving these resources for future

generations. Consequently, Big Data possesses the potential to offer practical and scalable solutions that can aid in the conservation of natural resources, thereby sustaining agriculture (Cravero, Pardo, Sepúlveda, & Muñoz, 2022).

2.5. Agriculture in Bahrain

The Kingdom of Bahrain is an archipelago consisting of 40 islands, covering an area of approximately 711.9 square kilometers. It sits in the middle of the Arabian Gulf, on the eastern coast of the Kingdom of Saudi Arabia. Agriculture played a significant role in the economy of the Kingdom of Bahrain. Prior to the emergence of the oil industry in 1932, the agricultural sector in Bahrain was primarily focused on cultivating date palms. This cultivation was sufficient to meet both local needs and export purposes. There are at least 23 different types of dates that are grown, and various parts of the palm tree, including its leaves, branches, buds, and flowers, are widely utilized (Albilad, 2012). Bahrain has long been famous for its abundant pure water, fruit trees, and exceptional palm trees, particularly in the cultivation of dates. In fact, it was historically referred to as the land of a million palm trees (Albilad, 2012). Nevertheless, in the early 1980s, Bahrain initiated efforts to diversify its agricultural sector by focusing on the cultivation of other commodities like vegetables, flowers, and trees (Shebeeb, 2006).

Agriculture's maximum contribution to the gross domestic product is 0.23%. Agriculture accounted for a mere 1.1% of the overall workforce in 2016 (IGA, Bahrain Open Data Portal, 2016). The agricultural productivity of the Kingdom of Bahrain and the contribution of the agrarian sector to the country's economic growth between 2006 and 2018 are detailed in Table 2.4.

Table (2.4): Agricultural productivity.

Source: (Al-Zayani & Karaymeh, 2020)

Year	Gross Domestic product (Million BD)	agricultural production and fishing (Million BD)	Contribution of the Agriculture Sector to the GDP (%)
2006	4,109	21.3668	0.52%
2010	5,100.20	28.0511	0.55%
2016	11,974.70	34.9	0.29%
2017	12,430.10	34.6	0.28%
2018	12,651.10	35.9	0.28%

Furthermore, it is worth noting that the agricultural land in the kingdom experienced a substantial decline from 6,460 hectares in 1956 to around 4,230 hectares at the start of the 1980s. The per capita area designated for agricultural purposes has decreased from 0.036 hectares in 1924 to 0.006 hectares in 2005, as reported by the Supreme Council for the Environment (Directorate of Environmental Assessment and Planning, 2009). A minority of substantial landowners own agricultural properties. Around 60% of all arable lands are leased to farmers by institutions and individuals under the terms of three-year contracts. 70% of the approximately 2,400 producers do not own the land on which they grow. On the other hand, certain agriculturalists and farm owners implemented a hybrid approach by integrating protected farming techniques, such as the utilization of greenhouses, which were initially implemented in Bahrain in 1976 (Bani, 2020). In recent times, the implementation of hydroponic farming systems has been limited in scope, and the adoption of aquaponic farming as a viable solution to the challenges of water scarcity and land limitation in Bahrain has been observed on even fewer farms. Irrespective of the farming system employed (open land, greenhouses, hydroponics, aquaponics), all farming operations are conducted through manual management. It is uncommon to come across a

farm that utilizes ambient air and humidity sensors to regulate cooling and ventilation systems.

When it comes to agricultural resources, the Kingdom of Bahrain faces a challenge with the availability of water that is suitable for agriculture and irrigation. Providing pesticides and fertilizers that cater to the specific requirements of farmers can be quite a challenge. Furthermore, Bahrain heavily depends on imports to meet its local crop requirements (El-Khoury, 2016). The Urban Planning and Development Authority estimated the total agricultural area in March 2022 to be 2.86% of the country's total land area (UPDA, 2022). Based on the special report (Millennium Development Goals) issued by the Bahraini government in 2015, Bahrain is currently a non-agricultural country. It relies mainly on imported crops (General Directorate of Statistics, 2015). However, the government and citizens of the Kingdom of Bahrain are always proactive in keeping pace with technical and scientific developments in all fields, which contributes to the advancement of the various sectors of the state.

Bahrain, similar to many other countries worldwide, attempts to overcome the challenges arising from population inflation, which encompasses food security. This has been accomplished by implementing and adopting modern agricultural and farming techniques, funding, training and development initiatives, and unifying efforts. As an illustration, the National Initiative for Agricultural Development (NIAD) was launched in 2010 with the objective of supporting producers, advancing the agricultural sector, and enhancing food security within the Kingdom (NIAD, 2019). Furthermore, in the Hoorat Aali area, a permanent market known as the Farmers' Market has been established by the government to accommodate the agricultural endeavors of the local farmers. The purpose of the market is to promote regional agricultural

goods and to draw attention to the progress made by the agrarian sector in Bahrain in order to expand employment opportunities for its citizens.

On the other hand, the government provides easy financial support for farmers to advance their agricultural development through the Bahrain Development Bank (BDB). Heavy agricultural machinery such as tractors is not popular on Bahraini farms due to the relatively small land area. This reduces farms' initial and running costs and poses an opportunity for alternative investment subsidies such as ICTs, especially with the ability to get up to 15,000 BD (approximately 40,000 USD) as a business loan with 0% interest (BDB, 2019). Initial and running costs of a farming business (the government covers 60% of the inputs such as fertilizers) include but are not limited to land rental (relatively low), farm labor, irrigation materials and fittings (manual irrigation), greenhouses structures and covers, IPM tools, cooling machines, fertilizers and hydroponic nutrients, hybrid seeds, and irrigation water (treated water) in case there is no groundwater in the farm.

Bahrain established a resilient infrastructure that enables the adoption and integration of ICT in the agriculture sector. Bahrain's technology adoption was more than 150 years ago, as it was the first country in the Middle East to connect through the telegraphic cable in 1864 (Bahrain Economic Development Board, 2015). Since then, Bahrain has employed the latest technology to provide sustainable public and private services. ICT penetration in Bahrain's market is growing rapidly; adopting emergent technologies such as cloud computing and the IoT led the country to achieve the highest ranks in technology empowerment globally. The UN Urban Agenda Platform's report (2021) reveals that the Kingdom of Bahrain has emerged as the top performer in the Telecommunication Infrastructure Index, securing the first position in the Arab world and ranking 31st internationally in the ICT Development Index (IDI).

Moreover, according to the 2021 International Telecommunication Union (ITU) Report on digital developments in the Arab States, Bahrain achieved the highest ranking among Arab countries in terms of the proportion of its people possessing advanced Information and Communication Technology (ICT) capabilities. Ranked first in the Gulf Cooperation Council (GCC) region for the percentage of GDP derived from trade in ICT services, according to the Technology and Innovation 2021 study by the United Nations Conference on Trade and Development (UNCTAD). According to the Inclusive Internet Index 2021 published by Facebook and The Economist, the Kingdom ranked first worldwide in digital literacy education and first worldwide in training women in digital skills and STEM education (Bahrain, 2021). Moreover, in 2015, the ITU presented the prime minister with the ICT in sustainable development award (IGA, 2018).

ICT penetration in Bahrain's market is increasing; the adoption of emergent technologies such as cloud computing and the IoT led the country to achieve the highest ranks in technology empowerment. It has a robust ICT infrastructure with strong internet coverage nationwide.

As for the digital initiatives for the agriculture sector, the Ministry of Municipalities Affairs and Agriculture (MMAA) has signed four agreements with the FAO to be implemented within the period (2018-2022); the first two agreements are directly concerned with improving crop production and achieving agricultural sustainability. The second agreement relates directly to enhancing the farming statistics and database systems in the Kingdom of Bahrain to support the information infrastructure and develop capacities for preparing, analyzing, and evaluating agricultural data (Municipalities, 2019). In 2021, the agriculture and marine resources agency, the Arabian Gulf University (AGU), and NIAD signed a collaboration agreement with the Dutch

company (Smart Farm Sensing) To combat the red palm weevil using smart sensing technology (Alwatan, 2021). Recently, and precisely on the 25th of May 2022, an online platform to strengthen the agricultural sector in Bahrain was launched (www.agro.bh). The curated database spearheaded by NIAD hopes to assist in accomplishing the kingdom's economic, social, and environmental development goals. The platform aims to provide up-to-date and reliable data that will help further develop principles, knowledge, and skills in the agricultural sector in the Kingdom of Bahrain. Moreover, it will support, encourage, and document agricultural activities and programs cooperating with the public and private sectors.

The agriculture sector in Bahrain comprises many sources for primary and secondary agriculture data that are either publicly available or not including:

1. The Ministry of Municipalities Affairs and Agriculture is the primary governmental entity responsible for the agriculture sector in Bahrain. It owns many data, including laws and legalizations, market data, soil and water data, and crop production and quality.
2. The National Initiative for Agriculture Development (NIAD) is a non-governmental organization that supports the agriculture sector, especially farmers. In the second quarter of 2022, they launched a website (<https://www.agro.bh/>) to share agriculture data they collected, such as available farms, vendors, and farmers. It is considered the first initiative to collect and share agriculture data in Bahrain.
3. The Bahrain data portal, accessible at <https://data.gov.bh/>, serves as the official platform for accessing open data pertaining to many industries in Bahrain. The site has recently undergone an overhaul

to enhance its functionality by including APIs and sophisticated interactive capabilities (IGA, 2023).

4. The National Space Science Agency (NSSA) was established in 2014 to support the achievements of the sustainable development goals in Bahrain by adopting space sciences and satellite technologies. They run different projects related to the collection of agriculture data; however, data is not accessible publicly nor anonymously for research purposes. In 2019, the agency launched a lab for processing space data and images, which adopted two agriculture projects:
 - Monitoring the agriculture area (Green Cover) in a matter of size.
 - Using remote sensing for inventory and health observation of palm trees.
5. Educational organizations engaging indirectly with agriculture activities and generating agriculture data, such as the AGU, one of the oldest academic institutions in Bahrain. With the launch of its post-graduate program (Program of Desert Farming Techniques & Soilless Agriculture), it will become the country's first research institution scientifically engaged with agriculture data.
6. Individual data owners such as farmers, farm owners, and even agronomists or agriculture scientists have run their projects where they collect data either manually or using technology; however, the data they collect is neither stored efficiently nor available for sharing due to many limitations.

The agricultural sector severely lacks ICT tools on farms or institutions. Technology can play a vital role in converting barren land into agricultural land with fertile soil. Although Bahrain's environmental

conditions do not comply with many crop production requirements, ICTs can minimize such challenges.

2.6 Summary

The second chapter highlighted the development of the agricultural field and how ICTs have affected the field. The first section of the chapter presented the related literature discussing an overview of ICTs and their development in the agriculture sector. It disputed its influence on extension services, smart agriculture initiatives, the impact on PA practices, and how it works as an enabler to achieve sustainable development goals. The second section defines the Big Data concept and reviews its related technologies. The third section overviews the literature investigating the Big Data in agriculture from different aspects, including (analytics, frameworks, and applications) and associated challenges and opportunities. In the last section, the agricultural situation in the Kingdom was discussed in terms of production, development, and use of ICTs.

The review of the literature helped in our investigation related to the state-of-the-art ICT applications in the agriculture field and how Big Data can be utilized in agriculture to achieve sustainable practices.

Chapter Three

Research Methodology, Tools, and Technology

3.1. Overview

Chapter 3 describes and elucidates the array of tools and methodologies employed in this study. Given the multifaceted nature of the research, a diverse set of research approaches were implemented. The methodologies encompass a closed-ended questionnaire, on-site field visits, semi-structured interviews, and an extensive review of pertinent literature, all framed within a conceptual framework. The tools and technological resources utilized encompass wireless soil sensors, a multispectral camera, cloud services (Google Cloud Platform), a remote sensing platform, and Visual Studio Code as the primary development environment.

3.2. Research Methods

The term research methodology is a holistic approach that is chosen to show the way that the researcher achieved the goals of his research through methods that are categorized as either qualitative or quantitative, which can be in the form of case studies, interviews, questionnaires, surveys or any other information gathering technique (Kirsch & Sullivan, 1992). These methods entail a prescribed series of steps governing the collection of information and the conduct of data analysis, intending to achieve research objectives. Qualitative methods are the prevailing choice among researchers due to their potential to yield advantages and precise outcomes when correctly applied. However, their effectiveness may diminish when employed in unsuitable contexts (Newman & Ridenour, 1998).

Qualitative and quantitative data are the two main types of data utilized in research. Qualitative research is the method of observation to collect non-numerical data where there is no use of statistics. As Mehrad & Zangeneh (2019) stated, qualitative research is a more exploratory method utilized to attain an explanative motivation, reason, and holistic view. Various methods, such as group discussions, interviews and observations, come under qualitative research techniques. On the contrary, quantitative research is a method that solves issues through collecting numerical data, which is then transformed and processed into an understandable and valuable statistical output.

Furthermore, Mehrad & Zangeneh (2019) mention that the quantitative method is utilized to quantify behaviours, attitudes, opinions and others. Unlike the qualitative approach, the quantitative approach is used when a large sample population is in focus and when a large amount of data must be collected to extrapolate results in a measurable output. Moreover, both can be merged into single research by a specific methodology and strategy (Newman & Ridenour, 1998).

The study is multistage research encompassing different objectives with different natures, which demand the adoption of both quantitative and qualitative methods to achieve the study objectives and ensure data reliability as well as covering all related factors as possible. The study consists of four phases:

- Phase 1: to achieve the first objective, an exploratory method to gather data, summarize information and describe the characteristics of the agriculture sector in Bahrain is applied. A closed-ended questionnaire was distributed among Bahraini farmers and farm, field visits to multiple farms with different features, and online research of agriculture-related publications for Bahrain.

- Phase 2: semi-structured personal interviews were conducted with selected stakeholders to identify Bahrain's current agriculture data lifecycle. Based on the results acquired from phase one and the understanding of the Bahraini agriculture sector along a SWOT/TOWS analysis, the following factors will be identified: available data, supply/demand of the stakeholders, the importance of the data, and the challenges of data collection and accessibility. These factors shaped the agriculture data lifecycle in Bahrain and proposed an owner/user data matrix to clarify the stakeholders' roles and privileges.
- Phase 3: designing a conceptual agriculture Big Data framework for the Kingdom of Bahrain to visualize and identify different approaches and technologies of data collection, transformation, storage, and presentation integrating Big Data tools and cloud services. The developed framework will enable multi-access levels of the stored data for the different stakeholders and must be flexible and resilient for future customization.
- Phase 4: a simulation model was developed on the cloud based on the suggested framework to assess its feasibility and imitate a smart agriculture platform's different processes and functionalities.

3.2. Phase I: ICT Awareness, Readiness, & Adoption

This phase utilizes both quantitative and qualitative approaches. Quantitative was applied through a series of questions involved in a questionnaire along with sampling techniques and measurement criteria to investigate the respondent's awareness, readiness, acceptance, and adoption of ICT in agricultural development in the Kingdom of Bahrain. This method was selected for its flexibility in collecting large amounts of data using different platforms within a limited period (Mehrad &

Zangeneh, 2019). Moreover, this approach allows the adoption of numbers in statistical tests to elaborate the results and ensure the findings have a statistical relationship and, hence, obtain statistical significance.

Waples and Gaggiotti (2006) state that the population is characterized as the number of individuals living in a region within explicit limits with a place with similar species. The study population comprised the stakeholders that control/own the primary agriculture data sources (farms), including farmers, farm owners, and enterprise farm managers. The Kingdom of Bahrain has approximately 2,400 farms and an estimated total agricultural workforce of 9,120 (Gulf Daily News, 2021). The sample size must be defined and calculated to know the number of responses required for the survey (Singh & Masuku, 2014). The sample size was calculated using a confidential level of 95% and an interval of 5, estimated to be approximately 300. However, due to the Covid 19 situation and the difficulties in reaching such a segment of individuals or even agriculture businesses and farm owners, only 100 responses were collected.

The questionnaire was designed using Google Forms and distributed online through emails and WhatsApp. To cover more respondents, the researcher printed out the questionnaire and spread it physically beside using telephone calls with target stakeholders. It was distributed from November 2020 through January 2021. The questionnaire was designed in Arabic (the native language in Bahrain), consisting of 19 questions divided into three sections (Appendix A). The first section is the demographic section, which collects basic demographic information about gender, age, educational level, monthly income, farm size, farming methods, etc. The following section evaluates the overall ICT literacy and the awareness about ICTs in agriculture (hardware and software). The third section is directed to collect data to assess the innovation

willingness, adoption, and usage of ICTs in agriculture along with the inhibitors which act as barriers to utilizing ICT and calculating the impact after the utilization of ICT in agriculture.

The collected data has been registered with Google Forms, which allows viewing the data status and exporting it to a spreadsheet operated by Microsoft Excel. The data is cleaned and transformed using Excel's custom and built-in functions such as (IF, ISNUMBER, SEARCH). Then, the transformed data has been analyzed utilizing the Statistical Package for Social Sciences (SPSS) to validate the data and come up with the study results.

The field study approach was adopted to form a holistic ICT model in Bahrain's agriculture. Field research/study, sometimes called ethnographic research, which the author will later refer to as "field visits", is a research approach characterized by observation and data gathering within a real-world context (Ven & Poole, 2017). This methodology encompasses activities such as observing, interviewing, and engaging with individuals in their native surroundings, which could range from a workplace and community to a natural habitat.

Four exploratory field visits were conducted from November 2020 through December 2020 to enhance the survey's findings and explore the current situation regarding agriculture behavior, production, and how ICT is used to support agriculture on selected farms. The limited number of visits is due to many reasons, including Covid-19 pandemic, difficulty in reaching the farmers/farms' owners, and social limitations on information sharing. The field visits allow the researcher to monitor closely the different agriculture activities while asking questions at the same time. Moreover, it presents opinions about the inhibitors and opportunities facing the agriculture sector. Within the research context, the chosen farms are referred to as Farm1, Farm2, Farm3, and Farm4. Different

sizes, characteristics, and cultures were considered for the selection of the farms.

In addition, a SWOT-TOWS analysis has been conducted to analyse the readiness of the Kingdom of Bahrain for adopting ICT in agriculture and its strategic direction.

3.3. Phase II: Agriculture Data Lifecycle

The second research objective aims to investigate the supply/demand side of the agriculture data needs by identifying the available data, the different needs of each stakeholder, the difficulties and challenges of acquiring the agriculture data, and the importance of the data. This will be done toward building a user/owner structure for the Kingdom of Bahrain agriculture Big Data. A qualitative approach was applied using a semi-structured interview to achieve the objective. A total of eleven stakeholders were personally interviewed between April and September of 2022. The interview consists essentially of eighteen questions (Appendix B); however, depending on the interviewee and the interview context, additional questions (support questions) may be posed in some interviews to help the respondent perceive the answer to a predefined question. Conversely, the defined agriculture stakeholders were determined in accordance with the exploratory study carried out in phase one and pertinent scholarly works. Agriculture users were categorised as follows by Gill et al. (2017): agriculture officers, agriculture experts, and farmers. The authors established connections between the user categories and various domains of agriculture data, such as equipment, crop, soil, and meteorological information. To ensure a comprehensive representation of all stakeholders within the agriculture ecosystem in Bahrain, interviews were carried out with a deliberate selection process. The participants included agricultural merchants, governmental

organisations, farmers, farm owners, non-governmental organisations, agronomists, and experts. As shown in Figure 3.1.

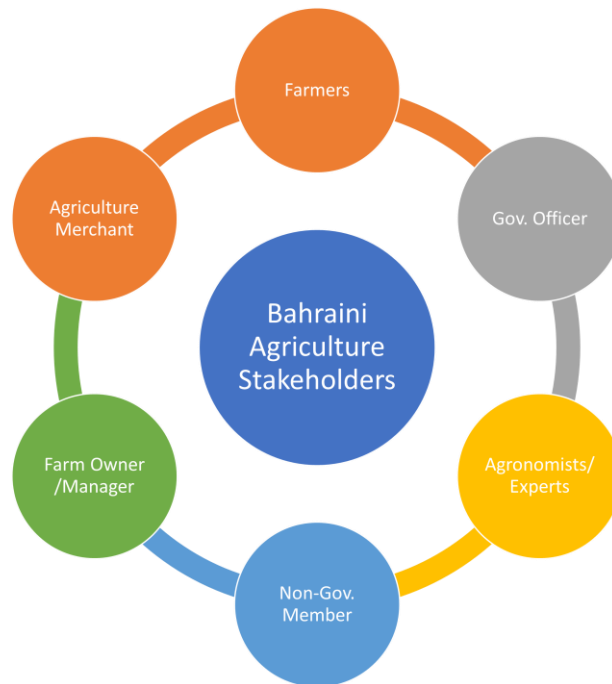


Figure (3.1): Agriculture Stakeholders

The agriculture ecosystem in Bahrain is unique, so you can find one person present as a farmer, farm tenant, and merchant at the same time. The mixture of roles in one stakeholder and the social limitation of sharing data formed the main selection criteria of the interviewees. Only one farmer agreed to be interviewed, and he owns two farms and works as a merchant for his agricultural products. Three agronomists, two members in two non-governmental organizations, one official officer in the MMAA, and four agriculture experts.

The interviews were conducted face-to-face; however, the interviewees were allowed to choose between physical visits or online platforms (MS Team, Zoom). All interviews were permitted to be recorded by the interviewees either through a sound recording device and the Mobile phone if physical, or through the online platform.

The questions focused on the data collection methods and tools, availability of the data, storage of the data, whether the data is shared, methods to access the data, and difficulties in obtaining the required data.

3.4. Phase III: Conceptual Framework

The results of the first two objectives were the foundation of the developed framework, and they contributed significantly towards understanding the available stakeholders, their roles, and the data they own/need. Moreover, the approaches conducted in the first two objects allowed the researcher to visualize the financial and educational status within the agriculture sector.

Frameworks for integrating Big Data in agriculture are available in a variety of scenarios that meet its needs. However, for Bahrain, the framework will be created with ICT knowledge and preparedness in mind, as well as social issues and limits, particularly among agriculture stakeholders, economic concerns, and environmental factors such as limited arable land. The majority of farms in Bahrain are characterised by their small size, typically spanning only a few hectares. As a result, the return on investment (ROI) associated with these farms is low (Shebeeb, 2006). Adopting complicated agricultural technology, such as IoT and cloud computing, may be an overwhelming financial barrier, particularly for small farms (Delgado et al., 2019). Furthermore, implementing IoT-based Big Data agriculture systems require major implementation factors to consider such as: 1- minimizing the cost of used technologies by building in-home sensors or purchasing complete modules; and 2- enabling cross-platform and standardisation of exchanged data due to differences between sensors and IoT gateways (Li & Niu, 2020). on the other hand, adopting and utilising ICT necessitates a certain degree of education and digital literacy. According to Bucci et al. (2019), farmers

with higher education levels (diploma, bachelor's degree) are more ready to digitise their businesses than farmers with lower education levels.

In the framework, Data and users' access levels will be considered for smooth interaction and improved privacy. Spanaki, Karafili, & Despoudi (2021) discussed using AI applications to control agriculture data sharing among different stakeholders. The authors suggested a scenario where the farmer is the data owner and collector, and other stakeholders are called data recipients. However, they also indicated that a third party (an official entity) processes the data on the cloud and is called a data processor.

The AgroBahrain framework is divided into two stages (ICT preparation & adoption and Data Collection, Storage, and Processing) aiming towards data collection, storage, transformation, and raw data delivery to the agriculture stakeholders. The first stage is responsible for studying, preparing, and establishing an ICT-ready agricultural environment. The second stage enables seamless collection of agriculture data from multiple sources using different approaches and managing the data storage and processing while allowing secure multidisciplinary access levels to the data under controlled privacy conditions.

Collected and processed data is expected to be integrated and delivered in different applications, such as:

- Change notifications for farmers and farm owners.
- Real time statistics on weather, soil, and crops conditions.
- Experts' knowledge.
- Running cold analytics by agronomists.
- Market information.

The framework is built upon the following data lifecycle (ETL/ELT):

1. **Ingestion:** extracting raw data from numerous sources, such as data streams from events, logs and IoT devices, historical data stores, and data from transactional applications.

2. **Storage:** storing the data in their raw format without or with minimal transformation in a flexible schema (schema-less) repository such as a data lake that is durable and easily accessible.
3. **Processing/Storage:** transforming the data (cleaning, aggregation, segregation) from its raw format into a format that enables usage and analysis. transformed data will then be loaded into schema-ready repositories such as data warehouses and data marts for analytics, visualization, and related operations.
4. **Analytics, Machine Learning, and AI Applications:** data analytics & visualizations, Machine learning, and AI applications are not considered in the simulation; however, a conceptual user-data interlinkage model is built to support data diffusion in future work.

3.5. Phase IV: The Simulation Model

Despite the term agriculture being broad and including many systems such as (crops, livestock, and forestry), the simulation focuses on agriculture data related to crop production, such as soil conditions, amount of inputs, and weather data; however, it is resilient to be scaled for other agriculture domains. The scarcity of livestock, the absence of other agriculture systems, and the differences in tools and data domains are among the primary reasons to focus on crops and their related devices, management approaches, and data domains.

A simulation model is adopted to imitate the processes of data ingestion, transformation, and storage for crop production. Farm owners and farmers are the most responsible for agricultural data among other stakeholders. Therefore, the researcher will simulate the agriculture data lifecycle from the perspective of a farm owner/farmer with respect to

different data domains, including (soil, weather, and farm data). The simulation scenario is suggested based on the developed framework and related literature. It includes the development of a Progressive Web App (PWA) to simplify communication with different agriculture stakeholders as well as the collection of data. The PWA is hosted on Google Firebase to streamline communication with other Google Cloud services. Various tools and technologies have been utilized (see Table 3.1); however, the primary workflow is implemented within the GCP.

Table (3.1): Tools and Technologies Used for the Simulation.

	Tool/Technology	Quantity	Description
1	Custom built PWA	1	The app enables data collection for demographic data of stakeholders, farms' data, adding new data sources, and it provides an Agriculture Enquiry Service (AES) functionality. On the other hand, it will provide easy access to agriculture data and enable the collection of legacy data such as Excel files and as a hub to import individual projects' data.
2	Pycno soil sensor	1	A field-based sensor to collect soil and weather data, and it can be installed in open and protective fields to collect and transmit data using cellular connectivity.
3	Remote sensing account	1	An account was created on the Earth Observation System (EOS) crop monitoring service to remotely collect data on the farm's weather, soil moisture, and NDVI index images. The satellite images are captured using SENTINEL-2 and Landsat 8 satellites.
4	Multi spectral Camera	1	Parrot SEQUOIA+ Multispectral camera to capture crop health by identifying (pests, diseases, and weed) and data on soil fertility by detecting nutrient deficiencies through leaves colors.
5	Google Cloud Platform (GCP)	1	A Google account were created to take advantage of cloud services such as (BigQuery, Cloud Storage, Cloud Functions).
6	Visual Studio Code	1	The main development environment for the simulation model.
7	Google Firebase	1	Used to host the PWA and the use of the NoSQL capabilities of the Cloud Firestore.

8	Front end development using Ionic framework	1	Ionic framework is well known development environment for developing hybrid mobile applications and PWAs.
9	Backend development using python and flask framework.	1	Python is an easy-to-use high-level programming language with strong support for data science. Flask is a python web framework adopted by google cloud services.

Finally, a custom Big Data management approach was applied and divided into three layers as follows (see Figure 3.2):

- Technical Metadata layer: a Meta catalog containing technical metadata about the data definition and structure such as (tables, fields and their description, and relation between tables).
- Business Metadata layer: business metadata such as data sources and description of data, supportive tools, data movements.
- Data layer: this layer is further divided into two structures: the first will contain sources-specific data generated, and the second is an aggregation of all sources based on data category. For example, weather data can be generated using soil sensors, remote sensing, and weather stations so that every source has a structured/unstructured storage mechanism. Then, all weather data will be transformed, aggregated, and stored in an operation-specific format.

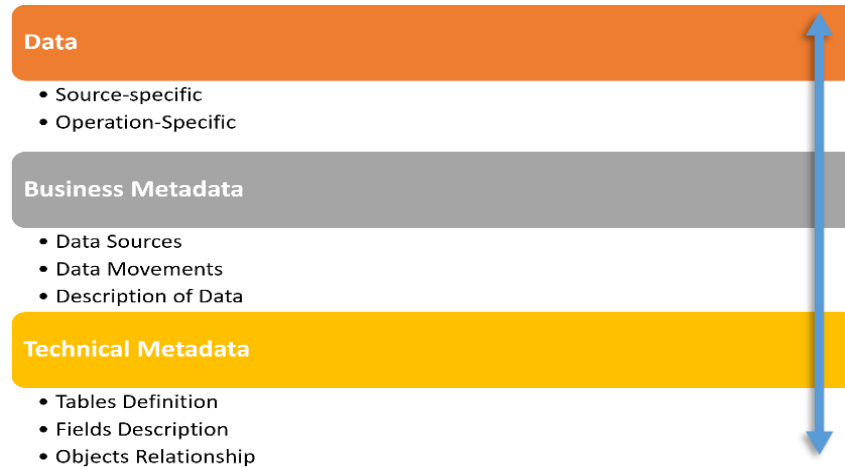


Figure (3.2): The conceptual agriculture Big Data management approach.

3.5.1. AgroBahrain Progressive Web App (PWA)

PWA operates as a web application but includes native mobile app capabilities such as push notifications, offline capability, and utilization of device hardware features. The concept of PWAs was introduced in 2015 by Google Chrome engineer Alex Russell and designer Frances Berriman, as highlighted by Steiner (2018). The origin of this concept may be attributed to the Firefox OS project that commenced in 2013. The project's objective was to enable online applications to function as native applications on mobile devices inside an open-source framework. According to Tandel and Jamadar (2018), although web apps are more easily accessible and discoverable compared to native apps, users typically spend more time on the latter. PWAs are being adopted at a fast pace because of their extensive functionalities and improved user experience. Prominent benefits include easy installation without the need for app stores, reduced file sizes compared to native apps, faster loading even in low network conditions, automatic updates, user engagement similar to native apps, and compatibility with different devices that support web browsers like Google Chrome, Safari, and Microsoft Edge.

PWAs benefit from many web technologies; however, two of these technologies are considered the heart of a PWA:

- a. The web app manifest, which is a simple JSON file responsible for identifying the app components such as icons, logos, and a splash screen.
- b. The service worker is the controller of the app, it is a script that runs on the background to manage features such as push notifications, app update, and enable the app to work offline by caching necessary data.

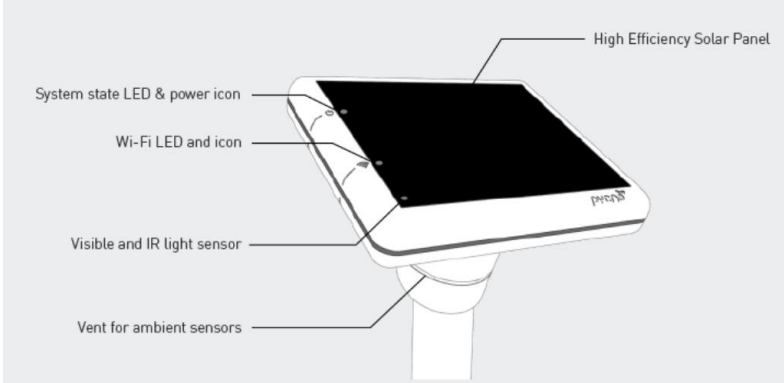
AgroBahrain PWA is a multi-user app coded once; however, it can run on both mobile and desktop devices to support different features for all agriculture stakeholders. The application enables agriculture stakeholders to communicate and explore agriculture data; however, the simulation will be limited to data collection, storage, and transformation features in addition to basic data visualization (for demonstration):

1. Basic profile data.
2. Farms' data including (size, location, name)
3. Crops data.
4. Crop Production.
5. Tools and technologies such as soil sensors.
6. Visualization of basic data and extracting custom data.

3.5.2. Pycno Soil Sensor

TERRA all-in-one soil and ambient sensor is a field-based sensor provided by Pycno company (<https://pycno.co/>) designed to collect soil and weather data, and it can be installed in open and protective fields.

Table (3.2): Technical Details – Pycno Soil Sensor

Device		
Sensor Length	30 cm/ 12 in	
Power source	High efficiency monocrystalline 2W Solar Panel, 3490mA/h Internal battery	
Connectivity	Cellular connection (4G)	
Data retrieval through:	a. API endpoints b. Online platform (limited)	
Data collection interval	the sensors combine multiple measurements every 15 minutes, and upload the measurements every 2 hours	
Data	Solar irradiance (IR + visible light LUX)	Accuracy: +/-10Lux, Resolution: 0.1 Lux
	Solar irradiance (Watts per m2)	Accuracy: 1 W/m2, Resolution: 1 W/m2
	Air Temperature	Range: -40~85°C – Accuracy: +/- 0.3°C, Resolution: +/-0.005°C
	Air Humidity	Range: 0~100%RH – Accuracy: +/- 2% , Resolution: 0.01%
	Soil Temperature	Range: -40~85°C – Accuracy: +/- 0.3°C, Resolution: +/-0.005°C
	Capacitive Soil Moisture	Resolution: 0.01mm, Accuracy +/- 1mm

3.5.3. Satellite Remote Sensing

Remote sensing using the EOS crop monitoring application. EOS data analytics is an online platform that uses the satellites' imaging capabilities to offer agriculture-related applications such as crop monitoring, forest monitoring, and the Land Viewer. For simulation

purposes, a paid subscription plan is selected for the crop monitoring platform with the following features:

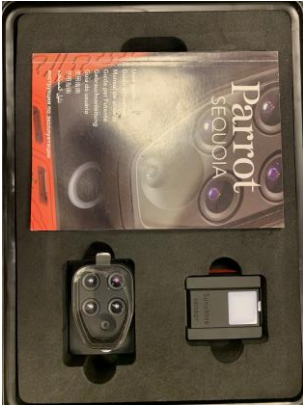
Table (3.3): Technical Details – EOS Crop Monitoring Platform

Subscription plan	Professional with coverage area of 10 hectares	
Platform	https://eos.com/	
Satellites (data sources)	a. Sentinel 2 – since 2015 b. Landsat 8 – since 2013 c. EOS SAT (not yet launched)	
Data retrieval through:	a. API endpoints b. Online interactive platform	
Data collection interval	Satellite images every 5 days.	
Data	Weather Forecast	a. Min/Max temperature 2m above the ground (K) b. Near-surface shortwave radiative flux (W/m^2) c. Precipitation (rain, snow) (kg/m^2) d. Snow thickness (m) e. Humidity at a height of 2m (%) f. Average wind (m/s) g. Cloud cover (%) h. Dew point temperature at a height of 2m (K) i. Min/Max temperature of the earth at a depth of 6 cm (K) j. Pressure (Pa) k. Wind direction (deg)
	Soil Moisture (not available in Bahrain)	a. Percentage of root zone soil moisture. b. Percentage of surface soil moisture.
	NDVI	Normalized Difference Vegetation Index
	NDMI	Normalized Difference Moisture Index
	MSAVI	Modified soil adjusted vegetation index
	NDRE	Normalized difference red edge index
	RECI	Red-Edge Chlorophyll Index

3.5.4. Multi Spectral Camera

Remote sensing using a drone-based Parrot SEQUOIA+ Multispectral camera, however and due to policies and regulations the camera was used manually (by hand) to capture images of crops with 5 different bands.

Table (3.4): Technical Details - Parrot SEQUOIA+ Multispectral camera

Device		
Sensors	1.2-megapixel monochrome sensors	• Green (550nm – 40nm bandwidth) • Red (660nm – 40nm bandwidth) • Red Edge (735nm – 10nm bandwidth) • Near Infrared (790nm – 40nm bandwidth)
	16-megapixel RGB sensor	
	GNSS/GPS	
	Magnetometer	
Power source	2.4 A USB battery	
Connectivity/ Data retrieval	a. USB ports b. Wi-Fi	
Storage	c. 64 GB internal storage. d. 32 GB external memory card.	
Data collection interval	the sensors combine multiple measurements every 15 minutes, and upload the measurements every 2 hours	
Data	Images in JPEG format	

3.5.5. Google Cloud Platform (GCP)

The decision to adopt GCP as a comprehensive tool for running an experimental simulation of an agriculture Big Data framework on the cloud was not influenced by comparing it to other cloud service providers. It offers a wide range of services that facilitate comprehensive data engineering and data analysis activities. In addition, it provides highly scalable and effective storage technologies for data lakes and warehouse operations (Bisong, 2019). Google provides customers with a complimentary cloud account for a duration of 90 days, along with a credit of 300 US dollars that may be used for all services. Some services provide a free tier with unlimited usage, while others, like Google Cloud SQL service, do not.

A full lifecycle of a farm owner/farmer account is implemented. It will be discussed in detail in chapter 5 to present the different pipelines and user interfaces used to complete the process of ETL/ELT on the data and the utilized data schema. Five GCP services were mainly adopted as depicted in Table 3.5.

Table (3.5): GCP services utilized for the simulation.

Service	Icon	Description
Google BigQuery		Cost-effective, serverless, cloud enterprise data warehouse for structured Big Data storage and analytics.
Cloud storage (Data Lake)		Infrastructure as a Service (IAAS): In GCP, Cloud Storage is a scalable and real-time storage repository that can handle both semi-structured and unstructured data with easy access operations.

Cloud SQL (MySQL)		Fully managed relational database service for MySQL for storing structured data.
Cloud Functions		Serverless Function as a Service (FAAS). Cloud-based functions to connect or extend cloud services.
Cloud Scheduler		fully managed enterprise-grade cron job scheduler to schedule pub/sub messaging services.

3.5.6. Implementation Environment

Due to limitations on funding and to simplify and speed up the development processes, different tools and technologies were adopted to develop and test the simulation model locally before running it on the cloud. For code editing, Visual Studio code version 1.42.1 is used. VS Code is a free source code editor created by Microsoft and was first released in 2015. It comes with built-in support for JavaScript, TypeScript and Node.js, as well as supporting remarkable extensions for other languages such as Python, C#, and PHP. XAMPP version 7.4.27 is used to provide a local MySQL database for creating primary structured entities. XAMPP is a free and open-source cross-platform web server solution developed by Apache Friends and released in 2002, and it supports mainly PHP, Maria DB, and Perl. The PWA user interface and features were developed using the ionic framework (HTML, TypeScript, and CSS) version 6.3.4. Ionic is an open-source software development kit (SDK) used for creating hybrid mobile applications and was released in 2013. The structure was modified to accommodate different user interface frameworks such as Angular, React, and Vue. The application utilises Capacitor and Cordova plugins to integrate native functionalities of mobile devices, such as the camera. Starting from version 4, it fully

enables the development and conversion of mobile apps into PWAs. The backend scripts were primarily developed using Python version 3.10.2, which is a high-level programming language that is easy to learn and supported by many libraries for building websites, task automation scripts, and data analysis. Flask, a micro web framework, offers the essential libraries and functions for web operations in Python web development. Table 3.6 is a comprehensive datasheet that includes information on all libraries, packages, and technologies.

Table (3.6): Technical Details – Development Environment

Tool/Library/Device	Version/Details
iMac late 2015	• Processor 2.8GHz Quad-Core intel core i5 • 8 GB DDR3 RAM • macOS Big Sur version 11.5.2
Lenovo ThinkPad laptop	• Processor 2.8GHz Quad-Core intel core i7 • 8 GB DDR3 RAM • Windows 10
Ionic	• Ionic CLI: 6.20.6 • Framework: 6.3.4 • Node JS: 16.13.2 • Npm: 8.1.2 • Capacitor: 4.4.0
VS Code extensions	• IntelliCode version 1.2.14 • Python IntelliSense 2020.3.71113
Main Python libraries and packages	• Flask version 2.2.2 • Faker version 15.3.4: a package that generates fake data in python. • Pandas version 1.5.2: powerful library for data analysis, time series, and statistics.

Finally, Google Firebase hosting services are adopted to host the PWA for free (limited usage) and provide NoSQL capabilities through cloud Firestore. Moreover, it seamlessly integrates with Google Cloud Functions and Cloud Storage.

3.6. Summary

This chapter details the methodological approach employed in this study and delineates the constituent elements of the developed framework intended for simulation as proof of concept. The research endeavor unfolds across four primary phases established to accomplish the outlined research objectives.

The initial phase encompassed a dual-method approach, integrating both quantitative measures through a questionnaire and qualitative methods via field visits. The primary aim was to comprehensively explore and gain an in-depth understanding of the agricultural landscape in Bahrain, assess the level of ICT awareness, and examine the adoption patterns among stakeholders within this sector.

Subsequently, the second phase involved the utilization of a qualitative approach, specifically semi-structured interviews, in conjunction with the insights and discoveries drawn from the initial phase. This phase aimed to elucidate and visualize the lifecycle of agricultural data within the Bahraini context.

Drawing upon the comprehensive findings of both phases, a conceptual framework was formulated to encapsulate the dimensions of agriculture Big Data specific to the Bahraini landscape.

The chapter also expounds on the technical components utilized in this study, encompassing the detailed descriptions of the devices employed, such as the soil sensor and multispectral camera. Additionally, the chapter elaborates on the online platform, namely EOS, utilized for remote sensing, the cloud services provided by GCP, and the specific development environment adopted for the simulation model.

Chapter Four

Stakeholders Awareness, Adoption, and Data Lifecycle (Results & Discussion)

4.1. Overview

This chapter investigates the readiness and acceptance level of integrating ICT in Bahrain's agriculture sector. The focus will be on analyzing the data life cycle as perceived by stakeholders involved in this area. Hence, the upcoming chapter will thoroughly examine and discuss the results obtained from the research methods used to accomplish the first two objectives. This examination will involve the use of SWOT-TOWS analysis, enabling an evaluation of both internal and external variables that affect the use of ICT in the agriculture industry.

The first objective was accomplished through the utilization of a dual methodology encompassing quantitative (utilizing a questionnaire) and qualitative (via field visits) approaches. This dual approach aimed to explore and comprehend Bahrain's agricultural landscape comprehensively. A qualitative approach was adopted for the second objective, employing semi-structured interviews in conjunction with insights garnered from the initial objective to represent the agricultural data life cycle visually.

4.2. ICT Awareness, Readiness, & Adoption

The following are the results and discussion of the quantitative method applied through a series of questions involved in a questionnaire to investigate the respondent's awareness, readiness, acceptance, and adoption of ICT in agricultural development in the Kingdom of Bahrain.

i. Demographic profile of respondents

Table 4.1 shows that most of the respondents have minimal land with a size of fewer than 0.5 hectares. Forty per cent of the participants have an estimated whole cultivation land and are adopting traditional farming (75.7%) for planting vegetables (83.8%) and fruits (56.8%). The result is unsurprising as the agricultural areas in the Kingdom of Bahrain are limited, and they have decreased by 0.25% since 1971 (Alayam, 2019).

As part of their agricultural behaviour, respondents use between 0 and 25 litres of pesticides yearly and almost use fertilizers for more than 76 tons—some of them use fertilizer for 25 tons or less annually, as shown in the table.

Table (4.1): Description of the farmers and agriculture business

Land		Cultivation land		Cultivation methods	
< 0.5	62.1%	< 50%	37.8%	A greenhouse	29.7%
0.5 - 1	27.0%	50% - 75%	21.6%	Cultivation without soil	16.2%
>1	10.8%	100%	40.5%	Traditional farming	75.7%
				Others	24.3%
Crops		Pesticides/year (liter)		Fertilizers/year (ton)	
Dates	35.1%	0-25	56.8%	<= 25	32.4%
Aromatic	40.5%	26-50	16.2%	26-50	18.9%
Vegetables	83.8%	51-75	8.1%	51-75	5.4%
Fruits	56.8%	76-100	18.9%	76-100	32.4%
Other	21.6%	> 100	0.0%	> 100	10.8%

ii. Digital Readiness

Digital readiness is about whether people have the skills to use IT and digital literacy tools to help people determine whether the online information they access is trustworthy. Digital readiness is essential in adopting IT in agriculture as it emphasizes the significant gap between the current state of farmers and the ideal farmer 4.0 (Klerkx, Jakku, & Labarthe, 2019). Therefore, digital readiness for farmers and individuals investing in the agriculture business was explored in terms of using smart

devices, social media, online services, and the Internet. Table 4.2 and Figure 4.1 present information on the digital readiness of farmers and agriculture businesses. The results revealed that 95% are using a smartphone, 97% are using WhatsApp, and 78% are using Instagram. Moreover, over 75% of the farmers use online services and can browse the Internet easily. However, only 50% of the respondents know the differences between hardware and software and can fairly use the computer devices.

Table (4.2): Digital readiness of the farmers

I use the following smart devices		Social Media used by Farmers	
Smart Phone	95%	WhatsApp	97.30%
Laptop	32%	Instagram	78.40%
Desktop	24%	Snap Chat	29.70%
Tablet	22%	Facebook	18.40%
I do not use smart devices	3%	Twitter	16.20%
		Others	8.10%
		I do not use social media Apps	2.70%

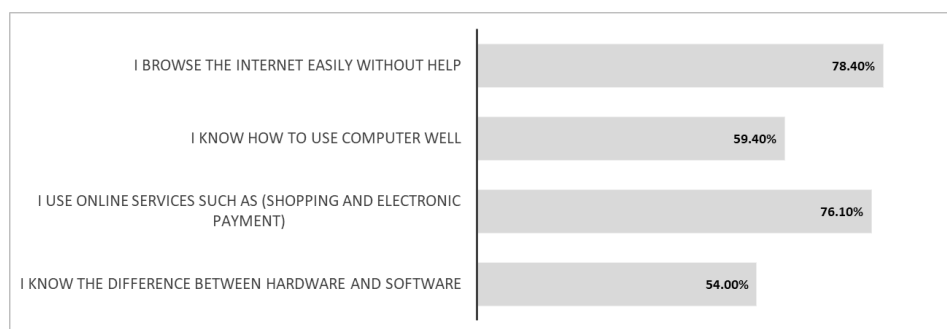


Figure (4.1): Digital readiness of the farmers

Results confirmed that mixed generations in the agricultural society live different lifestyles and mindsets in the Kingdom of Bahrain. Most of the community consists of regular farmers who are low-educated and illiterate and are using only WhatsApp for social communications and conducting their services physically (not using online services). Although they use smart devices, some online services, and social media, they have

yet to adopt sophisticated devices and applications such as crop sensing, smart irrigation systems, IoT, Augmented Reality, and AI. However, some farmers successfully adopt digital technology, which can be called the “Young Smart Farmers”. They are a group of farmers who use smart devices and social media intensively in their daily lives and adopt some sophisticated ICT in their agriculture to enhance their productivity (Sayruamyat & Nadee, 2020).

The findings revealed a reasonable awareness and low usage of ICT in agriculture among farmers and agriculture businesses in Bahrain. For example, results in Figure 4.2 show that there is little awareness of ICT use in agriculture as almost (40%) or less are aware of agriculture technologies such as databases, soil sensors, agriculture remote sensors or expressions such as IoT or PA. Although terms such as soil sensors, agriculture remote sensing, IoT, and Big Data in agriculture are fundamental tools and concepts in ICT in agriculture and digital innovation in agriculture, most farmers and agriculture business respondents are unaware of them.

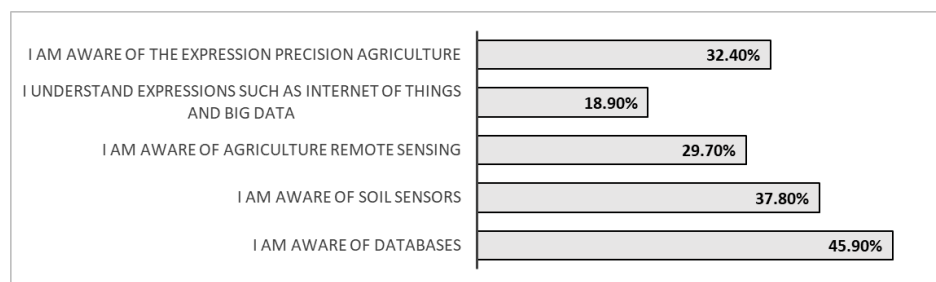


Figure (4.2): Awareness of farmers on the ICT used in agriculture.

iii. Adoption and Usage

The low awareness of ICT in agriculture has been reflected in the usage of these technologies, as shown in Table 4.3. The results show that almost less than 40% of the respondents have an internet subscription or feel they need the internet to manage their farm, using remote or soil sensors,

special software, or even applications to manage their farm. Moreover, they never use special information systems to manage their farm or consider it an important tool to advance their business. However, more than 75% of the respondents are curious about modern farming technologies, often search about them, and like to try contemporary technologies to develop their farms, as shown in Table 4.3.

Table (4.3): Usage of ICT in Agriculture by farmers.

I am very curious about how modern farming technologies work	72.90%	I love trying new methods to develop my farm	86.50%
I use computers to manage my farm business	37.80%	I often look for information about new products and technologies in the agricultural field	75.70%
I use special software and applications to manage my farm	35.10%	I have access to the resources I will need to use precision agriculture technologies	40.50%
I need internet to manage my farm	32.40%	I have the resources, the opportunities, and the knowledge to use the technologies for precision agriculture	37.80%
I use soil sensors in my farm	27.00%	I use remote sensing for my farm	24.30%
There is a dedicated internet subscription for the farm	24.30%		

Subashini & Fernando (2017) state that lack of knowledge in ICT and difficulties in using ICT are the main obstacles to adopting it in agriculture. Table (4.4) summarizes other challenges farmers and agriculture businesses face in adopting ICT in agriculture. Most of the respondents are facing difficulties in the supportive ICT infrastructure (43.0%), followed by a lack of knowledge and experience in using ICT (39.10%) and a lack of training and support (35.1%). Moreover, there are additional costs for operating and adopting digital technology, i.e., internet charges, mobile network charges, etc. These costs were automatically added to the farm management cost. As mentioned earlier,

most Bahraini farmers are considered family farming, not business farming entities. Therefore, the perspective on cost management is different. For example, they do not separate household expenses from the expenses in farming. As a result, farmers may refuse to use or adopt technologies since they realize the additional costs they must pay (Sayruamyat & Nadee, 2020). Government policies and regulations in Bahrain focus on enhancing agricultural productivity. They provide technical and financial support for individuals and farmers, including ICT services, equipment, and Agricultural loans (Alwatan, ٢٠٢٠). For instance, installing WSNs and developing smartphone applications by the government are vital and strategic actions for rapid growth and easy access to information for farmers and planters (Barh & Balakrishnan, 2018). Government support is essential for encouraging individuals' involvement in agriculture and adopting ICT for their productivity; however, most farmers and business agriculture need to be aware of government support and efforts to facilitate their business and enhance their productivity.

Regarding the lack of knowledge and training support, the agricultural sector in the Kingdom of Bahrain receives special attention from the competent authorities by preparing specialized courses to develop human cadres working in the farming industry. NIAD, for instance, intensifies training courses and advanced academic programs to attract the most significant number of agriculture enthusiasts (Alwatan, 2020). NIAD also provides educational and awareness programs in addition to the annual Bahrain International Garden Show (BIGS), which provides indirect awareness programs targeting farmers, agriculture businesses, and investors. Moreover, most higher education institutions offer different agriculture academic programs. For example, the University of Bahrain provides a bachelor-level program in landscaping and plans to establish a

college for Agriculture to build national agriculture capacity. The AGU offers a master's program in Desert Farming Techniques and Soilless Agriculture. The government offers vital efforts in the Kingdom of Bahrain to facilitate the challenges and difficulties in adopting ICT in agriculture and encourage individuals and businesses to take further steps in benefiting from the available agriculture technologies and ICT.

Table (4.4): Perception of the difficulties in adopting ICT in agriculture.

Perception on the difficulties in adopting ICT in agriculture	%
lack of supportive ICT infrastructure	43.00%
Lack of knowledge and experience in using ICT in agriculture	39.10%
Lack of training and support	35.10%
High cost for adopting ICT in agriculture	33.80%
lack of technical support	23.80%
Erratic power supply and fluctuating networks	13.90%

Table (4.5) finally identifies the respondents' perception of the impact of adopting ICT in agriculture. More than 50% of the respondents see that ICT can help save time and effort and increase production and income. Furthermore, there is a good perception of the effect of ICT in providing necessary information for better decisions and best resource utilization (67.5%). However, respondents disagree on the impact of precision farming on reducing the environmental factors and disasters on their farms (40%), and almost 40% of them perceived no benefits in using ICT in agriculture. The low perception of the positive impact of ICT on enhancing and improving agriculture productivity is due to the lack of knowledge and awareness and the low level of adoption of ICT in agriculture.

Table (4.5): The impact of adopting ICT in agriculture.

Perception on the benefits of using ICT in agriculture	%
ICT helping me in doing my work (planting, irrigating) more efficient	32.5
ICT helps in increasing my production and income	46.7
ICT helps in saving time and efforts	56.8
ICT helps in improving the quality of my products	31.8

Precision farming techniques can provide the necessary information to make better decisions that contribute to reducing resource consumption (such as water, pesticides, and fertilizers).	67.5
The data provided by precision farming techniques will reduce the environmental impact on my agricultural activities	40.5
None of the above	40.1

iv. Field Visits

The survey points out the lack of awareness and digital readiness among farmers; however, some of the young farmers indicated the opposite. Furthermore, the size of the farm significantly influences the extent to which digital technologies are used, in conjunction with the educational level of the farmer. Consequently, further investigation is conducted employing a qualitative methodology that relies on observation and monitoring to highlight the findings of the survey and the problem being researched. Four investigative field visits were undertaken to augment the survey's results and investigate the present state of agricultural practices, productivity, and the use of ICT in supporting agriculture on selected farms. The farm visits were conducted during the COVID-19 epidemic and the country's lockdown, making reaching farms in any location challenging.

Nevertheless, the assistance from NIAD has streamlined the process of reaching the designated farms and offered a superb chance to obtain information directly from the proprietors. Each visit has required around 4 to 6 hours. In addition, a short interview was conducted with the farm's owner or manager during the visit. The visits were analyzed based on the description of the farm and the adoption and awareness of ICT in agriculture, as shown in Table (4.6).

Table (4.6): Field visits analysis

Interviewer	Description of the farm	ICT adoption in agriculture	Awareness on ICT adoption in agriculture
Farm1	15000 hectares (ha) of land divided into 12000 ha cultivated using traditional agriculture methods and 3000 ha cultivated using Aquaponics agriculture. Aquaponics is a circular economy solution that is based on two basic systems: Aquaculture and hydroponics. The solution is built upon recirculating the water between the fish tanks and plants' growing area, where plants will convert the toxic residues of fish (ammonia and nitrates) into valuable nutrients and purify the water on its way back to fish tanks.	Within the 12000 ha, farm1 uses plain traditional farming, no automated practices nor technology usage. Moreover, the authors noticed that no sensors were attached to the fish tanks. The data were collected manually 2 to 3 times a day; however, there were temperature and humidity sensors on the Aquaponics planting area. The data gathered manually or through the sensors were not kept for future referencing nor analyses.	They believe using automation for irrigation and sensors to control the soil quality, and input are unnecessarily redundant costs, indicating a lack of knowledge of the technology advantages.
Farm 2	A small protective cultivation area (1 ha) containing a mixture of fruits, vegetables, and ornamental crops. The farm depends on the hydroponic method that uses the automatic air-conditioning system for temperature control.	Automatic Air-Conditioning System.	The owner of the farm is an average educated person that has a little knowledge about ICTs related to hydroponic farming. However, he expressed his willingness to experiment with these technologies in line with the needs of his farm.

	The owner of the farm mentioned the main challenge they face, which is the shortage of fertilizers. Farm 2 depends heavily on fertilizers since they do not use soil; this type of agriculture will be more sustainable when the inputs are controlled efficiently.		
Farm 3 and Farm 4	small to medium-size farms in Bahrain 0.5 ha and 2 ha, respectively. The farms are based on traditional soil farming, and both divided into protective and open land cultivation. The shortage of fertilizers and pesticides reported as a common challenge between both farms.	None (pure traditional farming)	The owners of the farmers' education are between elementary and intermediary, and they have no knowledge about ICTs in agriculture. The owner of farm 4 stated that he does not trust technology, and he runs his farm based on experience and practice.

Although few field visits were explored, they confirmed the survey's findings (the primary method for the first objective). The analysis of the field visits confirmed both low awareness, usage, and adoption of ICT in agriculture among the farms' owners and businesses. The results demonstrated a clear difference between younger, well-educated farmers and elder, less-educated farmers. The former are more knowledgeable about ICT in agriculture, including technologies such as smart irrigation systems and soil sensors, and they show enthusiasm towards adoption and usage. Moreover, the combination of young, well-educated farmers,

strong financial support, and large farm sizes is a significant driver to adopt digital farm technologies such as IoT sensors.

On the other hand, older and less educated farmers disagree with the adoption of technology and believe more in inherited experience. They lack awareness about the advantages of digital farm technologies such as UAVs and remote sensing, and they believe that farming should be managed through instincts and life experience. Moreover, they were less interested in ICT adoption even if they had financial support and technical capability (Al-Ammary & Ghanem, 2023).

4.3. Agriculture Data Lifecycle

The second research objective aimed to investigate the supply/demand side of the agriculture data by identifying the available data, the different stakeholders and their data requirements, the difficulties and challenges of acquiring the agriculture data, and the significance of the data. Moreover, it studied the available agriculture data lifecycle with respect to each stakeholder by identifying collection tools and methods, storage mechanisms, and whether the data is accessible by other stakeholders and how. This is done toward building a user/owner structure for the Kingdom of Bahrain agriculture Big Data. A qualitative approach was applied using a semi-structured interview to achieve the objective. Between April 2022 and September 2022, 11 personal interviews were conducted with different agriculture stakeholders in Bahrain; however, due to technical difficulties in recording the interviews, only ten interviews were considered. During a few interviews, some support questions are generated and added to the predefined questions based on the interviewee and the interview context.

The interviewed stakeholders include farmers, farm owners or managers, agronomists, government officials, non-government officials, and third-

party personnel involved in agricultural activities who have knowledge and experience in the field and will be referred to as agriculture experts. Every interviewee is assigned a code to be used in the context that will only refer to their role as a stakeholder while maintaining personal data privacy (Table 4.7).

Table (4.7): Interviews and the Interviewees details

Serial	Interviewee Code	Description	Interview date/time	Interview Method	Notes
1	Int_1	an agriculture expert and an academic. Participates in many official and non-official agriculture activities.	10-April-2022	Online through Microsoft Teams	
2	Int_2	An agronomist and a senior officer in the MMAA.	14-April-2022 @09:00	Face to Face (visiting the interviewee office)	
3	Int_3	an agriculture expert and an academic. Participates in many official and non-official agriculture activities	24-April-2022	Online through Microsoft Teams	
4	Int_4	Senior member in a non-profit organization that cooperates with official and non-official organizations to improve the agriculture sector in Bahrain	11-May-2022	Online (Zoom)	Excluded for technical issues in recording
5	Int_5	Traditional farmer and owns two farms in two different locations in Bahrain. Depends mainly on traditional farming and inherited experience	6-June-2022	Face to Face (farm visit)	
6	Int_6	An agronomist and academic specialised in agriculture resources management.	6-Sep-2022 @13:30	Face to Face (visiting the interviewee office)	
7	Int_7	An agronomist and	8-SEP-	Face to	

		academic specialised in plant physiology.	2022 @08:30	Face (visiting the interviewee office)	
8	Int_8	An academic and senior government official deals with technologies related to agriculture data.	18-Sep-2022 @16:00	Face to Face (visiting the interviewee office)	
9	Int_9	Retired official government and an agronomist specialised in horticulture	20-Sep-2022 @09:30	Online (Zoom)	
10	Int_10	Senior member in an agriculture association (non-governmental official).	28-Sep-2022 @10:00	Online (Zoom)	
11	Int_11	An agricultural expert manages a large enterprise farm. The farm utilizes smart agriculture methods including (sensors, computers, and database)	4-April-2022 @10:00	Face to Face (farm visit)	

The structured interview is divided into three sections. The first section contains questions about the availability of the data through a set of general questions, such as the number of farms and the percentage of imported crops yearly. The second section provides information on the various ICTs utilized for data exchange among agriculture stakeholders such as farmers, business owners, and researchers. It covers topics like the collection, storage, and delivery of agriculture data, the preparation, storage, and delivery of extension services, and the different ICT platforms that facilitate effective interaction between agriculture stakeholders, including forums, expert systems, and crop-sharing systems. In the last section, the questions concentrate on agriculture data and technologies that extract and benefit from these data, including the storage mechanisms used. Additional questions were added based on the

interviewee's job role and the work organization. For instance, farmers and official agents were asked about the data collected from the weather stations available in A'Ali and Budaya.

Table (4.8): Interview Questions (Availability of data).

	Q1	Q2	Q3	Q4	Q5
	Number of Farms (individuals/ corporates)	Agriculture Stakeholders	Yearly Imported Crops	Average/Minimum/Maximum Farm area	Number of open farms/protected farms
Int_1	Corporates may be 16 and individuals may be double the corporates	NIAD, Tamkeen	Most of the crops are imported. No actual percentage given	No data	No data
Int_2	Corporates around 5 or less. Individuals more than 200	The Farmer, MMAA, NIAD, farmers society, Agriculture engineers' association	More than 90%	Average: 2.5 ha, Min: 0.5 ha, Max: 15 ha	No data
Int_3	Corporates around 5	NIAD, MMAA, Ministry of Commerce	More than 90%	Min: 20 by 20 meters, no data	No data
Int_5	Corporates 4	The farmer, NIAD, MMAA, farmers	No data	Avg: 2 ha, Min: 0.5 ha, Max: 15 ha	No data

		society			
Int_6	Total of more than 140	The Farmer, MMAA	More than 80%	No data	No data
Int_7	No data	The farmer, NIAD	More than 80%	No data	No data
Int_8	No data	NIAD	No data	No data	No data
Int_9	No data	No data	More than 90%	No data	No data
Int_10	Total more than 600	The Farmer, MMAA, NIAD, farmers society, Agriculture engineers' association	More than 90%	Avg: 0.5 ha	Open farms less than 10% of total farms.
Int_11	Corporates between 4 and 5	MMAA, Farmer, Ministries of (Health, Commerce and industry, and Interior)	More than 90%	No Data	Almost all farms are mixed. No actual statistics.

Almost none of the interviewees has the exact number of total farms in Bahrain (Table 4.8); however, most of them agree that the known large corporate farms that produce crops for the market are between 4 and 5 farms. As for farms operated by individuals or as a total number of farms, only three interviewees gave disparate estimations of 140 (Int_6), 200+ (Int_2), and 600+ (Int_10) farms with an average size of 1 – 2 hectares. Moreover, almost all interviewees agree that Bahrain depends heavily on imported crops, with a yearly import percentage that exceeds 90%.

According to the interviewees, this happened due to many reasons, including the limited cultivation area, the scarcity of water resources, limited farming seasons due to difficult weather conditions, and, according to Int_11, the harsh soil structure restricts the farming of many products such as potatoes.

Identifying primary stakeholders is necessary to understand the agriculture data lifecycle and better serve each stakeholder with the required data. There was a consensus among the interviewees that the MMAA presents the main stakeholders in the sector as the only official organisation, NIAD, and the farmers are represented through the Farmers' Society. Interviewees (Int_2, Int_5, Int10) referred to other non-governmental organisations such as the Farmers' Society and Agriculture Engineers' Society. Moreover, Int_3 and Int_11 suggested additional stakeholders such as Ministry of Commerce, Ministry of Health, and Ministry of Interior. However, none of the interviewees, not even the academics, mentioned scientific and academic institutions as major partners in the agricultural field.

Table (4.9): Interview Questions (ICTs used to exchange agriculture data between stakeholders).

	Q6	Q7	Q8	Q9	Q10
	Type of Extension Services Provided	How the Services are Delivered	Availability of Agriculture Knowledge Networks	If yes, what are they	Participation on external agriculture platforms
Int_1	I don't know	I don't know	yes	Private WhatsApp groups between farmers	I don't know

Int_2	Delivering updated farming practices such as hydroponics, use of solar energy	Physical visits and workshops	No	I don't know	ICARDA (CGIAR) through the peninsula project
Int_3	Farmers market, garden show, training by NIAD	Physical workshops, posting information through NIAD website and Instagram account	Yes	Private accounts and groups on Instagram and WhatsApp	None
Int_5	Advice on pesticides, fertilizers, etc..	Physical visits through official agents	Yes	Private WhatsApp groups between farmers	I don't know
Int_6	Guidance about pests and fertilizers	On site visits and MMAA offices visit	Yes	Private WhatsApp groups	As an agronomist I participate on CGIAR, FAO, ICARDA for scientific research participation
Int_7	I don't know	I don't know	Up to my knowledge none	I don't know	International conferences, ICARDA

Int_8	I don't know	I don't know	Yes	The web page (agro.bh) by NIAD	I don't know
Int_9	Regular visits to farms and farmers guidance	Physical visits through official agents and doesn't exist anymore	Yes	YouTube channel by NIAD, personal Instagram and WhatsApp groups	I don't know
Int_10	There is but not enough	A national TV program called (Alsedrah)	Yes	The web page (agro.bh) by NIAD, private WhatsApp groups	I don't know
Int_11	I don't know	I don't know	We have experts but no networks	I don't know	I don't know

Despite the importance of agricultural extension services in transferring modern technologies to farmers and guiding and training them on their uses, almost 50% of the interviewees did not know the expression or knew it through the explanation of its content (Table 4.9). Extension services can be either through agriculture agents that provide training, knowledge transfer, and expert advice to farmers or as facilitation to connect farmers with different agriculture stakeholders, such as the market and the government, through mobile apps or web platforms (FAO,

2017). According to the interviewees, agricultural extension services in Bahrain were limited to personal visits by advisors from the ministry to the farms based on a prior request from the person in charge of the farm to perform specific tasks such as taking samples of crops and soil to help eliminate pests and diseases that affect crops, to distribute new seeds, or to promote new initiatives such as hydroponics. Moreover, in cooperation with NIAD, the national TV presents a weekly program that provides general guidance and advice to Bahraini farmers. According to Int_3, there are personal initiatives with the support from NIAD carried out by agriculture experts and agronomists to hold workshops and events for farmers. However, Bahrain lacks usage of ICTs in delivering agriculture extension services and as per Int_2, Int_9, and Int_10, unfortunately, due to the shortage of employees in the ministry, many of the available services are at risk of extinction.

According to most of the interviewees, Bahrain contains many agriculture experts, and basically, they refer to the farmers only and no other stakeholders such as the agronomists. This results from the presence of some social challenges in sharing information with others, which led to the absence of agriculture knowledge/experts' local networks except for a few private WhatsApp and Instagram groups among farmers. Moreover, Int_8, Int_9, and Int_10 identify the web page recently launched by NIAD (agro.bh) as a knowledge network that offers valuable information about the agriculture sector in Bahrain. On the other hand, external platforms that support agriculture and food security data, such as FAO and the Consultative Group on International Agricultural Research (CGIAR), Bahrain is still in the early stages of participation with a few projects, such as the peninsula project by the International Centre for Agriculture Research in the Dry Areas (ICARDA). The project was launched in 2007 with three main objectives: (1) rangelands, forage and

livestock; (2) protected agriculture; and (3) water resources management. It covers the Arabian Peninsula countries, Including Bahrain, Kuwait, Qatar, Saudi Arabia, the Sultanate of Oman, the United Arab Emirates, and the Republic of Yemen. The project is now in its fifth phase, which aims to introduce emergent technologies to reduce climate change effects. All project data is hosted by the agriculture Big Data platform (CGIAR) (ICARDA, 2007).

Table (4.10): Interview Questions (Collection of Agriculture Data)

	Q11	Q12	Q13	Q14
	Internal/external projects to collect agriculture data	Type of collected data	Collection Tools and Methods	Collection Intervals
Int_1	I think there is between Bahrain and Abu Dhabi	Water, soil, fertilizers, crop types	Manual observations and questionnaires	I think weekly bases
Int_2	The statistical project with FAO	Production data, farms data (area, farming methods, etc.), pests, soil, water data	Site visits and interviews	Almost no data being collected for more than 10 years
Int_3	Not up to my knowledge	Production and fertilizers data	I don't know	I don't know
Int_5	I don't know	Water, crops, fertilizers, and pesticides	Official agents do farm visits and ask questions	Not anymore for long time
Int_6	The NSSA project for palm	Soil data, quantifying	I use soil sensors and	Project based

	and green index cover in Bahrain	palm trees, water quality, and weather data	weather stations. As for the NSSA project remote sensing was used.	
Int_7	I don't know	Soil, palm, and crops data	I visit the farm and collect the data using my tools	Per project or research
Int_8	Palm trees health, water stress, trees classification	Number of palm trees, vegetation index (NDVI), GHGs emissions, surveying the mangrove area	Remote sensing (satellite imaging and spectral analysis), ARC GIS	Project based
Int_9	I don't know	Soil and water data	Site visits and manual surveys	Project based
Int_10	I don't know may be the NSSA have	Farms area, crop types, water data	Using manual and traditional tools	Previously agriculture data was collected almost yearly
Int_11	Yes, we collect data in our farm	Temp, Soil AC, wind, humidity, production data	Using different types of sensors manually and with IoT.	Daily and monthly

Questions 11 to 18 (Tables 4.10 and 4.11) Some interviewees answered based on what they actually do. Others answered based on what the official entities do.

The interviewees point out that public agriculture data in Bahrain is primarily outdated or does not exist. Data collection initiatives and projects are currently limited and exclusive to a few data requirements. According to Int_2 (senior government officer), there are no automated mechanisms or procedures for collecting and storing agriculture data; he even stated that no data has been collected for more than ten years, and Int_10 confirmed that as well. Although there is a project with the FAO that aims to collect and store agriculture data digitally in Bahrain, the tools and technologies have not yet been identified. Int_8 said that they collect agriculture data based on customers' requests and for paid amounts, and then the customer will receive the collected data on a storage media like a CD. Agronomists such as Int_6, Int_7, and Int_9 collect their data per research or project by physically visiting target farms. Int_6 and Int_7 stated that they sometimes face difficulties in acquiring the data due to a lack of cooperation.

As per Int_6, difficulties such as missing the appropriate data collection and storage technology are major challenges in advancing agriculture research. Moreover, he stated that Bahrain has a lot of data; however, it needs to be fed into an information system and become accessible through some portal to enable easy data exchange and to benefit scientists and farmers. According to Int_7, Int_2, Int_9, and Int_1, data such as (soil, water, and crops) are usually collected and transferred manually for testing and analytics. On the other hand, Int_11 stated that the farm he runs contains many technologies, including wireless sensors to collect, transfer, and store data automatically. However, the data are not utilised efficiently for analytics and future predictions.

Moreover, the farm's data is not shared with others, even anonymously, due to the lack of incentives and the low financial outcome of the business. Most interviewees agreed that the only national initiative currently available to collect and display agriculture data publicly is a project sponsored by NIAD that uses a static website to display the data (Agro.bh). Other projects may exist; however, either there is no news about them (private), or no one knows about them. For instance, the weather stations in Hoorat A'Ali and Budayi were installed in 2019. However, interviewees Int_5 and Int_10 never heard about them, so Interviewee Int_2 was asked if they were still functioning, and he said yes; however, the data is uploaded directly to a web server accessed only through the ministry, and there are no ICT means to deliver the collected data to interested parties such as the farmers. The interviewees' major challenges in data collection and access are social boundaries, shortage of personnel, and lack of projects directed by the government. Many of the interviewees (Int_1, Int_2, Int_6, Int_7, Int_10) believe that it is mainly the MMAA's responsibility to initiate data collection projects, assign dedicated personnel, and then share the collected data under their governance to mitigate the current challenges. Finally, a question is used to evaluate the interviewees' knowledge about agriculture data-related technologies such as (soil sensors and remote sensing) and concepts such as precision agriculture and Big Data. Only 50% of the interviewees (Int_2, Int_3, Int_6, Int_8, Int_11) knew or heard the expression precision agriculture, and they stated that, unfortunately, it is not yet implemented in Bahrain. Few others have little to medium knowledge about soil sensors and remote sensing, such as Int_6, Int_8, and Int_11 and even used them, such as Int_2 and Int_8. However, only one interviewee has heard about the expression Agriculture Big Data and its related technologies.

Table (4.11): Interview Questions (Storage, Access, and Challenges)

	Q15	Q16	Q17	Q18
	Data storage mechanisms	Access to data	Challenges in collection and access	PA technologies, Soil sensors, Remote sensing, Big Data
Int_1	I don't know	Through manual and physical procedures (paper based or using a CD)	Social boundaries and physical collection effort and time consuming	I did not hear about PA expression before; I knew about soil sensors but not remote sensing and agriculture Big Data
Int_2	Collected data stored on Excel files	Certain governmental entities have a private access to dashboard for agriculture market data. Farmers and other stakeholders access data through manual and physical procedures.	Lack of specialised personnel and dedicated data projects	I am aware of PA however it is not applied in Bahrain; I am also aware and have used soil sensors and remote sensing, however the expression Big Data is familiar but never heard about it in agriculture.
Int_3	I do not know	Directly through MMAA reports	I need environmental and production data; however,	I am aware about PA and it is partially implemented in

		(offline)	it is difficult due to privacy concerns, accountability, and lack of projects and action plans	only 1 or 2 private farms.
Int_5	I do not know, may be in the MMAA.	If I need data, I usually access it through printed copies from the MMAA offices	Not Applicable	I am not aware of any of these expressions.
Int_6	No data currently stored	There is no data to access. If I need data, I collect it myself	Lack of data centralisation, shortage of tools and projects.	I am well aware of PA and other technologies, and I have heard about agriculture Big Data applications.
Int_7	I store the collected data on my computer usually in excel sheets	Directly from my computer. For other data through manual procedure by contacting the owner of the data	Physical and social challenges	I do not know about PA, but I am aware of soil and remote sensing.
Int_8	Data saved on local servers	Data accessed through storage media	Limitations of remote sensing	I am aware of PA and related technologies such

		(flash drives, DVDs, etc.)		as remote sensing
Int_9	I don't know.	I access the data through published research.	I collect my own data based on the project, lack of governmental projects.	I never heard about PA and its related technologies before.
Int_10	Local storage on computers	The collected data previously were published on a yearly paper based periodical available to anyone.	Lack of personnel	I never heard about PA and its related technologies before.
Int_11	On local computers just to keep records for future referencing. Not shared and we don't analyze the data	We only access our local data directly through our devices	Currently we have the required data without any challenges	I heard the PA expression before, I am aware about soil sensing and remote sensing, I know the Big Data in general but not in agriculture.

i. Data Owners/Users' Matrix

Based on the structured interviews along with the results of the different research approaches adopted in objective 1 (questionnaire and field visits), a proposed Data owners/users' matrix for the Kingdom of Bahrain (Data Diffusion) is presented in (Table 4.12) visualize the data needs for each stakeholder. The data will become accessible through the (AgroBahrain) web app through customized interfaces and will be

delivered in different formats such as alerts and notifications, downloadable data sets, and graphs. Agriculture stakeholders are divided into data owners (e.g., farm owners) or data users (e.g., agronomists); however, owners can become users and vice versa.

The following emphasizes different stakeholders in the agriculture sector in Bahrain and related data domains:

1. Farm Owner/Tenant: the stakeholder responsible of one or more farms and can add data such as:
 - Register owned farm(s) data such as farm size and location.
 - Add different cultivated crops details along their used inputs (pesticides, fertilizers, water, nutrients) and amounts.
 - Registering IoT devices and remote sensing technologies
 - Post/Answer enquires to agriculture experts.
 - Add farmers as labour to a farm.
2. Farmer: an employee works on one or more farm and has farming experience. A farmer can provide data related to a farm he/she works in such as cultivated crops details along their used inputs (pesticides, fertilizers, water, nutrients) and amounts. Post/Answer enquires to agriculture experts.
3. Government official: employees in the Ministry of Municipalities Affairs and Agriculture. They have access to many data domains including market data, post/answer enquires, official data (laws and regulations), farms data, and soil and crops data.
4. Non-Government Official: members of non-government organizations including NIAD and the agriculture society. Most of them are farm owners and farmers, others will have limited data access such as secondary data and post/answer agriculture enquires.
5. Merchant: business owner or a salesperson that will sell and buy agriculture products such as crops and provides mainly market data.

Agriculture market data are generated through the parties involved in the food supply chain. In Bahrain, the farm owner, in many cases, is a farmer and a merchant, which plays a significant role in identifying crops and inputs prices.

6. Agronomists/Expert: a scientist in agriculture-related fields such as crops and soil, or an expert on agriculture practices. Usually, he/she collects data for a specific task either directly, and in that case, he/she becomes co-owners of the data with the farm owner or the study area or indirectly by using available data.

Table (4.12): Agriculture data owners/users' matrix for the Kingdom of Bahrain.

		Data Owners					
		Farm Owner	Farmer	Gov. Official	Non-Gov. Official	Merchant	Experts, Agronomists
Data Users	Farm Owner	farms' data, crops' data, sensors' generated, market data, experts' data	experts' data	policies & regulations, sensors' generated, market data, experts' data, secondary data	secondary data	market data	experts' data, secondary data
	Farmer	crops' data, sensors' generated, experts' data	experts' data	policies & regulations, sensors' generated, experts' data, secondary data	secondary data		experts' data, secondary data
	Gov. Official	farms' data, crops' data, sensors' generated, market data, experts' data	experts' data	policies & regulations, sensors' generated, market data, experts' data, secondary data	secondary data	market data	experts' data, secondary data

	Non-Gov. Official	farms' data, crops' data, market data, experts' data	experts' data	policies & regulations, market data, experts' data, secondary data	secondary data	market data	experts' data, secondary data
	Merchant	crops' data, market data		policies & regulations, market data	secondary data	market data	secondary data
	Experts/ Agronomists	farms' data, crops' data, sensors' generated, experts' data	experts' data	policies & regulations, sensors' generated, experts' data, secondary data	secondary data		secondary data, experts' data

4.4. SWOT-TOWS Analysis for strategic options for adopting ICT in agriculture in Kingdom of Bahrain

SWOT Analysis is a strategic analysis method used to evaluate the 'strengths', 'weaknesses', 'opportunities' and 'threats' involved in an organization, a plan, a project, a person or a business activity. It is a tool used to measure the internal factors presented as strengths and weaknesses of an organization (along with the external aspects of opportunities and threats that may support or undermine the implementation of new projects (Parra-L'opez, Reina-Usuga, Carmona-Torres, Sayadi, & Klerkx, 2021). It helps to choose the most matching strategies and strategic options that allow the enterprise to perform its duty and develop concerning the interaction between its internal and external environment.

Adding the relationship between the internal and external factors makes TOWS a much more helpful matrix than a standalone SWOT and an obvious next step. While the SWOT is considered an internal-external analysis, the TWOS is the other way around and is derived from the SWOT as an external-internal analysis tool. The TOWS analysis serves as

a tool for crafting tactics required to execute strategies effectively and identifying specific actions that bolster these tactics. This analytical approach begins by assessing Threats (T) and Opportunities (O), followed by an examination of Weaknesses (W) and Strengths (S). The TOWS framework outlines four potential strategy types, (Parra-L'ópez et al., 2021):

- Growth strategies (SO strategies): These strategies are founded on the idea of optimizing both strengths and opportunities (Max–Max), with a focus on amplifying strengths and capitalizing on available opportunities.
- Reorientation strategies (WO strategies): These strategies aim to minimize weaknesses and maximize opportunities (Min–Max), by mitigating weaknesses and leveraging existing opportunities.
- Defensive strategies (ST strategies): These strategies concentrate on maximizing strengths while minimizing threats (Max–Min), emphasizing the reinforcement of strengths and the mitigation of risks.
- Survival strategies (WT strategies): These strategies are built upon the principle of minimizing both threats and weaknesses (Min–Min), with an emphasis on reducing threats and addressing weaknesses.

Consequently, both SWOT and TOWS analysis techniques were employed to assess the efficacy of ICT adoption in the agriculture sector in the Kingdom of Bahrain. The research will assist in identifying the strategic pathways for using ICT in agriculture for both farmers and agricultural businesses in Bahrain. Furthermore, they comprehensively address all facets of Bahrain's agricultural industry, hence enhancing the efficiency of the adoption process. The information is obtained from on-

site visits (Table 4.6), direct observations, and statistical data extracted from the literature, as presented in Tables (4.3 – 4.5).

Table (4.13): SWOT analysis for the readiness for adopting ICT in agriculture in Kingdom of Bahrain

STRENGTHS • The adoption and implementation of modern agriculture practices such as hydroponics and aquaponics. • Level of awareness and experience regarding suitable agricultural methods among farmers. • The new generation of farmers is more ICT oriented. • Academic and professional agriculture programs and certificates	WEAKNESSES • Low ICT literacy among Bahraini farmers. • Lack of knowledge about ICTs in agriculture. • Nonexistence of a market information system. • Lack of reliable data. • Inadequate water management and harsh soil conditions. • Insufficient use of ICTs for agricultural purposes.
OPPORTUNITIES • The establishment of the Bahraini space agency (remote sensing. • Community and governmental support through different programs and organizations such as NIAD. • Resilient ICT infrastructure. • Sufficient domestic energy supply. • Strategic geographical location for exports in the region • Strong financial support. • Advanced Information Technology (IoT, AI, Data Science, Big Data and etc.) • Relatively low labor cost. • Digital native and digitalization and digital transformation	THREATS • Relative high prices of technologies. • Lack of technology experts. • Land scarcity and growing population • Low return over investment (ROI).

The SWOT analysis revealed that agriculture in the Kingdom of Bahrain awaits a promising future, as there are many opportunities for developing agriculture. The emerging agriculture IT and advanced infrastructure can directly contribute to agricultural productivity and indirectly empower farmers to make informed, quality decisions that will positively impact

how agriculture and allied activities are conducted. However, the weaknesses or challenges, such as low ICT literacy among Bahraini farmers and lack of reliable data, create barriers to exploiting such opportunities, specifically those related to ICT.

Therefore, TOWS analysis was conducted to set a relationship between the internal and external elements and provide the best ways to assess, create, compare, and finally decide upon the strategic strategies to utilize inner strengths to make optimum use of the external opportunities available to the agriculture sector in the Kingdom of Bahrain. In other words, farmers and agriculture businesses must utilize their strengths to cash in on potential opportunities. For instance, by cooperating with financial institutions such as the Bahrain Development Bank (BDB), they can take advantage of the available and emergent IT and expand their usage of modern agriculture practices. Moreover, they have to detect various alternatives to look past the weaknesses and take control of the opportunities in the course. For example, investing in ICT education for farmers and agriculture-related technologies will produce a new generation of agriculture technology experts.

Table (4.14): TOWS matrix for the strategic directions for adopting ICT in agriculture.

	Opportunities	Threats
Strengths	S-O (Strengths – Opportunities): aim to maximize both by mutually benefit from each other (Max – Max). ● Cooperation with financial intuitions such as the Bahrain Development Bank (BDB) will aid the expand and usage of modern agriculture practices. ● Incorporating ICTs courses into the agriculture certifications offered by educational institutions such as the Arabian Gulf University (AGU) will improve the level of awareness among farmers and pave the way to more ICT farming practices.	S-T (Strength – Threats): how to utilize the internal strengths to minimize external threats (Max – Min). ● Encouraging the adoption of modern agriculture practices such as hydroponics, which supports vertical farming as a solution to limited farm size. ● New generations of farmers are more technology oriented; this can open new careers in the agriculture sector related to ICTs. ● Hydroponics initial setup cost is relatively high, however at the long run and with the implementation of ICTs such as Internet of Things (IoT) the cost of human labor, fertilizers, and pesticides will be avoided
Threats	W-O (Weaknesses – Opportunities): Seize external opportunities to overcome weaknesses (Min – Max). ● NIAD as a support organization for the agriculture sector facilitates partnership with educational institutions to provide specialized courses in ICTs for farmers. ● Inducing agriculture related governmental entities and farmers to implement emerging technologies such as Big Data and cloud computing to improve data integrity, reliability, and accessibility. ● Strong ICT infrastructure in Bahrain enables telecommunication companies to easily provide reliable network coverage for every farm.	W-T (Weaknesses – Threats): controlling internal weaknesses to mitigate threats (Min – Min). ● Utilizing ICTs to build a Market Information System (MIS) to help farmers efficiently manage their expenses. ● Investing in ICTs education for farmers and agriculture related technologies will produce a new generation of agriculture technology experts. ● Enhancing data reliability through emergent technologies will enhance production for small farms as well as large ones.

4.5. Summary

In this chapter, the author presented and discussed the results of the methods adopted to achieve the first two objectives. Quantitative and qualitative methods were utilized, and a questionnaire and field visits were conducted to study ICT awareness, readiness, and adoption among agriculture stakeholders. On the other hand, personal interviews and the results from the first objective were undertaken to understand the agriculture data lifecycle.

Although the majority of the questionnaire respondents use a smartphone, they have low ICT literacy and poor awareness about IT in agriculture. However, the country's infrastructure is built to accommodate ICT solutions for all sectors, and the stakeholders show willingness to advance their agriculture business. On the other hand, agriculture data is primarily owned by farm owners/tenants and the Ministry of Municipalities and Agriculture Affairs. However, these data are rarely collected and never shared for public use. Moreover, the data is collected manually and usually stored in Excel files on local devices. Finally, a matrix was built to identify data supply/demand needs among agriculture stakeholders.

Chapter Five

“AgroBahrain” Framework Design and Simulation

5.1. Overview

This chapter introduces a framework called “AgroBahrain” for implementing Big Data applications within the agricultural domain in Bahrain. The focus lies on an in-depth analysis of the framework’s structure and the outlining of its stages. The framework is divided into two primary stages. The initial stage is a preparatory phase aimed at enabling agricultural stakeholders to adopt ICTs, along with the integration of emerging technologies such as IoT, cloud computing, and Big Data, significantly amplifying the agricultural industry’s capabilities. The following stage illustrates the information lifecycle of the system, consisting of three interconnected stages. The initial stage entails gathering a wide range of agricultural data from multiple sources and utilising various tools to accommodate varied formats. The second procedure focuses on creating a storage schema to preserve the collected data in its original, unprocessed format. Storing data in this format improves the effectiveness of future data transactions. The final stage entails converting raw data into various formats that are appropriate for analytical insights and visual representations. Afterwards, the modified data is kept in highly scalable repositories to ensure rapid and efficient data retrieval.

Concurrently, the chapter focuses on creating a simulation model that mimics the workflow of the framework's second stage. The model simulates different processes related to a particular stakeholder. The simulation involves the construction of a user interface and the establishment of a backend environment for the storing and manipulation of data. In addition, as a result of organizational limits and financial

constraints, this simulation will require the use of both authentic and mimicked data sources.

5.2. The Framework

As per the outcomes and discussions pertaining to the primary research objectives—ICT Awareness, Readiness, and adoption, alongside the secondary objective concerning the Supply/Demand of Agriculture Data—it is evident that Bahrain is at an early stage in integrating ICT within its agricultural sector. The level of ICT literacy among farmers in Bahrain remains notably deficient, with minimal engagement with technology. Moreover, the average farm size in the region is relatively small, approximately two hectares, which suggests that certain modern technologies, such as autopilot tractors and variable rate technologies, may not substantially enhance agricultural productivity. This implies that farm-level solutions might prove inefficient, considering their associated costs and managerial demands. Most of the farms consist of open and covered cultivation areas. Open farming is limited to one annual cycle due to challenging weather conditions. However, several farms employ year-round operations using either traditional greenhouses or hydroponics, which are covered structures. To effectively execute the AgroBahrain framework, it is crucial to increase awareness, preparedness, and acceptance of ICTs among stakeholders in the agricultural sector. The established framework's objective is to operate nationwide, guaranteeing enhanced availability of data and adaptability in its structure, ultimately promoting agricultural sustainability within the Kingdom of Bahrain.

The AgroBahrain framework is devised to facilitate the semi-autonomous collection and dissemination of significant agricultural data by fostering collaboration and integration among various stakeholders, including

farmers, agronomists, merchants, governmental and non-governmental organizations, and cloud services.

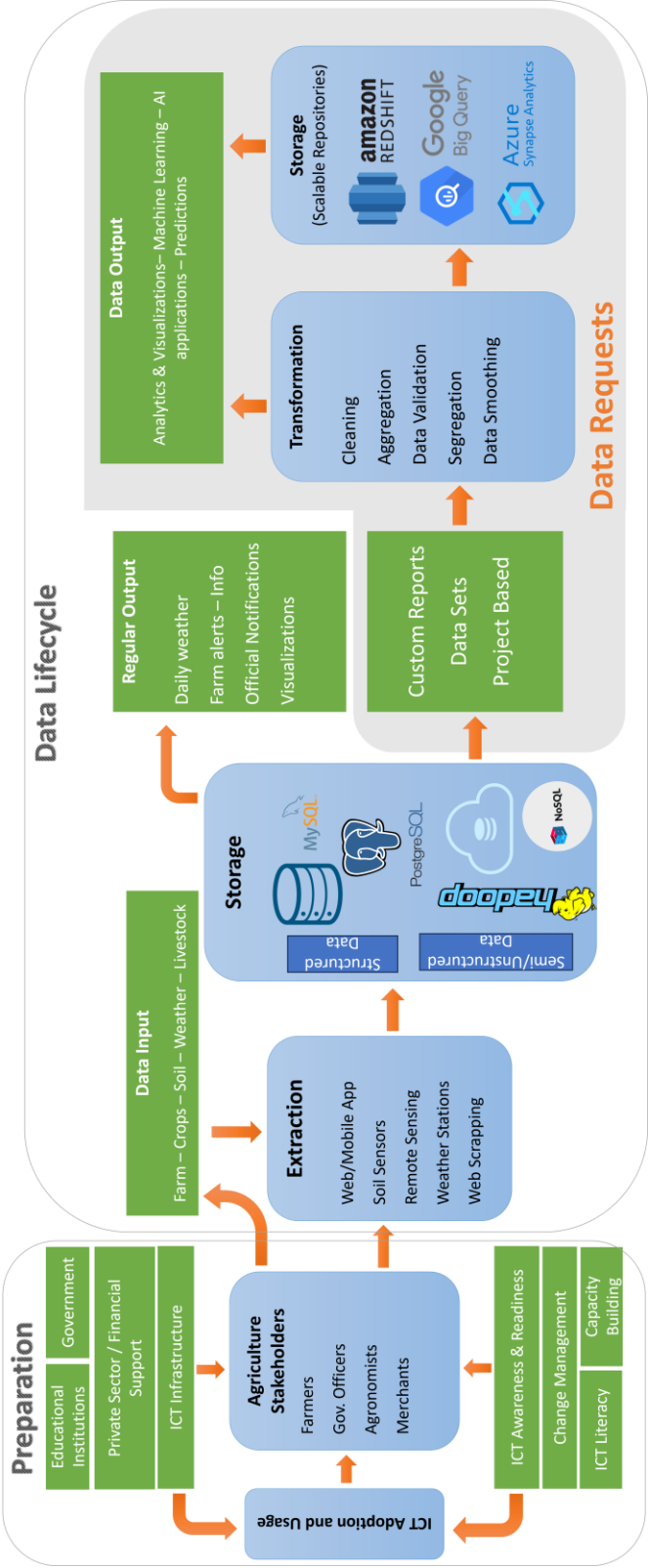


Figure (5.1): The AgroBahrain Framework

Figure 5.1 illustrates the AgroBahrain framework, which is divided into two stages. The first stage, elaborated in section 5.2.1, presents a recommended set of requirements aimed at improving the agricultural sector's knowledge and readiness to adopt and utilise ICTs in agricultural activities. The following stage depicts the data lifecycle within the framework, which is organised into three interconnected processes, as described in sections 5.2.2 and 5.2.3. While the framework predominantly leverages Google Cloud services, it remains adaptable and robust, capable of integrating with other cloud service providers or employing a complete Big Data stack. The adopted approach primarily focuses on seamlessly integrating ETL and ELT operations to maintain a continuous and sustainable flow of various data types.

5.2.1. Preparation Stage

The effective deployment of the AgroBahrain framework relies on the incorporation of ICT tools and concepts, such as IoT devices, process automation, and mobile and web-based apps. In order to guarantee the effective execution of the framework, it is imperative that users are sufficiently equipped in terms of technical expertise, financial resources, and psychological readiness.

Integrating ICT into agriculture requires careful technical preparation to ensure its successful implementation. It is a complex procedure that requires careful planning and execution. Essential components of this preparation include constructing the required infrastructure, supplying hardware and software, delivering training and capacity building, assuring data management and security, tackling connectivity obstacles, and incorporating IoT devices. Furthermore, it is essential to take into account the varying degrees of ICT proficiency among individuals involved. The level of ICT literacy among farmers in Bahrain varies from poor to intermediate (Al-Ammary & Ghanem, 2023). Comprehensive

technical training should include essential skills, such as computer proficiency and basic application knowledge, as well as more advanced abilities, such as remote sensing and data science expertise. It is crucial to offer training to farmers, farm tenants, and official agents about the installation and maintenance of devices such as soil sensors and digital cameras. Additionally, equipping them with the knowledge of effective data management using a variety of digital technologies, including mobile and web applications, is of paramount importance.

i. Infrastructure

The first step in supporting the integration of ICT into the agriculture industry entails establishing fundamental infrastructure. This involves the delivery of reliable and affordable internet connectivity, which acts as the fundamental element for smooth integration of ICT.

Through field trips and personal interviews, it was noted that numerous farms suffer from inadequate power supply and network access. Establishing high-speed broadband connectivity in farms is crucial to ensure that farmers and other stakeholders in agriculture have unobstructed access to digital tools, relevant information, and necessary services. Simultaneously, the establishment of this infrastructure requires the guarantee of a consistent power source, as ICT equipment depend on electricity for their functionality.

ii. Digital Tools and Devices

After the establishment of the infrastructure, it is crucial to equip farmers with the requisite hardware and software. These encompass technology like smartphones, tablets, and computers, which facilitate access to digital tools and information. Furthermore, it is imperative to create or acquire agricultural software applications and platforms that are tailored to meet the precise requirements of farmers. These may encompass applications for predicting weather, software for managing crops, and systems for

tracking livestock. Moreover, the integration of ICT in agriculture yields a substantial volume of data, encompassing details concerning crops, soil conditions, weather patterns, and additional factors. Effective data management is essential to ensure the systematic organisation, secure storage, and timely accessibility of information. Furthermore, implementing robust data security procedures is crucial in safeguarding sensitive agricultural information against unauthorised access or cyber threats. Ensuring data privacy and integrity is vital to building trust in the ICT systems.

iii. ICT Literacy and Capacity Building

Based on the SWOT analysis conducted in section 4.4, major weaknesses involve poor ICT literacy and knowledge among Bahraini farmers. It is essential for farmers and agricultural stakeholders to possess the necessary knowledge and expertise to properly utilise ICT tools and technology. Capacity building and training programmes are crucial for equipping the stakeholders with the requisite digital literacy. These programmes should encompass the use of hardware, software, and the internet, together with the particular applications pertinent to agriculture. Furthermore, the sustained effectiveness of integrating ICT in agriculture depends on continuous support and maintenance. Immediate access to technical support should be provided to resolve any hardware, software, or network connectivity issues. Moreover, it is imperative to consistently upgrade and enhance the ICT systems in order to maintain their relevance and efficiency.

iv. Financial Preparation

Effective incorporation of ICT into agriculture necessitates meticulous financial planning. To incorporate ICT into agricultural practices, a thorough financial preparedness strategy is necessary. This plan should include various crucial elements, beginning with investments in

infrastructure. This encompasses the acquisition of critical equipment such as sensors, drones, GPS devices, and software. Farmers must efficiently arrange resources and investigate viable financing methods to cover this crucial initial expense. Training and Capacity Building should be prioritised alongside other efforts, as farmers and agricultural labourers require thorough guidance on the operation and maintenance of ICT equipment. Therefore, when considering the budget, it is important to take into account the related costs of training.

Furthermore, the integration of ICT relies heavily on Connectivity and Internet Access. In order to enhance connectivity in farms, securing finance may be imperative. The vast amount of data produced by ICT requires careful management and storage, with financial allocations necessary for secure data storage systems and cybersecurity measures.

Agricultural stakeholders must develop contingency strategies and allocate financial resources to address unforeseen expenses or difficulties. In addition, farmers may consider utilising Government Support and Grants, as Bahrain provide technical and financial support for individuals and farmers, including ICT services, equipment, and Agricultural loans (Alwatan, ٢٠٢٠).

v. Cultural Preparation & Change Management

Farmers and stakeholders must undergo a cognitive transformation in order to fully harness the promise of ICT in agriculture. In order to properly utilise the potential of ICT, it is essential to engage in mental preparation. Farmers and stakeholders should be receptive to change, enthusiastically use digital tools, and adjust to new techniques. Education and awareness efforts are crucial for narrowing the digital divide and ensuring that farmers can fully use the myriad benefits of ICT in agriculture. Given the appropriate mindset and assistance, the agriculture sector has the potential to flourish in the digital era, attaining enhanced

efficiency, sustainability, and global market reach. Prior to implementing ICT in agriculture, it is essential to engage in mental preparation through education and awareness. It is crucial to educate farmers about the prospective advantages of ICT and its ability to revolutionise their agricultural methods. Education programmes should prioritise digital literacy, with a specific emphasis on utilising mobile phones, apps, and other ICT technologies. Workshops, training events, and outreach programmes can proficiently instruct farmers on leveraging technology for their benefit. In addition, governments, non-governmental organisations (NGOs), and private sector stakeholders have the ability to significantly contribute to creating a conducive climate for the adoption of ICT in agriculture. This includes the provision of subsidies to encourage the adoption of technology, the establishment of incentives to promote sustainable farming methods, and the guarantee of dependable network connectivity.

Conversely, the field visits and personal interviews revealed that Bahraini farmers frequently adhere to conventional practices that have been inherited across generations. Resistance to the adoption of new technology may arise due to apprehensions over expenses, intricacy, and uncertainty. An alteration in thinking is necessary to embrace the potential of ICT. A significant number of farmers and farm tenants in Bahrain express scepticism towards technology, perceiving it as an avoidable expense that adds little value. Preexisting assumptions and prejudices can significantly impede the acceptance of new concepts. Effective mental preparation entails addressing preconceived notions, scrutinising their legitimacy, and adjusting or eliminating them as needed.

5.2.2. Data Lifecycle: Overview

The second stage of the framework represents the agriculture data lifecycle within the Big Data context. It focuses on three processes: 1- data extraction, 2- storage of raw data, and 3- transformation and storage of processed data ready for sharing and usage by different stakeholders.

i. Data Extraction

Agriculture data extraction refers to the systematic collection of diverse data pertaining to farming activities from many sources, including event data streams, logs, IoT devices, secondary data, and transactional applications. The identification of data sources was determined by the study methodologies employed to address research objectives 1 and 2, which encompassed interviews with various agriculture stakeholders, visitation, and relevant literature. The AgroBahrain framework focuses on retrieving data pertaining to several categories of agriculture data, including weather, soil, crops, markets, farms, and stakeholders.

ICTs provide farmers with access to a variety of agricultural data, such as up-to-date weather forecasts, past climatic patterns, and microclimate data. This enables farmers to strategically plan their planting and harvesting schedules and adopt resilient methods based on weather circumstances. In addition, they provide assistance in soil testing and analysis, helping farmers understand nutrient levels, pH, and moisture content. This empowers farmers to make more informed decisions about fertilisation and irrigation methods. By utilising sensors, drones, and satellite imagery, farmers have the ability to precisely observe the condition of their crops. This allows them to rapidly manage problems such as pest infestations, diseases, and plant stress. Consequently, farmers may decrease their dependence on chemical inputs and enhance agricultural yields. Furthermore, the availability of market data through mobile applications and internet platforms empowers farmers with the

necessary information to make informed decisions on crop selection and the most favourable selling periods, thus improving their profitability. Data extraction in livestock management enables the monitoring of animal health, growth, and production, leading to enhanced livestock management techniques and increased productivity. ICTs facilitate resource management by enabling the efficient utilisation of essential resources such as water and electricity through data extraction. This, in turn, contributes to sustainable farming practices and the conservation of resources.

The framework's objective is to gather diverse agricultural data semi-autonomously and thereafter build scalable pipelines for transmitting the data to the cloud for storage, processing, and preparation for analytical operations.

ii. Data Storage

Extracted data is either stored in its original, unprocessed version or prepared for further analysis, such as analytics, querying, or visualisations. Data in its unprocessed form is gathered directly from its primary sources, such as sensors, satellite photos, legacy files, and so on. The data will be kept in a schema-less storage repository known as a Data Lake for the purpose of archiving and pre-processing.

The chosen storage schema is a bespoke model that offers flexibility for complex structures such as agricultural data. It enables the use of numerous joins and accommodates many facts and dimensions tables. The schema serves as the fundamental basis for every data storage system. Within the context of agriculture, the data storage schema delineates the arrangement and storage format of data. It encompasses multiple elements and factors. Firstly, it incorporates the various sources of agricultural data, such as remote sensors, drones, satellite photography,

manual field observations, and others. The schema should be thoroughly designed to effortlessly collect data from these varied sources.

Agricultural data comprises a wide variety of formats, including quantitative data like temperature and humidity, qualitative data like crop descriptions, geographical data like GPS coordinates, and multimedia data such as photographs and videos. In order to accommodate this wide range of possibilities, the schema must possess the ability to adapt. In addition, the schema can utilise other data storage technologies, such as relational databases, NoSQL databases, or distributed file systems, depending on the specific requirements of the farm.

The schema plays a vital role in enabling the integration of data, which involves consolidating information from many sources to provide a comprehensive perspective of the farm's activities. This integration is crucial for making educated decisions.

To ensure the protection of agricultural data, it is essential to incorporate strong security features into the schema in order to prevent unauthorised access and any data breaches. Last but not least, the schema should prioritize accessibility, ensuring that authorized stakeholders can conveniently access the data through user-friendly interfaces. Timely and well-informed decisions rely on this crucial accessibility for farmers, agronomists, and researchers.

iii. Data Transformation

To make raw agricultural data actionable, it must be transformed into meaningful information. Algorithms, machine learning, and artificial intelligence are utilised to convert raw data into useful knowledge. These technologies have the capability to analyse extensive datasets in order to offer farmers recommendations regarding crop management, irrigation scheduling, and pest control.

Soil sensors have the capability to gather information on moisture levels, nutrient content, and pH data. Through the process of converting this unprocessed data, farmers can obtain precise instructions regarding the optimal timing and quantity of irrigation, the appropriate selection of fertilisers, and even the suitable choice of crops to cultivate in specific regions of their fields. Such a high degree of accuracy enables farmers to minimise the squandering of resources, diminish their environmental footprint, and optimise crop yields.

The process of transformation will be conducted iteratively, employing tools such as Google Cloud Functions (GCFs) and customised SQL queries, and it will be divided into:

- Initial transformation: text data collected from sources such as soil sensors, weather stations, and satellite remote sensors is presented in different formats such as JSON and XML. To unify the storage format and enhance access to raw data, all text data extracted from automated sources such as soil sensors will be transformed to CSV format and stored in (.csv) files.
- Transactional transformations: data is transformed to suit platform requirements to present updated data such as (weather and soil) using web or mobile applications. Transformations encompass a wide range of actions, including but not restricted to retrieving the most recent data connected to modules. Data pertaining to many modules, such as weather, soil, and crops, may be gathered by certain sources, necessitating the segregation of this data.
- Custom transformation: the transformation process is executed in response to user requests for specific data sets, encompassing a multitude of processes. Performing data cleaning by identifying and eliminating duplicate entries, deleting unnecessary spaces and null

values, and verifying the data range and presence of outliers. Standardizing data types, measurement units, and field names. Data aggregation is the process of collecting and combining data in order to produce visualizations and prepare data for analysis, with the intention of providing valuable information to stakeholders.

5.2.3. Implementation Scheme

The methodologies employed for data collection uphold continuity in data acquisition across varied formats—structured, semi-structured, and unstructured data (Table 5.1). Structured data were mostly obtained through the PWA platform (Figure 5.2), specifically for tasks such as registering new stakeholders and adding new farms, crops, and related inputs. The semi-structured data obtained from sources such as the Pycno soil sensor or the crop monitoring platform is received in JSON format. However, a part of this data undergoes transformation using Python during the collection phase to convert it into CSV files, thereby improving its practicality in the future. Unstructured data includes photos obtained from the multispectral camera and NoSQL records produced by the agriculture enquiries service (AES).

Table (5.1): Agriculture data categories and sample sources

Category	Source	Raw Format
Structured	PWA Forms	String, number, date and time, etc..
	Web Scraping (peninsula farms)	
Semi-Structured	Pycno soil sensor	JSON (string and numbers)
	EOS Crop monitoring platform	
Unstructured	Multispectral Camera	Images (JPG, TIF)
	PWA agriculture enquiries service (AES)	Text, images, other files

As explained in section 5.2.2, the data flow process consists of three primary phases: extraction, storage, and transform/store. The proposed techniques for extraction and data sources consist of the following components: PWA, soil sensors, weather stations, spectral and regular cameras (with or without unmanned vehicles), satellite remote sensing, secondary data, and web scraping.



Figure (5.2): The PWA Main Page

The PWA platform serves as the central connection inside the system, addressing the various needs of different stakeholders. The preference for PWA over other development technologies arises from its unique structure, which enables flexible data representation using different tools and formats. The platform will collect agricultural data, encompassing demographic information of stakeholders, farm details, crop production data, and enable the direct transfer of historical data from external files.

In addition, it will provide registered users with an inquiry service that allows for the collection of various data categories and formats.

Simultaneously, the PWA will provide fundamental information customized for certain stakeholders. As an illustration, owners of farms have the ability to obtain previews of soil information that is collected via farm devices that have been registered, such as soil sensors. The PWA will be accessible to both registered stakeholders and guest users, but the latter will have limited features. The collected data will be associated with a stakeholder ID to guarantee the legitimacy of the data and enhance its traceability.

Registering new users is a crucial first step that allows stakeholders to enter data and utilize all the features of the platform. However, guest users can only access a limited number of features, including market prices, weather information, and government-issued laws and regulations. The registration process is a multi-stage procedure designed to authenticate new users and verify the accuracy of the submitted data through connecting with eGovernment services. Please refer to Figure 5.3 for a visual representation. The Information and eGovernment Authority (IGA) in Bahrain functions as the authorized government entity responsible for delivering electronic services. It is responsible for managing and updating comprehensive digital records of individuals, both citizens and residents, which encompass various information such as demographic details, contact information, and employment records. These records are interconnected through the use of a national identification number.

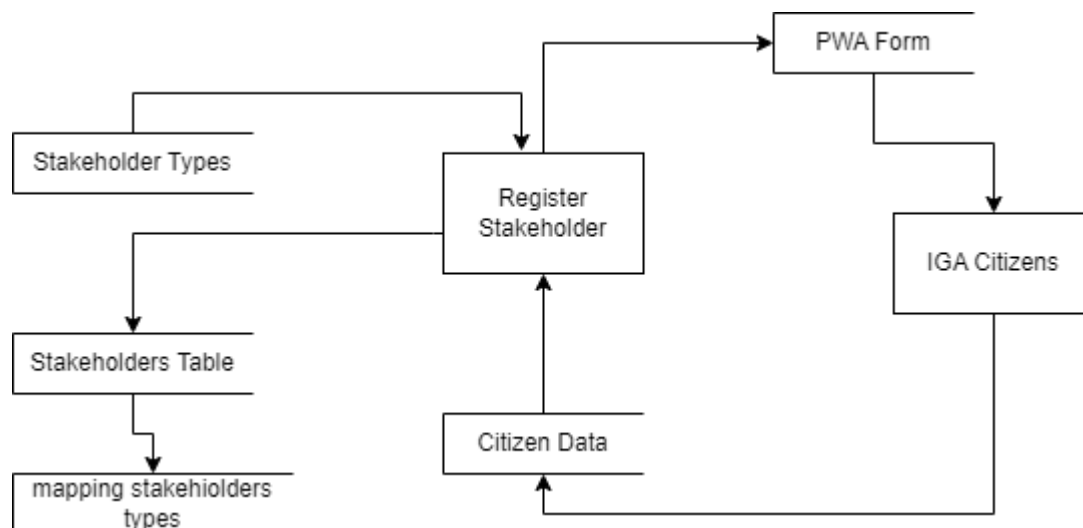


Figure (5.3): Agriculture stakeholder registration scheme

PWA users will have the option to register by providing their national identification number, ID expiration date, residential address, mobile phone number, and selecting their account type (which includes options such as farm owner, merchant, farmer, non-governmental agricultural expert, official agriculture staff, and agronomist). While not mandatory, the user optionally chooses to supply an email address. The registration data will be sent to the IGA API for the purpose of verifying and authenticating the selected account by comparing it with employment data. After confirmation, essential user information such as name, date of birth, and education level will be retrieved from the IGA, and a new stakeholder record will be created. Afterwards, the platform will send an SMS with a link to set up the account password and activate the account. The selection of SMS as a primary means of communication is based on the limited level of ICT literacy within the farming community, as well as the lack of functional email addresses.

i. Direct Data

Information is collected from both primary and secondary sources using the PWA. Three separate approaches are used to directly collect data. The first approach entails the direct input of unprocessed data via the

platform's forms on farms, crops, and production parameters. Furthermore, this procedure adheres to a hierarchical structure, requiring the addition of a farm and crop prior to a stakeholder's ability to incorporate agricultural production cycles. Data obtained through forms is saved directly in a structured format using a relational database, such as MySQL.

The second approach involves manually inputting previously obtained or legacy data into the PWA using an external file service. Active stakeholders are able to manually upload data files to Cloud Storage through this approach. The uploaded data can be categorized as either primary or secondary. Primary data are data originated from sources such as soil sensors and user input. Secondary data, although initially identified as such, is considered primary data when requested by other stakeholders through the platform, yet it may have limited customization options. Supported file formats encompass, but are not limited to, CSV, Excel, pictures, and PDF documents. During the upload process the user is required to provide the following meta data:

- Data category (weather, soil, crop image, etc.).
- Since most of the collected data is usually related to a certain crop production, the crop name along the farming method can be defined using drop down lists if the data linked to a predefined crop.
- Data period (from – to).
- Dynamic data fields (name, unit of measurements).
- Data Interval (hourly, monthly, etc.).

The file is directly uploaded to a data lake within a folder labelled with the stakeholder ID. Simultaneously, the metadata collected is appended to a table located along the path of the uploaded file.

The third approach entails gathering data using a question-and-answer format using the AES. This service is an important feature integrated into the PWA, enabling various agricultural stakeholders to exchange knowledge and experience through an interactive question-and-answer format. In order to either submit an inquiry or provide a response, users must have an active stakeholder account. However, all responses are made available anonymously to the platform guests, serving as public knowledge.

A NoSQL database will be adopted to store and manage inquiries and their corresponding responses. The choice of tool is determined based on its ability to manage diverse data forms, such as text and files, which ensures enhanced versatility in data manipulation.

Stakeholders who are actively involved will be given permission to access three separate entities: the list of inquiries, the ability to submit an inquiry, and the ability to view detailed inquiry data. The main component, known as the enquiries list, serves as the core hub, showcasing all inquiries, whether they have been responded or not, and offering several search tools to filter the results. Stakeholders can interact with an enquiry by viewing its details, providing a response, and modifying or removing it if they are the owner of the inquiry.

Within the enquiries interface, users have the ability to start a new inquiry. They can input text, capture images of crops or leaves, or attach files. Images and files are stored in a specified Cloud Storage directory, with each directory being identified using the stakeholder ID. Furthermore, the full file path in the storage system is included as a document field in the NoSQL database. Every document is allocated a unique ID (primary key). Enquiries will be assigned auto-generated IDs, while answers will use the enquiry ID as a custom identifier, followed by a defined answer order, enabling numerous responses per enquiry.

ii. Indirect Data

On the other hand, indirect data sources facilitate automated and remote data inputs into the system, including wireless sensors deployed on farms, external online platforms, and weather stations. Automated data sources must go through several stages to ensure the credibility and establishment of seamless, flexible data collection procedures. During the initial stage, stakeholders are required to register the data sources using PWA forms before they can make use of them. The registration process is a series of steps that requires the input of essential metadata pertaining to the technology being registered, such as soil sensors and remote sensing platforms. This includes providing information such as technology names, descriptions, and types. Upon registration, the technology is included to the agriculture technologies and tools page, granting stakeholders the ability to modify or remove the data source as desired.

Automated data sources should be equipped with different APIs to streamline data interchange among various platforms and user interfaces. APIs have become essential components in the software development environment, particularly in online applications since the year 2000. Ofoeda, Boateng, and Effah (2019) state that they enhance interoperability, promote component reusability, and streamline software development processes

APIs consist of many endpoints that allow for the retrieval of different categories of data. An endpoint refers to a URL that is created by the supplier of a data source. It usually requires special authentication, such as a user token or key, in order to access specific sets of data. The endpoints can be accessed using either GET or POST requests. The latter sometimes requires a body in specific formats, such as JSON. The variables present in the body may vary among endpoints within the same API, which would make automated management more complex.

For example, the APIs linked to the EOS crop monitoring web platform include a range of endpoints that use the POST request approach. Each endpoint has its own set of body variables. Concurrently, the weather API of the platform consists of three distinct endpoints, with one of them necessitating constant body data (geolocation points of the specified area) to deliver an aggregated weather forecast for today and the subsequent two days. Automating the last two endpoints is tough due to the requirement of using variable data with different variable names. Therefore, the manual upload capability offered by PWA forms is a more convenient choice for managing data produced by such endpoints. Stakeholders have the option of capturing data in CSV or Excel formats and manually submit them, ensuring accuracy in the transfer of data.

During the next phase, the user will focus on editing the technical details to complete the registration process. The modifying feature linked to each technology will allow stakeholders to register API endpoints through a PWA form, including the necessary response fields. Stakeholders must enter the exact field names as specified in the endpoint response, while also indicating the data type for each field and providing descriptions from a predetermined list (such as Temperature, Humidity, Soil Temperature, and others). After registration, a cloud script will be triggered to create a new storage folder in the Cloud Storage system (data lake) using the stakeholder's ID, if it doesn't already exist. Additionally, it will initiate the creation of a new CSV file with the response fields as column headings.

The next stage begins when the stakeholder assigns a technology for a currently operating crop production, which then activates and automatically becomes part of the endpoints through the initial execution of a cloud script. Due to the automated nature of this process, it is important to acknowledge that it does not consider variations in the

interval of data source updates and depends on incremental extraction and loading of data according to fixed intervals. Therefore, cloud scripts will be consistently triggered through a URL, utilising a cloud scheduler at predetermined time intervals, to collect data and place it into specific CSV files.

Stakeholders have the ability to initiate an active production via PWA forms, enabling subsequent modifications to in-production crop details and the activation of a data source. When the cloud script begins functioning, it will regularly analyse a database table to identify active technologies. It will then request the corresponding endpoints to retrieve the responses. The captured responses will be added to the designated CSV files.

Web scraping, also known as web crawling, represents an indirect but valuable source of data acquisition. This method entails the automatic retrieval of data from various web pages, either in real-time or in batches, which is then used for analytics and AI applications. The growing implementation of this technology in various industries and scientific domains, such as Big Data, cybersecurity, and artificial intelligence, highlights its significance (Bifulco, Cirillo, Esposito, Guadagni, & Polese, 2021). Web scraping has become a crucial data source in the field of agriculture, particularly for market price predictions and traceability (Fernandes & Reite, 2021). Nevertheless, the legality surrounding web scraping practices remains a concern (Khder, 2021), and it is recommended that official entities should oversee and execute such operations.

Due to the unique structure of every web page, two primary approaches are recommended for extracting data. The first approach entails thoroughly examining individual websites to identify their content, target fields, and associated web pages. Afterwards, a custom scheme is created,

consisting of a table that identifies the target fields and a cloud repository for storing unstructured data, such as images and files. Conversely, the second strategy chooses not to analyse each website or develop a tailored framework. Instead, it uniformly scans all websites by consolidating their content into a single file, from which data is collected according to user demands, using semantic analysis through Natural Language Processing (NLP) algorithms.

Semantic analysis, a branch of linguistics, explores the process of interpreting language by examining the meanings of individual words and phrases, which are fundamental in our understanding of the universe (Salloum, Khan, & Shaalan, 2020). Although the second strategy is more time and resource efficient, the first method is more suited for Bahrain because there are just a few websites that provide agricultural data.

In summary, web scraping is an indispensable method for collecting data from websites. Nevertheless, it is crucial to thoroughly evaluate the legal aspects and various approaches to implementation in particular circumstances and businesses (Khder, 2021).

The third phase of the implementation process within the framework (transform/store) is responsible for transforming the collected data into a format that users can easily understand. The data flows through the PWA in its original or unprocessed form, or it can be converted into knowledge and forecasts by utilising personalised datasets. The original or raw data, derived from sources like sensors, transactions, and Software as a Service (SaaS) platforms, pertains to real-time data specifically gathered for ongoing research purposes. Its defining characteristics include originality, reliability, and proprietary rights (Ajayi, 2017). On the other hand, secondary data is historical information collected for diverse applications, available for reuse by various individuals or groups (Hox & Boeijs, 2005).

Active stakeholders have the capacity to preview their own primary data using the PWA interface, which includes information about their profiles, farms, crops, and agricultural production data. In addition, data produced by technology associated with active crop production will be displayed in a raw data format, presented in tables and simple charts. The default display will show the most recently acquired data, while users will have the option to filter and examine earlier data. Simple charts format will be used to visually show weekly data points.

Following that, stakeholders other than the original data owners will have access to the data acquired from other sources, as previously mentioned in the data owners/users' matrix. Stakeholders have the option to view data either anonymously or identified. Stakeholders' roles and access levels limit the use of identified data. For example, a farmer who is officially employed by a farm owner and assigned to a farm that is currently engaged in production will be granted full access to all relevant data related to that farm, including information about the soil. On the other hand, stakeholders with restricted access, such as agronomists, can retrieve anonymous data. The data will be extracted and organised into datasets for analytical purposes, future studies, projections, and other similar applications.

A dataset consists of empirical, quantitative, and figurative data pertaining to a certain subject, such as the average monthly humidity for a specific timeframe. There are five types of information that can be presented as datasets: numerical, bivariate, multivariate, categorical, and correlation (Niu, 2016).

The PWA form will empower registered stakeholders and guest users to generate datasets in various formats, including CSV, Excel, and images. The form contains different filters to allow for advanced queries including:

- Data categories (Weather, Soil, Crops, Market, Stakeholders, and Farms). Based on the selected category a predefined list of fields will be presented with the option to remove or add another field.
- Date/time intervals.
- Fields and units.
- Aggregation functions (total, mean, maximum, minimum).

The task of generating a data set requires the use of the ETL procedure, where data is extracted in raw format from different sources on the data lake (relational database, CSV files, excel, images), in which transformation operations such as data cleaning, unifying data units (Celsius or Fahrenheit) for similar fields and aggregation is required before loading the data. Generated data sets are either downloaded into the user device in formats such as CSV or stored on the cloud as a sharable data for analytics and visualisations by others.

In the following section, the various processes that are explained in the implementation scheme will be systematically and visually demonstrated.

5.3. Model Simulation

A simulation model was created to imitate the operational sequence of the framework's second stage. This model is designed to replicate the specific functions connected with predetermined stakeholders' and guest users' perspectives.

The development operations primarily took place on an iMac using VSCode as the code editor. The Ionic framework was used for the front-end user interface, while Python was utilized for backend scripting. The PWA paradigm effectively meets the needs of all stakeholders. Nevertheless, as a result of limitations in both time and budget, the implemented functionalities predominantly replicate the experiences and

responsibilities of the farm owner stakeholder while only partially incorporating the roles of a farmer and an agronomist.

The adoption of the Google Cloud Platform facilitated the coordination of data extraction, storage, and processing, making use of features like Cloud Scheduler, Cloud Functions/Run, and Cloud Storage. The simulation utilizes the data categories outlined in Table 5.2. Data collection was performed using automated Python scripts that were initiated by Google Cloud Scheduler (Cron Job) or through URL trigger and executed using GCFs. Alternatively, data was collected via PWA forms.

To mitigate the risk associated with the extraction or utilization of untrusted data sources, all data must be added and linked to a registered stakeholder account created through the AgroBahrain PWA. It is worth mentioning that a stakeholder can serve various roles, such as being both a farmer and a farm owner. Nevertheless, data obtained via web scraping activities lack any association with a specific party, and it is recommended that they be collected and monitored exclusively through official government channels.

The data collection process comprised either mimicked data or using authentic data sources, including legacy data from governmental and non-governmental organizations in CSV format, taking into account both organizational and technical limitations. Two historical weather files in CSV format, spanning from 1901 to 2021, were obtained from the World Bank repository at the following link: <https://climateknowledgeportal.worldbank.org/country/bahrain/climate-data-historical>. The initial file depicts Bahrain's average monthly temperature, while the second one showcases the average monthly precipitation.

Table (5.2): Data used for the simulation model.

Data Domain	Collected Data	Real/Simulated
Stakeholders	Demographic	Simulated for 500 farmer /farm owners using the faker library in python
Farms	Geolocation, area, name, governorate	Simulated for 100 farm using faker library in python
Crops	Images	Real data using a multispectral camera
	Name, description, farming type, harvest amount, cultivation area, inputs (fertilizers, pesticides)	Simulated using the faker library in python
Soil	Soil temp, Soil moisture	Real using the Pycno soil sensor
Weather	Air temp, Air humidity, Solar radiation, Rainfall intensity, pressure, wind	Real using: Pycno sensor, legacy data from the world bank
Market	Local vegetables prices	Real data using web scrapping on the peninsula farms web site (https://peninsulafarms.com/product-category/vegetables-fruits/)

In order to automate the collection of data in real-time, an IoT soil sensor called Pycno was utilised. This was done in combination with a subscription to a remote sensing platform that uses satellite imagery called EOS. The purpose of this online tool is to provide crop monitoring services. The Pycno soil sensor was strategically placed in a greenhouse located on a tomato farm. The deployment took place from November 2021 to January 2022, as shown in Figure 5.4.



Figure (5.4): The Pycno soil sensor positioned in the selected farm.

During the same period, images of the tomato crop within the farm were manually captured on two separate occasions using a multispectral camera (Figure 5.5).

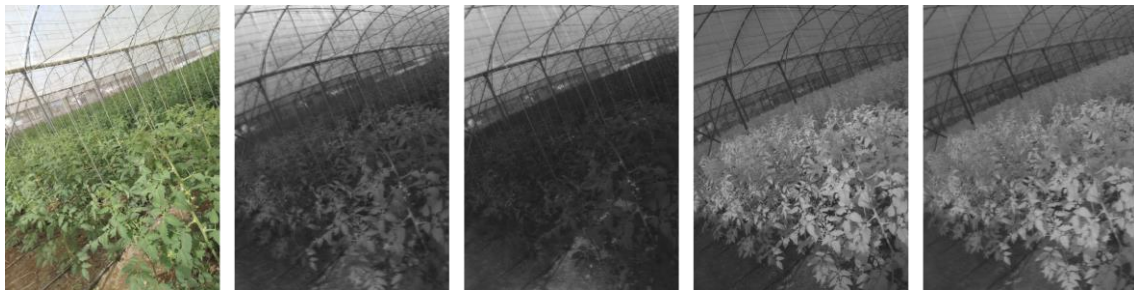


Figure (5.5): Multispectral images of tomato. Bands from left (RGB, Green, Red, Red-edge, Near Infrared).

```

"S1T": 66.2,
"S2T": 61.81,
"HUM": 96.34,
"TEMP": 15.18,
"LW1": 3.28,
"LX1": 5896,
"RAINH": 0,
"name": "pycno 1",
"LFREQ": 868,
"NS": 320,
"BAT": 4.15,
"CME": 0,
"GR": 1,
"GRF": 0,
"VE": 4.73,
"MDM": 68,
"TXT": "No GPS",
"loc": {
  "type": "Point",
  "zoom": 14,
  "valid": 1,
  "updatedAt": "2021-12-25T06:24:46.231Z",
  "satellites": 6,
  "coordinates": [
    "50.532875000",
    "26.171695000"
  ]
},
"epoch": 1641878076662,
"polledAt": "2022-01-11T05:14:36.662"
}

[
  {
    "2022-10-30T00:00:00+00:00": {
      "t_2m__min": 298.2,
      "t_2m__max": 300.6,
      "asob_s__sum": 0,
      "tot_prec__sum": 0,
      "h_snow__sum": 0,
      "relhum_2m__avg": 85.7,
      "u_10m__avg": -0.9,
      "v_10m__avg": 0.8,
      "w_speed__avg": 1.2,
      "w_speed__min": 1,
      "w_speed__max": 1.5,
      "clct__avg": 1.3,
      "td_2m__avg": 296.6,
      "t_so__min": 298.6,
      "t_so__max": 300.5,
      "ps__avg": 101196,
      "w_dir__avg": 312.2,
      "ww": 0,
      "ww_human": "clear sky"
    }
  },
  {
    "2022-10-30T01:00:00+00:00": {
      "t_2m__min": 297.7,
      "t_2m__max": 300.4,
      "asob_s__sum": 0,

```

Figure (5.6): sample of data collected. On the left sample JSON data from Pycno soil sensor, and on the right sample JSON of EOS satellite data.

A rectangular area of 1.7 hectares was located and chosen within the 'Horaat A'ali' region using the crop monitoring platform. This region encompassed both open and covered agricultural plots, which included the presence of greenhouses. Starting in November 2021, the monitoring platform consistently collected data, despite the constraints of satellite remote sensing, especially in areas with covered agricultural plots or when there was heavy cloud cover. Nevertheless, data collection was effectively carried out for the open field region, providing either numerical data (Figure 5.6) or visual maps displaying vegetation indices (Figure 5.7).

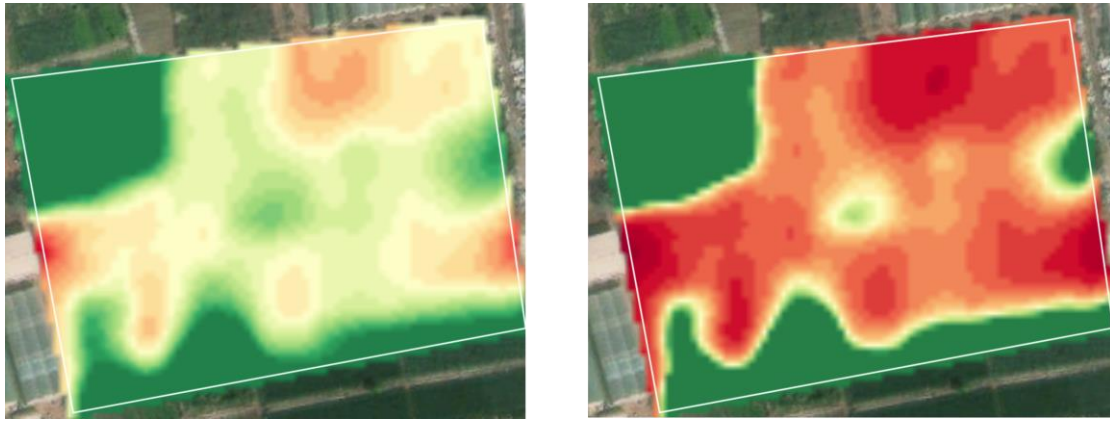


Figure (5.7): Satellite remote sensing images captured through the EOS crop monitoring platform. Left image represents the NDVI index, and the right image represents the RECI index.

On the other hand, data is either stored in raw format or for processing (analytics, querying, visualizations). The primary data originates from several sources such as sensors, satellite images, and legacy files. Initially, it is kept in Google Cloud Storage, as the Data Lake, for the purpose of archiving and preliminary processing. The metadata, transactional data, and other information collected by AgroBahrain PWA forms are stored in Google Cloud SQL using the MySQL database. Nevertheless, the AES data will be stored in Google Firestore, a dynamic NoSQL document database. Cloud SQL has comprehensive security features and flexible SQL commands, but it has a less intuitive interface and does not offer application compatibility. In contrast, Google BigQuery, which is a data warehouse, offers a great degree of adaptability, rapidity, and dependability to enable data analytics, machine learning, and visualisations. Stakeholders will be able to use the PWA interface to create customised datasets, which can be downloaded or stored directly using Google BigQuery.

The simulation begins by showcasing the essential features of the application that are relevant to farm owners' accounts. This will encompass the registration process, adding new farms and crops, and initiating crop production. Afterwards, the simulation will demonstrate

the inclusion of indirect data sources, specifically emphasising the use of the Pycno soil sensor, as well as the integration of external files and their corresponding procedures. Moreover, a thorough examination of AES functionalities will be conducted, with a focus on behaviours such as examining question lists, submitting questions, and replying to inquiries. The subsequent phase will focus on exhibiting web scraping methods using the (peninsula farms) website. A final section will follow, focusing specifically on data visualisation choices. This component of the simulation will provide a preview of the relevant data that has been identified for a farmer operating on a particular farm, as well as the owner of that farm. This will cover details on farms, currently cultivated crops, and meteorological charts. Furthermore, the simulation will illustrate the process of generating an anonymous dataset using the agriculture data interface, accessible through an active agronomist account.

5.3.1. Registering New Stakeholders

Enrollment serves as a mandatory procedure to establish an operational stakeholder account and access the functionalities within the application. The utilized data schema (Figure 5.8) comprises three tables along with external data retrieved from the IGA.

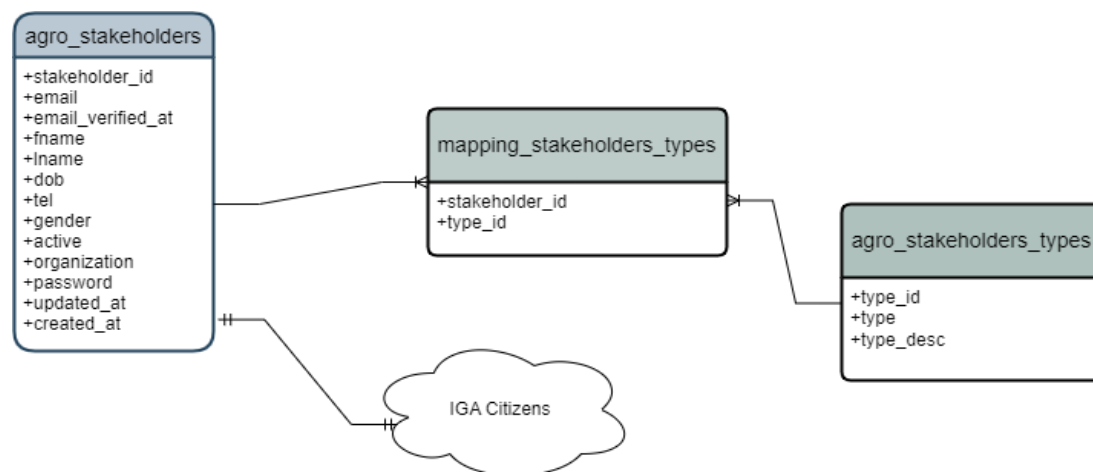


Figure (5.8): Adding new stakeholder data scheme

The first table, 'agro_stakeholders_types,' is designed to maintain a catalog of predefined roles for stakeholders, including 'farm owner,' 'official agent,' and 'agronomist.' The second table, 'agro_stakeholders,' contains demographic information pertaining to the stakeholders who have registered, while the third table, 'mapping_stakeholders_types,' links active stakeholders to their respective roles. These roles are instrumental within the application to control access to its features.

The procedural workflow starts with the entry of preliminary user information, including the CPR, block number, and account type, through a registration form (Figure 5.10). Following the user's selection to register, a GCF ('agro_add_stakeholder') is initiated, which executes two actions in a sequential manner.

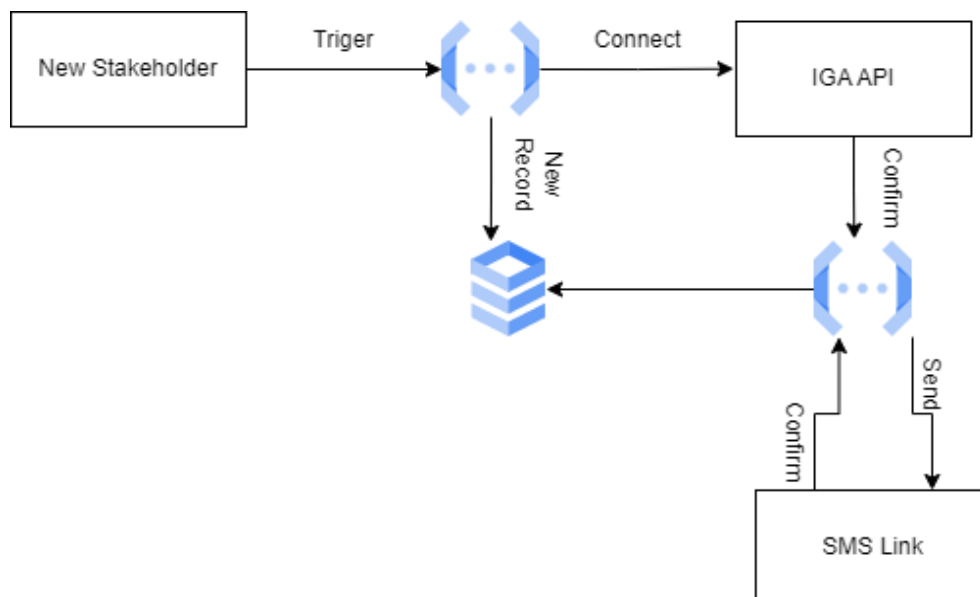


Figure (5.9): Adding new stakeholder pipeline using GCP tools

Figure (5.10): New user registration form

First, it will create a new entry in the 'agro_stakeholders' table, setting the default value of the 'active' field to zero based on the provided details. Subsequently, the system will transmit the data to the IGA API, utilizing the simulated API endpoint "https://domain_name.com/agrofiles/api/get_citizen_resident.php". If the request is successfully acknowledged, the API's response will supply additional user details, encompassing information such as the individual's name, occupation, and date of birth (DOB). The cloud function will then proceed to update the relevant entry in the 'agro_stakeholders' table, setting the 'active' field to 1, dispatch a confirmation SMS, and, if available, send an email to the user. Additionally, it will insert a new record in the 'mapping_stakeholders_types' table to establish the user's association with a stakeholder role. Succeeding to these processes, the stakeholder will be fully enabled to utilize the application.

5.3.2. Farm owner features

The management of agricultural data by farm owners involves a comprehensive range of activities, both directly and indirectly, facilitated

through a system of interconnected operations, as illustrated in Figure 5.11.

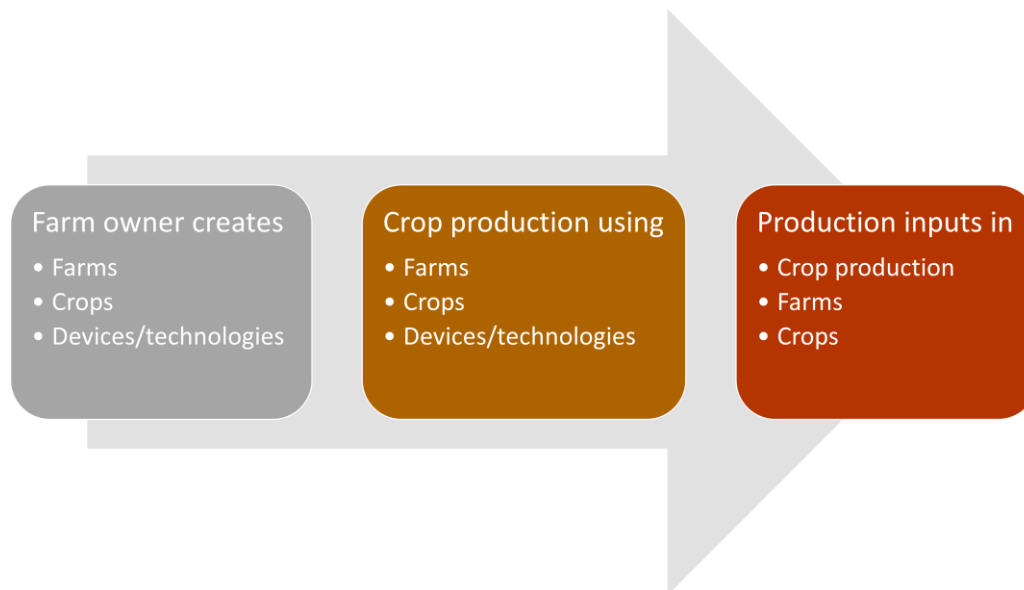


Figure (5.11): Farm Owner Interlinked Operations

Upon activation of the account, farm owners gain the capability to initiate a structured workflow for data generation. This includes tasks such as adding farms and crops, registering external data sources, and establishing a cycle for crop production. Furthermore, farm owners can access and modify all related data associated with their farms, crops, and information gathered through owned technologies. They also possess the ability to review externally provided data, such as market statistics, weather forecasts, and agricultural advice and methods made available via the AES interface. The functionalities allowing simulated farm owner actions and the manipulation of accessible data are visualized in Figure 5.12.

Conversely, farm owners have the authority to assign registered farmers to specific farms and delegate particular responsibilities, such as the documentation and updating of production inputs. Additionally, an assigned farmer is granted access to view farm-specific data, crop details, production information, and device-generated data.

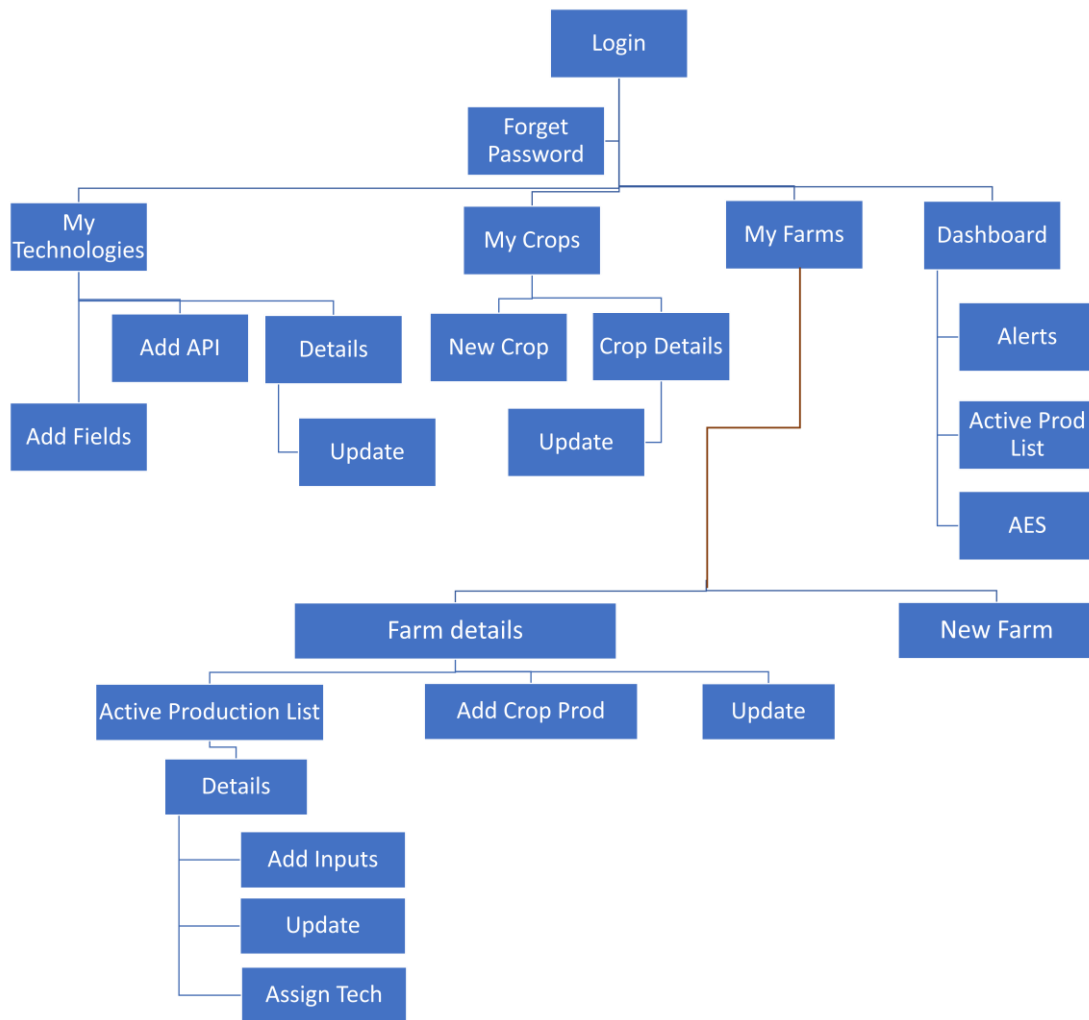


Figure (5.12): Simulated farm owner features. Emphasized farms features to include adding and updating crop production.

The simulation consists of essential independent components such as farms and crops, as shown in Figure 5.12. These components are derived from several data sources, including both direct and indirect channels. The data is obtained by automatic and human inputs, as well as the AES. Nevertheless, these entities are all strongly interconnected with the stakeholder ID. Farms and crops are fundamental data elements that are essential for establishing relationships with other attributes. These connections are enabled through the corresponding fields, 'farm_id' and 'crop_id'.

The figure displays two side-by-side screenshots of the AgroBahrain web application interface. Both forms feature the 'AgroBahrain' logo at the top, which includes a green palm frond icon. The left form, titled 'اضافة محصول' (Add Crop), contains a text input for 'اسم المحصول' (Crop Name), a text area for 'وصف المحصول' (Crop Description), and a blue button labeled 'اضافة المحصول' (Add Crop). The right form, titled 'اضافة مزرعة' (Add Farm), includes a text input for 'اسم المزرعة' (Farm Name), a text input for 'مساحة المزرعة بالهكتار' (Farm Area in Hectares) with a note '* ١ هكتار = ١٠٠٠٠ متر مربع', a location picker for 'موقع المزرعة' (Farm Location), a text area for 'يجب التواجد داخل المزرعة للحصول على البيانات بدقة' (Must be inside the farm to get data accurately), a dropdown for 'المحافظة' (Governorate), and a blue button labeled 'اضافة المزرعة' (Add Farm).

Figure (5.13): Adding Farms and crops.

In the case of technologies and devices, the user starts by inputting core information such as name, type, and description. Subsequently, via the technology details page, the user can furnish in-depth data field specifications for each technology, along with associated API endpoints containing relevant credentials.

5.3.3. Indirect Data Sources

Indirect data sources encompass two primary categories: external data files manually integrated into the system (comprising legacy data and manually collected data) and automated data sources (such as wireless soil sensors and UAVs).

The process of integrating external files involves enabling active stakeholders to manually upload data files to Google Cloud Storage via Firebase functionalities and GCFs. However, only specific file formats—namely, Excel, CSV, or images (whether captured or from a gallery)—are accepted for simulation purposes. Notably, CSV and Excel files must adhere to a predefined structure, prohibiting in-file metadata and necessitating the inclusion of headers for all fields. An example illustrating rejected/accepted CSV files is depicted in (Figure 5.14). For

instance, if uploaded in its current state, a dataset containing average monthly temperature data spanning from 1901 to 2021 would be rejected due to additional metadata (such as variables and country) and a column lacking a header (year). Consequently, any file must undergo an essential transformation before being stored.

Variable: :tas
Country: Bahrain

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1901	15.56	17.41	21.94	26.53	31.60	33.19	35.24	34.71	33.84	29.56	24.57	20.45
1902	17.56	18.61	23.30	26.63	30.10	33.19	33.54	34.01	32.14	27.79	22.67	17.57
1903	14.99	16.16	19.64	23.73	30.00	32.89	33.54	35.31	32.64	27.26	21.83	17.57
1904	16.92	17.11	20.44	24.73	29.00	31.89	34.14	33.81	31.34	28.06	24.87	17.97
1905	15.46	15.81	19.24	25.17	29.30	32.39	33.94	34.33	32.24	29.16	24.07	17.77
1906	15.72	16.48	19.64	23.53	30.10	31.79	33.64	33.21	31.24	28.56	24.67	19.69
1907	16.66	15.86	20.04	23.93	30.30	32.09	34.04	33.61	31.61	28.86	21.97	18.07
1908	16.52	16.03	20.74	25.93	29.33	33.59	33.47	34.41	32.84	28.56	24.17	18.97
1909	15.94	18.76	22.14	26.33	30.70	31.89	34.54	34.41	32.04	28.66	24.07	18.37
1910	15.66	17.11	18.74	23.67	29.30	31.29	33.94	34.46	31.94	27.46	23.07	16.57
1911	12.86	14.70	18.24	23.63	28.68	30.69	34.02	33.38	30.74	27.66	22.75	19.19
1912	15.72	17.11	20.70	24.40	28.90	33.39	33.44	33.01	31.64	28.06	23.95	17.57
1913	15.86	15.51	18.54	25.63	30.48	31.69	33.12	32.28	31.90	28.56	23.57	18.47
1914	18.06	17.01	20.64	25.73	29.68	31.29	33.14	33.78	31.04	29.56	24.47	18.47
1915	17.66	18.01	21.24	24.93	29.43	34.19	33.67	34.31	33.04	27.79	23.97	18.27
1916	16.26	15.71	22.14	23.85	30.60	33.09	34.94	34.71	31.36	26.76	24.17	18.77
1917	16.46	18.01	19.37	26.22	30.41	33.84	30.26	24.17	15.67	12.86	14.70	18.24
1918	16.46	17.21	19.27	23.95	28.93	30.79	33.84	34.11	32.14	29.96	24.67	19.17
1919	18.44	18.51	20.44	26.23	29.33	31.09	33.47	34.41	32.44	29.66	24.97	19.37
1920	17.06	15.31	22.04	25.33	30.50	33.19	33.57	34.41	33.04	30.26	24.17	15.67
1921	16.56	17.01	21.84	27.70	29.13	32.89	35.24	34.51	32.94	28.26	23.97	17.77
1922	17.76	17.41	20.44	25.43	29.13	33.79	34.54	34.91	31.94	28.56	25.55	19.57
1923	16.76	16.81	22.24	23.85	29.60	32.99	34.64	33.13	31.64	28.96	22.70	19.47
1924	17.94	17.11	23.70	25.63	30.90	33.99	34.94	34.01	32.24	30.26	24.47	17.97
1925	13.46	14.61	22.32	26.43	29.90	31.62	32.97	34.51	33.04	29.66	23.57	19.57
1926	17.76	18.71	21.14	23.85	30.20	33.09	33.67	34.31	31.06	28.36	25.07	18.27
1927	17.06	16.33	20.24	24.25	31.20	33.69	34.64	34.01	33.44	30.16	24.77	17.60

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1901	15.96	17.41	21.94	26.53	31.60	33.19	35.24	34.71	33.84	29.56	24.57	20.45
1902	17.56	18.61	23.30	26.63	30.10	33.19	33.54	34.01	32.14	27.79	22.67	17.57
1903	14.99	16.16	19.64	23.73	30.00	32.89	33.54	35.31	32.64	27.26	21.83	17.57
1904	16.92	17.11	20.44	24.73	29.00	31.89	34.14	33.81	31.34	28.06	24.87	17.97
1905	15.46	15.81	19.24	25.17	29.30	32.39	33.94	34.33	32.24	29.16	24.07	17.77
1906	15.72	16.48	19.64	23.53	30.10	31.79	33.64	33.21	31.24	28.56	24.67	19.69
1907	16.66	15.86	20.04	23.93	30.30	32.09	34.04	33.61	31.61	28.86	21.97	18.07
1908	16.52	16.03	20.74	25.93	29.33	33.59	33.47	34.41	32.84	28.56	24.17	18.97
1909	15.94	18.76	22.14	26.33	30.70	31.89	34.54	34.41	32.04	28.66	24.07	18.37
1910	15.66	17.11	18.74	23.67	29.30	31.29	33.94	34.46	31.94	27.46	23.07	16.57
1911	12.86	14.70	18.24	23.63	28.68	30.69	34.02	33.38	30.74	27.66	22.75	19.19
1912	15.72	17.11	20.70	24.40	28.90	33.39	33.44	33.01	31.64	28.06	23.95	17.57
1913	15.86	15.51	18.54	25.63	30.48	31.69	33.12	32.28	31.90	28.56	23.57	18.47
1914	18.06	17.01	20.64	25.73	29.68	31.29	33.14	33.78	31.04	29.56	24.47	18.47
1915	17.66	18.01	21.24	24.93	29.43	34.19	33.67	34.31	33.04	27.79	23.97	18.27
1916	16.26	15.71	22.14	23.85	30.60	33.09	34.94	34.71	31.36	26.76	24.17	18.77
1917	16.46	18.01	19.37	26.22	30.41	33.84	30.26	24.17	15.67	12.86	14.70	18.24
1918	16.46	17.21	19.27	23.95	28.93	30.79	33.84	34.11	32.14	29.96	24.67	19.17
1919	18.44	18.51	20.44	26.23	29.33	31.09	33.47	34.41	32.44	29.66	24.97	19.37
1920	17.06	15.31	22.04	25.33	30.50	33.19	33.57	34.41	33.04	30.26	24.17	15.67
1921	16.56	17.01	21.84	27.70	29.13	32.89	35.24	34.51	32.94	28.26	23.97	17.77
1922	17.76	17.41	20.44	25.43	29.13	33.79	34.54	34.91	31.94	28.56	25.55	19.57
1923	16.76	16.81	22.24	23.85	29.60	32.99	34.64	33.13	31.64	28.96	22.70	19.47
1924	17.94	17.11	23.70	25.63	30.90	33.99	34.94	34.01	32.24	30.26	24.47	17.97
1925	13.46	14.61	22.32	26.43	29.90	31.62	32.97	34.51	33.04	29.66	23.57	19.57
1926	17.76	18.71	21.14	23.85	30.20	33.09	33.67	34.31	31.06	28.36	25.07	18.27
1927	17.06	16.33	20.24	24.25	31.20	33.69	34.64	34.01	33.44	30.16	24.77	17.60

Figure (5.14): Sample of rejected/accepted CSV file.

Stakeholders will be facilitated in uploading data files via the "add data file" form (as shown in Figure 5.15) by furnishing the required metadata.

اضافة ملف بيانات

نوع البيانات

تربة

المزرعة

المحرق

المصدر

طماطم

نوع القياس

درجة Celsius

نوع القياس

(%) Percent

ارفاق ملف

No file chosen [Choose File](#)

[ارسل](#)

اضافة ملف بيانات

AgroBahrain

نوع البيانات

تربة

المزرعة

المحرق

المصدر

طماطم

نوع القياس

درجة Celsius

نوع القياس

(%) Percent

ارفاق ملف

No file chosen [Choose File](#)

[ارسل](#)

Figure (5.15): The user interface for uploading data files manually.

The data collected will be subsequently inserted into the meta table (agro_manual_files) and (agro_manual_fields) through invocation of the GCF (agro_manual_source), thereby inserting the complete file path.

Data files are uploaded directly through the app without any backend scripts intervention. This uploaded file will be stored in a Cloud Storage folder named after the stakeholder ID within the (agro_manual_upload) bucket, utilizing Firebase storage tools (as depicted in Figure 5.17).

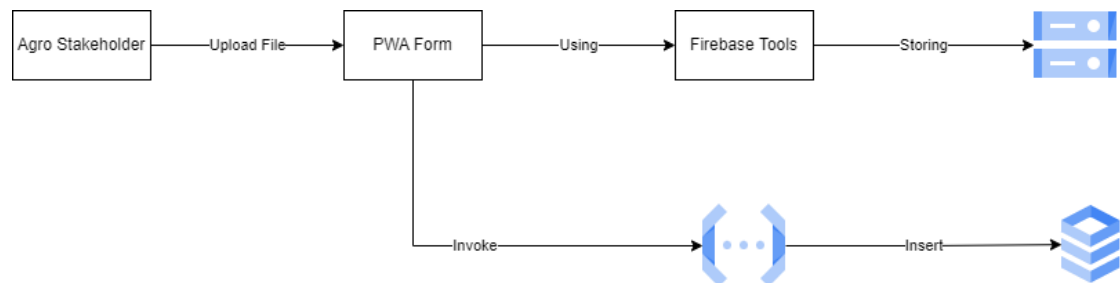


Figure (5.16): Uploading data files manually pipeline using GCP tools

gs://agro_manual_upload > 148523221			
☐	Name	Size	Type
☐	 1681904739870_tas_timeseries_monthly_cru_1901-2021_BHR.csv	9.18 KB	text/csv

Figure (5.17): Sample storage bucket output

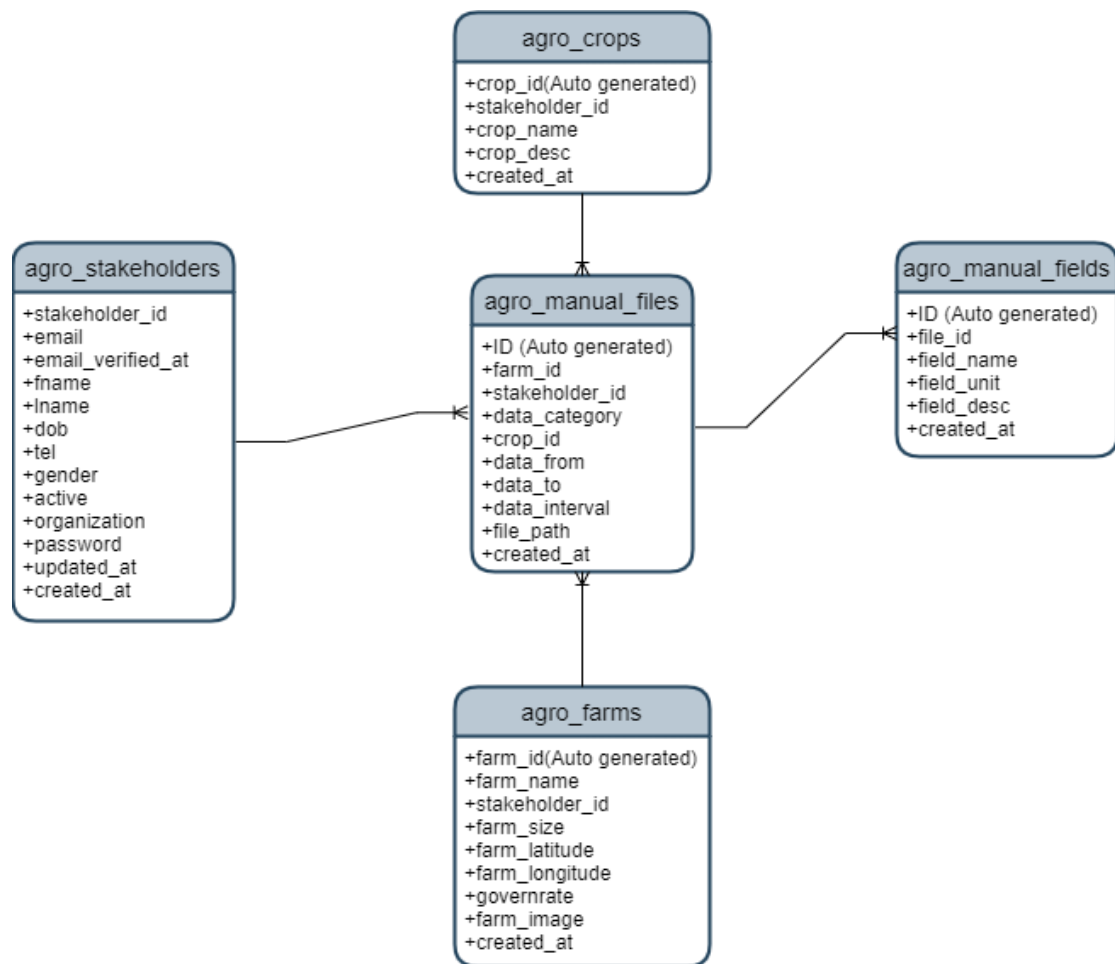


Figure (5.18): The add data file (manual upload) data scheme.

On the other hand, integrating an automated data source represents a workflow (see Figure 5.19) that impacts three entities (agro_tech, agro_tech_apis, agro_tech_fields).

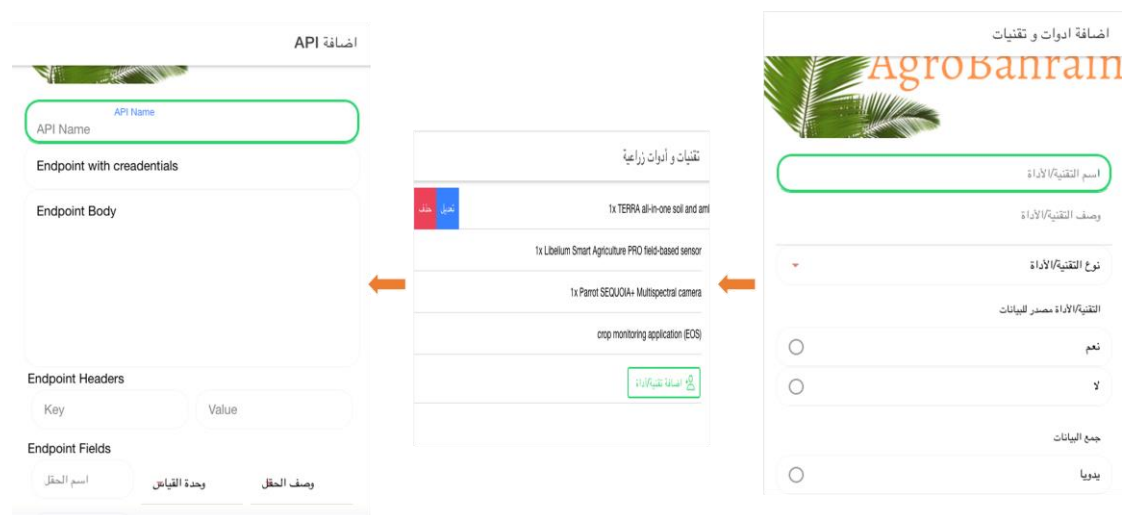


Figure (5.19): Adding automated data sources - interfaces sequence.

To initiate this process, the stakeholder must initially input basic metadata concerning the technology (such as soil sensors, remote sensing platforms, etc.)—including the technology's name, description, and type—using the add tool/tech PWA form. This form enables the registration of devices and technologies that allow remote data collection or manual upload. Once the technology is registered, it will be listed under the agriculture technologies and tools page.

The subsequent step involves the user editing the technology details to finalize the registration process. Within this step, the user utilizes the edit option for each technology to register API endpoints through the "add API" form, inputting the required response fields. The stakeholder is mandated to input the exact field names as presented in the endpoint response, specify the data type for each field, and provide a description using a predefined list (such as Temperature, Humidity, Soil Temperature, etc.).

Adding an API as a data source triggers updates in two entities: `agro_tech_apis`, which records the API URL, header, and description, and `agro_tech_fields`, which stores the API fields and their respective units. Given that registered technologies may possess multiple APIs, the data fields are interconnected using both (`api_id`, `tech_id`) as foreign keys for the `agro_tech_fields` table. The registration procedure triggers a Cloud Function (`agro_init_auto` – refer to Appendix C) that generates a new folder in the storage bucket (`agro_auto_extract`) with the stakeholder ID, if it does not already exist. Afterwards, it creates a new CSV file named (`techid_apiid`) with the response fields as headers, along with an extra header called 'created_at' to record the timestamp.

Two technologies will be recorded for the purpose of simulation. The online weather platform utilises the real-time weather API (https://api.weatherapi.com/v1/current.json?key=user_key&q=Bahrain) to

provide current weather information. It requires a GET request with two mandatory parameters: the user API key and the country or city. The response includes the temperature in Celsius, wind speed, pressure, humidity, and precipitation. The second technology is the Pycno soil sensor, which consists of a single API with three endpoints that utilise the GET request. The initial endpoint retrieves a collection of sensors, the subsequent endpoint exhibits the most recent recorded data (such as temperature, humidity, soil temperature, etc.) from the specified URL (https://portal.pycno.co.uk/api/v2/data/no/sensorID.json?TK=user_token), and the final endpoint necessitates manual intervention since it requires different parameters each time it is triggered.

It is important to note that registered automated data sources remain inactive and do not collect data within the system unless they are allocated to active crop production.

5.3.4. Crop Production

Crop production refers to the intentional allocation of cultivated crops within a farming environment, designating an active status for their production. This feature allows for the activation of pre-registered automatic resources and provides for the manual input of used resources, such as nutrients and pesticides, using PWA forms. Moreover, it enables farm owners and agricultural labourers to monitor the production status remotely and instantaneously using the PWA interface.

The simulation primarily focuses on three primary operations: firstly, creating a production entity using the "add crop production" form (see Figure 5.20). Subsequently, connecting an automated source to the production and specifying how the backend scripts automate data collection from registered devices, such as IoT soil sensors, irrigation systems, and satellite remote sensing. Lastly, previewing data related to soil moisture collected through the Pycno soil sensor.

The figure displays two screenshots of the 'Add Crop Production' form in the AgroBahrain application. The left screenshot shows the 'Crop' dropdown menu with options: 'Traditional Farming', 'Soilless Farming', and 'Soilless Farming - Water'. The right screenshot shows the 'Crop Type' dropdown menu with options: 'Traditional Farming', 'Soilless Farming', and 'Soilless Farming - Water'. Both screenshots show a calendar for February 2023 and a date picker for the expected harvest date (November 2023).

Figure (5.20): Adding crop production form.

The workflow initiates with the addition of new production, contingent upon the availability of farms and crops. A successful transaction results in the creation of a new entry within the "crop_production" table. This entry is assigned a production ID, and its production status is designated as 1, denoting an active state (where 0 signifies inactive and 2 stands for completed). All productions in an active or completed state are viewable on the main crops page (refer to Figure 5.21).

The subsequent phase involves activating automated sources by choosing an available active product, triggering the display of a production details page. This page exhibits basic production information such as crop name, farm name, and farming type, alongside real-time data sourced from automated systems like soil moisture (refer to Figure 5.22). Within this interface, stakeholders are provided with options to activate a source or manually input production elements like fertilizers and pesticides.

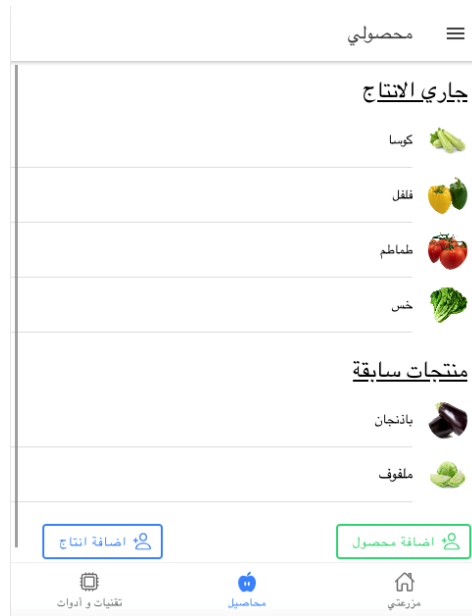


Figure (5.21): Crops main page. Active production and completed.

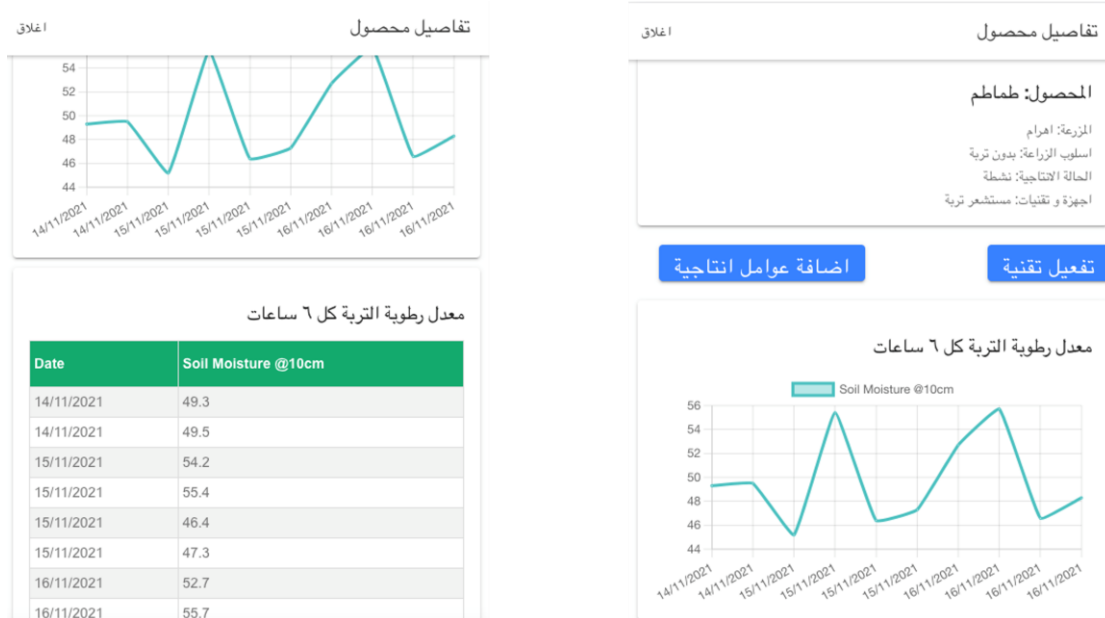


Figure (5.22): Basic data visualization for active production

Once a technology is assigned to an active crop production by a stakeholder, it becomes operational and is recorded in the catalogue of endpoints automatically activated via the Cloud Function (see Appendix C, section "agro_auto_extract" for further information). Due to the automated process's indifference to the variability of data source update notifications and its reliance on incremental data extraction and loading, the Cloud Function is executed periodically at 6-hour intervals via the

Google Cloud Scheduler. The purpose of this schedule is to gather data and append it to designated target CSV files.

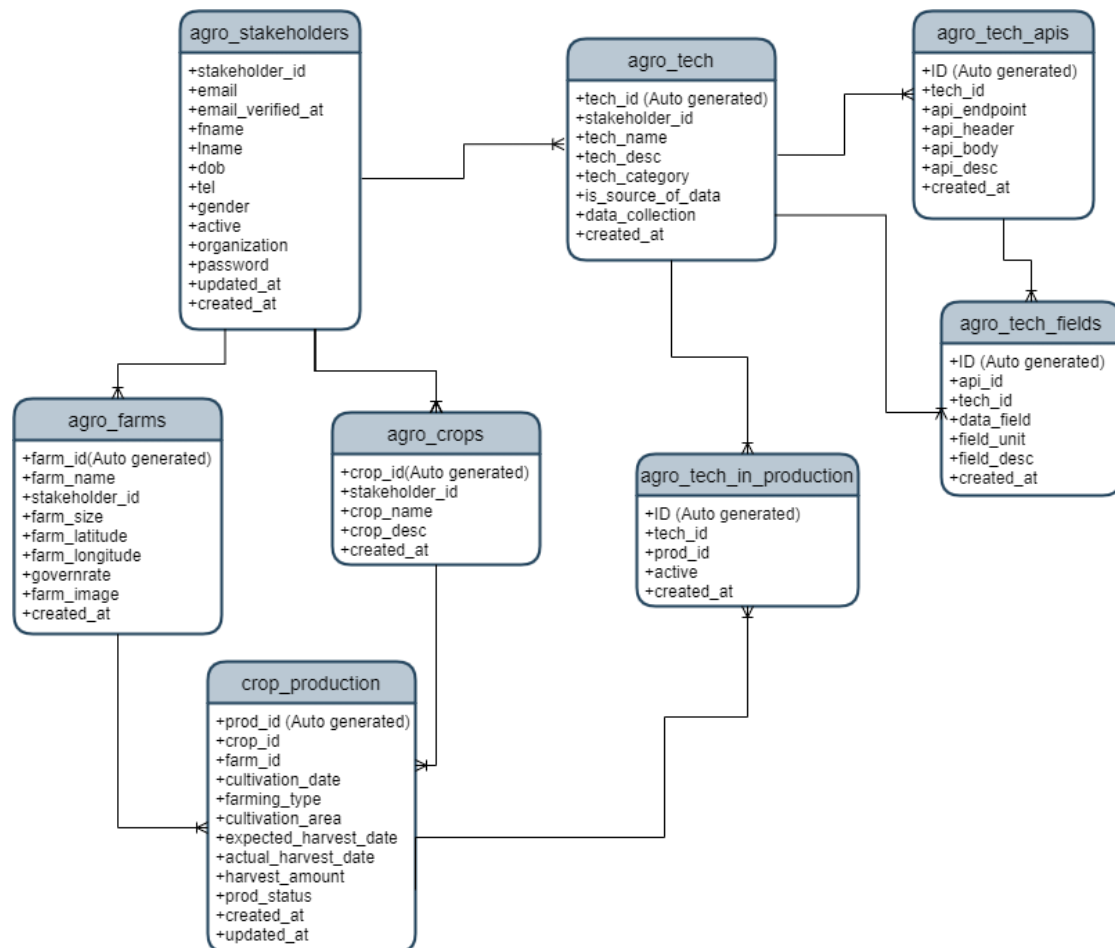


Figure (5.23): Crop production and activating automated data source data scheme

Activated technologies are recorded in the "agro_tech_in_prod" table within the data schema (depicted in Figure 5.23), marked with an active status (1). When the production cycle concludes, the technology's status is not removed but rather updated to inactive (0). The Cloud Function scans the table during every invocation for active technologies, subsequently triggering related endpoints and capturing their responses. These captured responses are appended to the assigned target CSV files.

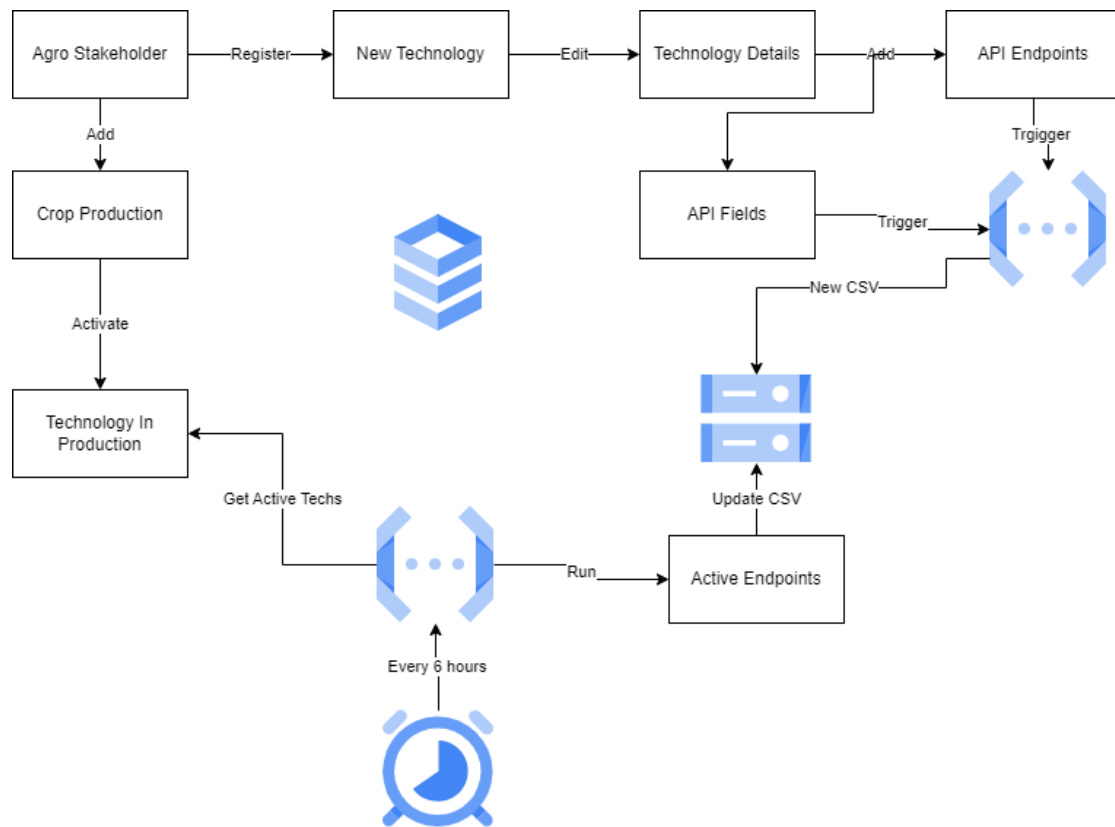


Figure (5.24): Pipelines for crop production with active data sources.

5.3.5. Agriculture Enquiries Service (AES)

The Agricultural Enquiry Service is an extension service integrated into the application, aimed at providing support to various stakeholders by facilitating the exchange of knowledge and experiences through a question-and-answer interface. This interface allows users to submit their queries in various formats, including text, images, files, and voice recordings. The response interface also offers these same format options. Additionally, user posts are categorized based on factors such as crop type, soil conditions, and fertilizers, enabling users to filter and search for specific agricultural practices and guidance.

To store both enquiries and answers and ensure flexibility in handling different data formats, the system utilizes the NoSQL Google Cloud Firestore document database. Firestore is a database service provided by Google Cloud Firebase, available to users with valid Google accounts within certain usage limits. The database organizes data by creating

collections, which are analogous to tables in a traditional database, and allows users to add records as documents within these collections. Each document is assigned a unique identifier, either auto-generated or custom, and data is stored in the form of key-value pairs.

Active stakeholders with registered accounts have access to three main components: the inquiries list, the ability to add new inquiries, and detailed information about existing inquiries. The enquiries list, as depicted in Figure 5.25, serves as the central interface for displaying all enquiries, both answered and unanswered, with built-in search options for refining the results. Users can click on a specific enquiry to view its details, provide answers, or, in the case of inquiries they own, edit, or delete them.

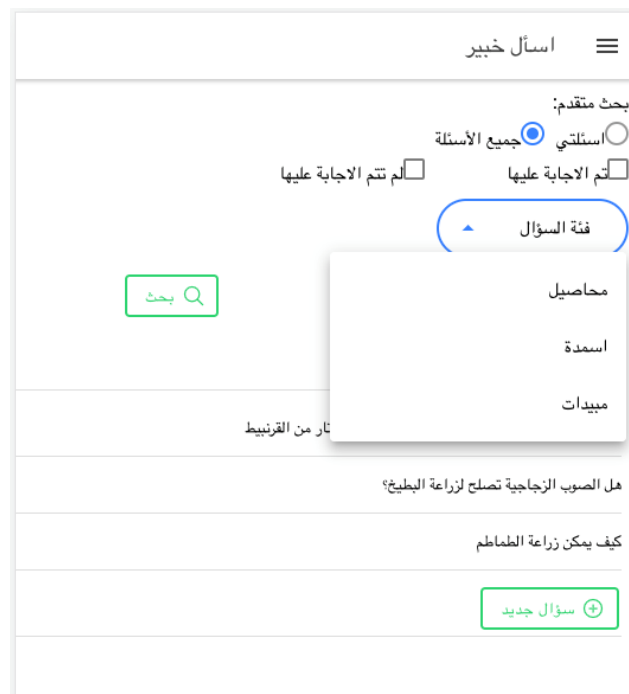


Figure (5.25): Enquiries main page (list of enquiries).

Within the enquiries list entity, users can initiate the creation of a new enquiry by clicking the "New Enquiry" button, as illustrated in Figure 5.26. This action offers the options to input text, capture images of crops or leaves, or attach files while specifying the enquiry's category.

سؤال جديد

AgroBahrain

فئة السؤال

السؤال

التقاط صورة

ارفاق ملفات

No file chosen Choose File

ارسال

Figure (5.26): Add new enquiry form

Images and files are stored in a Google Cloud Storage bucket known as "AES." Each stakeholder is allocated a dedicated folder within this bucket, named after their unique stakeholder ID. The complete path to the stored file is added as a document field in the Firestore collection. Two Firestore collections, "aes_questions" and "aes_answers," are employed. The questions collection includes fields for document ID (auto-generated), stakeholder ID, answered status (yes/no), the enquiry text, file path, and creation date (in epoch format), as depicted in Figure 5.27. In contrast, the answers collection employs document IDs derived from the questions collection, followed by an answer number to facilitate multiple responses to a single enquiry. The "answered by" field records the stakeholder ID responsible for providing the answer.

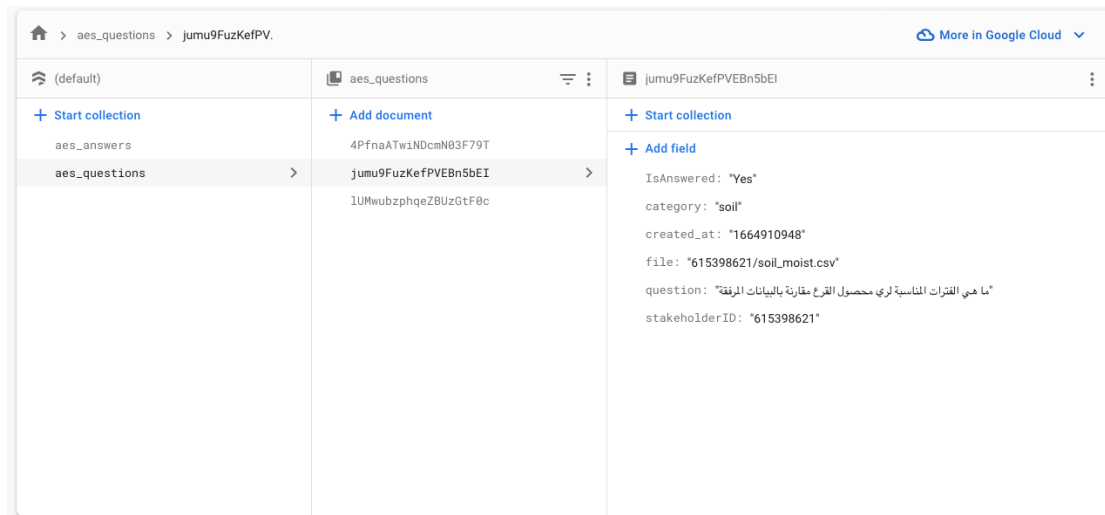


Figure (5.27): Sample of stored enquiry as a NoSQL document on Google Firestore (key value pairs)

5.3.6. Web Scrapping

For the purpose of simulation, the public pricing information for specific vegetables available on a Bahraini website (<https://peninsulafarms.com/product-category/vegetables-fruits/>) (refer to Figure 5.28) will be systematically extracted for demonstration.

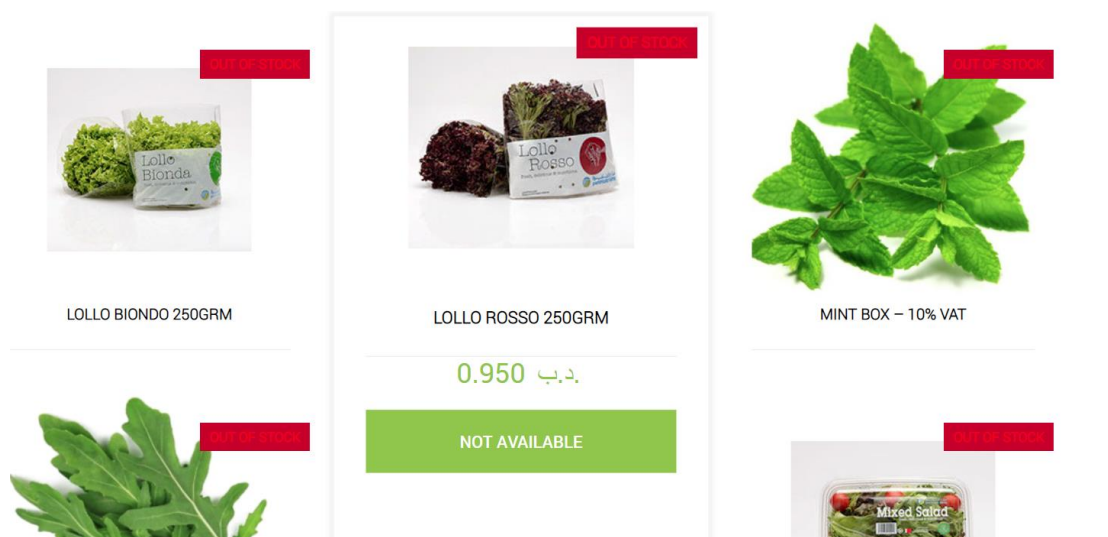


Figure (5.28): The target webpage used in web scrapping.

Python offers a diverse range of libraries suitable for web scraping, such as selenium, regular expressions, BeautifulSoup, and Lxml. The simulation methodology does not encompass a PWA interface but instead relies on Google Cloud Scheduler to activate a Python script (refer to

Appendix C). This script is designed to retrieve pricing data from the designated webpage and subsequently store it in a CSV file within a designated Cloud Storage bucket titled "web_scrap". The scheduling mechanism will execute the function on a daily basis to capture any alterations in the prices of the listed products on the target webpage.

The process involves a manual exploration of the website to pinpoint the specific webpage and its associated fields. The identified target page pertains to the section dedicated to 'vegetables & fruits,' and its structure is examined using the 'inspect' function within the Google Chrome browser to identify the HTML tags containing the relevant data fields. The script will utilize the BeautifulSoup library to isolate two specific tags, namely 'h2' and 'bdi,' which respectively encapsulate the product names and their corresponding prices. The accumulated pricing information will be organized within a data frame alongside the product names. Subsequently, the script will transform this data frame into a CSV file, facilitating its upload to the designated storage bucket.

5.3.7. Data Extraction and Presentation

Extracted data is presented in real-time utilizing visualization tools such as tables and charts (see Figure 5.22 in section 5.3.4) or by producing datasets for analytical, predictive, and other research purposes. The application facilitates both active stakeholders and guest users to generate datasets in various formats, including CSV, Excel, and images, accessible through the interface referred to as the Agriculture Data Platform (see Figure 5.29). This interface not only showcases the latest agricultural statistics but also provides users with diverse filter tools to extract specific data.

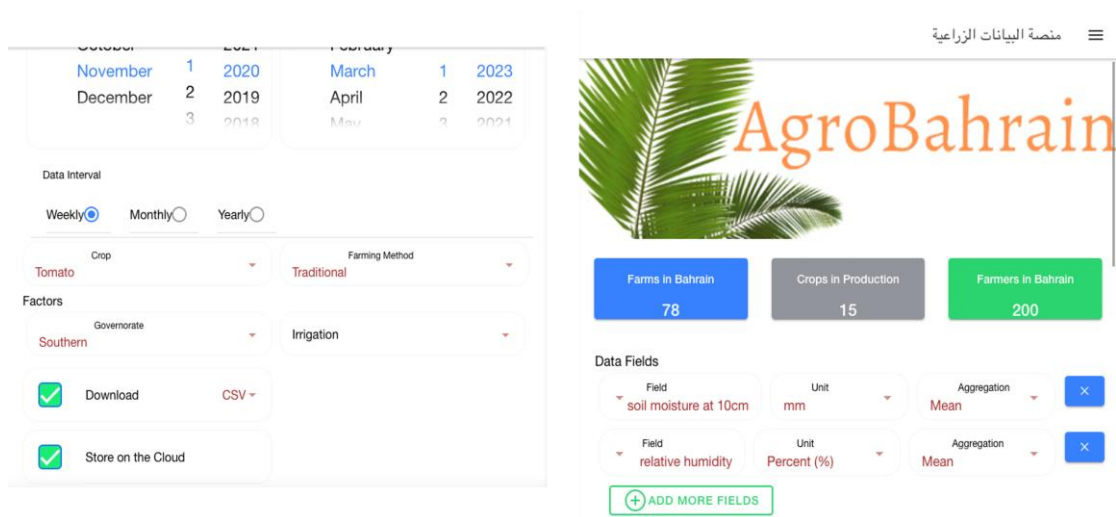


Figure (5.29): Agriculture data interface, displays statistical data along the ability to extract custom data sets.

In the simulation scenario, an agronomist with an active stakeholder account will request data concerning the average weekly humidity and soil moisture at a depth of 10 cm for tomato production on open farms (utilizing traditional farming methods) between November 1, 2020, and March 1, 2023. These requested factors will be limited solely to the farm's geographical location. Factors encompass both internal and external details impacting field values, such as geographical location, irrigation methods, and fertilizers. For example, soil temperature variations are contingent on irrigation time and amount, differing from one farm to another. Factors (inputs) are recorded during crop production either manually via PWA forms or automatically through technologies such as soil sensors integrated into active production.

Following the constraint to solely southern governorate farms (as a location factor), the agronomist must choose a storage method for the dataset. The default selection for downloading in CSV format is supplemented by an additional option to store the dataset in the cloud. Cloud storage is executed via the Google BigQuery service, ensuring a reusable copy of the dataset for further analysis and visualization by other users.

i. Extract, Transform, and Load

The initial step involves the selection of fields with descriptions matching those in the request from specific entities (agro_manual_fields, agro_tech_fields). This extraction process includes capturing requested crop information, farming methodologies, and data within the specified timeframe from entities such as crop_production, agro_tech_in_prod, and agro_crops. Furthermore, the search query is configured to confine outcomes to farms situated in the southern governorate by utilizing the agro_farms entity.

Subsequently, upon user initiation by clicking the 'generate' function, the script (agro_data_etl) executes a series of operations in a systematic order:

1. Extraction of target files' names and paths along the creation of a data frame is the initial step, with headers comprised of the requested fields collected from various sources, alongside an additional timestamp field. This simulation involves two sources: manually uploaded data files and data automatically uploaded via technologies and devices. Given that manually collected data may include historical or secondary information and their scheme is different, all files even with duplicates will be selected utilizing the queries in Figure 5.30. The first query addresses automated data sources employing the data scheme represented in Figure 5.31, and the second deals with manual data files following the data scheme provided in the indirect sources section (Figure 5.18).

```

SELECT atf.data_field, CONCAT_WS ('_',atf.tech_id, atf.api_id) AS file_name From agro_tech_fields atf JOIN
agro_tech_in_prod atp ON atf.tech_id = atp.tech_id JOIN crop_production cp ON atp.prod_id = cp.prod_id
JOIN agro_crops ac ON cp.crop_id = ac.crop_id JOIN agro_farms af ON cp.farm_id = af.farm_id AND
af.stakeholder_id = "stakeholder_id" AND ac.crop_name = "crop_name" AND atf.field_unit = "field unit" AND
atf.field_desc = "field description" AND af.governorate = "governorate" AND cp.farming_type = "farming
method" AND cp.created_at BETWEEN (date_from , date_to);

SELECT amf.field_name, am.file_path FROM agro_manual_fields amf JOIN agro_manual_files am ON
amf.file_id = am.ID JOIN agro_crops ac ON am.crop_id = ac.crop_id JOIN agro_farms af ON
am.farm_id = af.farm_id AND am.stakeholder_id = "stakeholder_id" AND ac.crop_name = "crop_name" AND
af.governorate = "governorate" AND am.date_from >= "date from" AND am.date_to <= "date to" AND
amf.field_unit = "field unit" AND amf.field_desc = "field description";

```

Figure (5.30): The queries used to extract the data based on the provided filters.

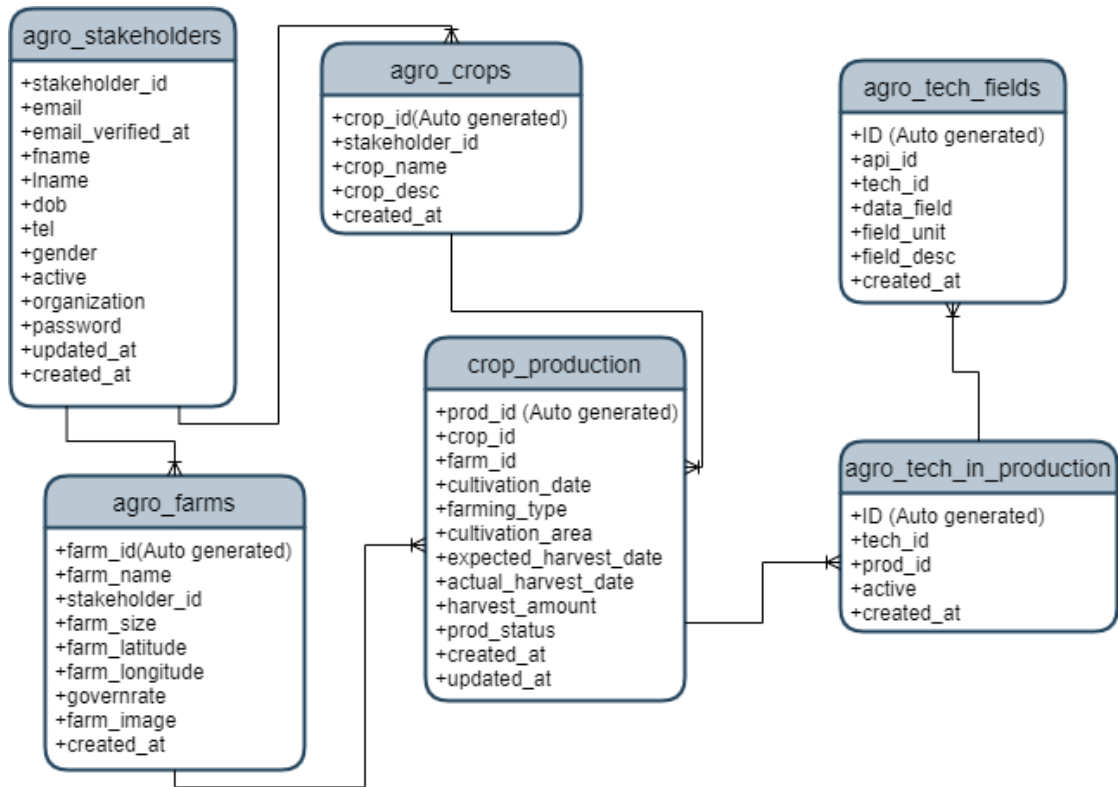


Figure (5.31): The data entities required for autonomus sources.

- The script iterates over files stored in Cloud Storage that contain similar fields based on the stakeholder ID, file path, or file name. The retrieved data is compiled and inserted into the data frame, structured akin to Figure 5.32.

Rel_humidity	humid	S1	Soil_moist_10	Created_at
66		52.3		2022-02-12 11:00:00
	68	51.2	51.1	2022-02-12 17:00:00
52	54	49.6		2022-04-15 10:00:00
	70		48.4	2022-10-12 14:00:00
48		43		2023-01-05 11:00:00
39		37.6		2023-02-01 11:00:00

Figure (5.32): Retrieved data frame prior transformation.

- Following the iteration and successful data insertion into the data frame, various transformation operations are executed to process and refine the data. Similar fields are merged or averaged under a single column with their descriptions as headers. If similar fields possess values at the same time stamp and exhibit considerable differences, the data will be merged into distinct record numbers (Figure 5.33). Conversely, if the differences are marginal, values will be averaged and then merged under the description column. Null values and outliers are omitted from this process.

No	Humidity	Soil Moisture at 10 cm	Created_at
1	66	52.3	2022-02-12 11:00:00
2	68	51.15	2022-02-12 17:00:00
3	53	49.6	2022-04-15 10:00:00
4	70	48.4	2022-10-12 14:00:00
5	48	43	2023-01-05 11:00:00
6	39	37.6	2023-02-01 11:00:00

Figure (5.33): Transformed data set.

- The data within the data frame is subsequently aggregated as per the user's specifications and exported to a CSV file for download onto the user's device. Opting to store data on the cloud results in the script creating a new table within the target dataset. In Google BigQuery, a dataset resembles a conventional database, containing a collection of tables and views. A default dataset, 'guest_users,' is created to store

tables added by guest users, while stakeholders have their own datasets formed using their stakeholder ID for future reference. Table names are automatically generated, utilizing formats such as 'guest_user_timestamp' or 'stakeholder_id_timestamp.' Upon successful data insertion into the new BigQuery table, a corresponding text file bearing the same table name and structure is generated in the Cloud Storage bucket ('data_set_metaData'). This text file, acting as metadata or a data dictionary, contains information describing the dataset, such as fields, sample periods, and geo-locations. The information is formatted in JSON or XML to ensure the datasets are discoverable, accessible, and comprehensible to others.

5.4. Summary

In this chapter, the conceptual framework has been introduced, delineating two distinct stages examined across subsequent sections. Section 5.2.1 emphasizes the initial stage by proposing various preparatory requisites to enhance the awareness and preparedness of agriculture stakeholders concerning the adoption and utilization of ICTs. Additionally, sections 5.2.2 and 5.2.3 expound upon the second stage, illustrating the agricultural data lifecycle within the framework and a suggested implementation paradigm.

Conversely, section 5.3 delves into the simulation of diverse platform features to exhibit the viability of the developed framework. The simulation predominantly concentrates on different data processes, such as ETL, by mimicking the subsequent features:

1. Enrollment and activation of stakeholders' accounts.
2. Highlighting farm owner functionalities encompassing farm and crop management along with the crop production cycle.

3. Manual uploading of data files and the methods to register and extract data from automated sources such as wireless sensors and weather stations.
4. Simulating an extension service that dispenses expert knowledge through a question-and-answer PWA interface named Agriculture Enquiries Service.
5. Outlining the process of extracting data from online platforms using web scraping tools; nevertheless, it is recommended to employ this feature under official oversight to circumvent concerns regarding data privacy.
6. Ultimately, leveraging the collected data for analytics and research purposes is realized by creating datasets via a PWA interface accessible to both stakeholders and guest users. This interface facilitates the application of diverse filters on available data to generate and download datasets in formats like CSV and images.

Chapter Six

Conclusion, Recommendations, and Future Work

6.1. Conclusion

ICTs such as Big Data, the IoTs, and cloud computing are widely recognized as significant contributors to the attainment of sustainable development goals. Specifically, those goals pertaining to ensuring access to sufficient and nutritious food (Goal 2), managing water resources (Goal 6), revitalizing biodiversity, and addressing the issue of climate change (Goals 12, 13, and 15).

Over the past decade, there has been a notable increase in the focus on Big Data, as many countries aim to acquire a competitive advantage in this growing sector. Furthermore, the agriculture industry is progressively depending on data-driven methodologies, necessitating the implementation of state-of-the-art technology to efficiently handle and evaluate the large volumes of data gathered. The application of Big Data analytics in agriculture has effectively tackled diverse difficulties related to crop cultivation. It has facilitated water conservation, early detection and control of agricultural diseases and pests, and reduced harmful emissions from certain agriculture practices. Nevertheless, the use of ICTs in agriculture has led to remarkable improvements in the quality of crops and a substantial increase in output. Agricultural Big Data applications will enhance the sustainability of agriculture, decrease natural resource usage, enhance crop production and quality, establish a national repository for agricultural data, improve scientific research in various agricultural fields, and provide enhanced information and guidance for stakeholders in agriculture, with a particular emphasis on farmers.

Big Data, cloud computing, and IoT are disruptive technologies that disclose great potential in the agriculture sector through collecting, storing, analyzing, and visualizing agriculture data for a more precise decision. In agriculture, Big Data applications will accurately detect wind direction, temperature, relative humidity, solar and evaporation, and plant transpiration.

The Kingdom of Bahrain is a non-agricultural country with limited agricultural land, harsh soil conditions, and water scarcity. The government has recognized the power of Big Data to drive economic growth, enhance decision-making, and create competitive advantages in both the public and private sectors. In general, Big Data is significant for improving and sustaining the agriculture sector in the Kingdom of Bahrain. However, Bahrain lacks agricultural data and has a low adoption rate of ICT in agriculture, moreover there is a lack in studies concerning Big Data frameworks that target non-agriculture countries with limited ICT knowledge and awareness among agriculture stakeholders. Therefore, implementing an agriculture Big Data framework for such a situation is complex and challenging.

The main research objective was to build a smart farming platform on a national scale that suits the Bahraini environment. To achieve that, the study was divided into two phases.

Phase 1

The primary goal of Phase 1 was to assess the level of preparedness of the Kingdom of Bahrain in adopting ICT in the field of agriculture. This study primarily examined the digital preparedness of farmers and enterprises, as well as the extent to which ICT has been embraced in the agriculture industry. Furthermore, the existing process of managing agricultural data and determining the parties involved are the fundamental components of the second phase. The findings demonstrated a noticeable

and evident enthusiasm for agricultural technology among individuals. The agricultural advancements in the Kingdom of Bahrain have created an encouraging environment for individuals and farmers, resulting in substantial improvements in the agriculture industry.

In addition, the Kingdom of Bahrain has been at the forefront of digital transformation in the Middle East, implementing it in education and many sectors, such as health care, government services, and the economy.

However, the study revealed that the agricultural sector in the Kingdom of Bahrain had not adopted digital solutions due to various challenges. These challenges include low ICT literacy among farmers, the absence of a curriculum to enhance farmers' IT skills (educational challenges), farmers' lack of awareness about smart agriculture tools used globally, the perception that investing in ICT is not cost-effective (cultural challenges), and the lack of programs promoting the use of ICT tools in agriculture. In addition, there is a lack of communication between farmers, individuals, government activities, and support for developing the agriculture sector.

The study methods employed to examine the farm data lifecycle uncovered several obstacles impeding the advancement of agriculture Big Data in Bahrain. The challenges encompass the lack of sufficient techniques and resources for gathering data, ineffective procedures for storing data, inadequate interfaces and media for accessing data, social and organizational constraints on data sharing, and legal limitations on specific automated technologies employed for data collection, such as UAVs. Furthermore, the absence of ICT expertise and acceptance among those involved in agriculture poses a substantial obstacle to the implementation of Big Data solutions in the agricultural sector.

Conversely, a proposed matrix was established to represent the relationship between data owners and users in Bahrain's agricultural

sector. The matrix facilitated the identification of distinct data domains and access entitlements for players in the agriculture sector.

Phase 2

A high-level conceptual framework has been created based on a thorough study conducted in the agriculture industry. This comprehensive architecture leverages the integration of Big Data, cloud technologies, and a web application. The primary objective of this system is to gather, store, and communicate agricultural data in a somewhat semi-autonomous way. Designed as an inclusive and cooperative platform, it aims to engage the various stakeholders in the field of agriculture. The primary objectives of this framework encompass equipping the agriculture sector (in terms of education, finance, and social aspects) to embrace ICT and expanding the availability of agricultural data to meet the needs of a wide range of beneficiaries. Enhancing measures for safeguarding the privacy of agricultural data. Enhancing the quality and productivity standards of crops, facilitating precise control over agricultural inputs, and promoting the development of a resilient and environmentally friendly agricultural ecosystem.

Furthermore, this framework has been carefully developed to cater specifically to the distinctive attributes of Bahrain while considering obstacles such as small-scale agricultural practices and farmers' low proficiency in ICT. On the other hand, the framework was adeptly designed for scalability to encompass different countries.

Conversely, a simulation model was developed, explicitly targeting the AgroBahrain framework's key elements. This simulation specifically examined the life cycle of farm owners, who are considered significant stakeholders. The data engineering procedures were replicated, encompassing collecting data, transformation, and storage. This simulation serves as an instructive model for constructing a sophisticated

agricultural ecosystem by using cutting-edge technology like Big Data and cloud services.

The simulation demonstrated the ability to develop an agricultural platform enabling seamless data intake and processing. This platform possesses the capacity to augment agricultural sustainability and guarantee food security in Bahrain significantly. The AgroBahrain framework leverages the GCP to make use of adaptable cloud services that are readily available and can scale as required. GCP's serverless cloud services facilitated the autonomous execution of all simulation model operations, eliminating the need for external services.

6.2. Contribution

ICTs serve as the primary instruments for developing intelligent agricultural solutions. It enhances agricultural decision-making by gathering and analyzing different data sets. Implementing precise decisions will allow for effective consumption management, resulting in an ideal setting for boosting production and improving quality. Consequently, this will promote sustainable agricultural practices.

The utilization of agriculture Big Data frameworks holds the capacity to fundamentally transform the attitude of policymakers towards decision-making in the agriculture sector. The AgroBahrain framework has several practical implications, such as improving data collection and analysis, enhancing policy formation, the implementation of precision agriculture and resource optimization, establishing early warning systems and risk management strategies, and facilitating market intelligence and supply chain management. The AgroBahrain framework serves as a comprehensive manual for policymakers, providing them with the necessary information to make well-informed decisions, tackle significant obstacles, and encourage the use of sustainable agricultural methods.

Ultimately, this framework aims to enhance food security and promote environmental sustainability. The study outlines a strategy for policymakers to establish an agriculture Big Data platform in Bahrain. It focuses on preparing the primary stakeholders of the agriculture community, identifying the necessary data needs, and exploring the many possible technologies for implementation.

Conversely, the AgroBahrain framework will impact individual stakeholders by enhancing their ICT proficiency and increasing their engagement with digital technologies. Furthermore, it improves information communication among various parties involved and it will provide farmers and farm owners with more effective techniques and instruments to carry out their everyday agricultural activities.

The GO FAIR project states that data must possess the qualities of findability, accessibility, interoperability, and reusability. Obtaining and utilizing agriculture data in Bahrain is a formidable task due to its scarcity and lack of accessibility. The smooth integration of this data across many platforms is nearly unattainable. The developed framework offers a partially automated system that manages various data sources (data agnosticism) and serves as a data repository for all stakeholders with varying levels of access privileges. The platform is designed at a national level to facilitate the cooperation of all participants involved in agriculture, whether they are providers or beneficiaries of the data. The platform facilitates the integration of diverse data sources and technologies, enabling farm owners to seamlessly include various interfaces and access and evaluate the data through a unified interface. Furthermore, it will enhance agricultural business intelligence by constructing robust business environments. Conversely, the framework is specifically designed to facilitate the accessibility of reusable data for the

purposes of research, analytics, visualizations, and catering to individual requirements.

The analysis uncovered the advantages, disadvantages, potential benefits, and risks associated with the implementation and utilization of technology in the agriculture industry. Furthermore, it sheds light on related policies and regulations, especially the usage of agriculture technologies and related data policies.

The study unveiled a noticeable disparity in the extent of ICT proficiency and implementation among farmers in Bahrain, highlighting a substantial obstacle within the agricultural industry. The research aimed to fill the gap by providing significant insights and recommendations, acknowledging the significance of ICT in modernizing agricultural techniques and improving output. The study sought to analyze the precise constraints and difficulties encountered by Bahraini farmers when it came to adopting ICT tools and practices. Its objective was to enable focused interventions and efforts to address and overcome these barriers effectively. The study suggested various ways to improve farmers' awareness and use of ICT in Bahrain's agriculture sector. These included customized training programs, access to appropriate resources and technologies, and policy assistance. Implementing these measures would ultimately contribute to the sector's progress and long-term viability.

On the other hand, this study addressed the scarcity of agricultural scientific research in Bahrain, which is a severe limitation in the field of agriculture. It covers a vital gap in agricultural research in the country. The study's contribution became even more significant due to the lack of current publications on agriculture in Bahrain, specifically on the use of ICT in agricultural activities. It provided a unique viewpoint by focusing on the agriculture industry in Bahrain and analyzing the extent to which farmers have used ICT. This perspective is currently absent in the

available literature. By conducting empirical research and analysis, this study addressed a gap in existing research and offers valuable insights into the adoption of agricultural technology in Bahrain. The study also laid the foundation for future academic exploration and practical interventions in this sector.

The simulation model demonstrated the potential to revolutionize the agriculture sector by embracing technological progress. An advantage of implementing a progressive web app is its compatibility with any device. The App will guarantee the participation of a larger number of users, particularly stakeholders with limited proficiency in computer usage, who will be able to access it through their smartphones or tablets. In addition, the framework enables the future development of several agriculture Big Data applications, such as crop yield projections, crop diseases and pest control, supply chain management, and fertilizer and nutrient recommendations.

This study is the initial documentation on the implementation and utilization of ICT tools inside the agricultural industry in Bahrain. Furthermore, to the author's knowledge, it is the first study to integrate Google Cloud services with agricultural Big Data applications.

6.3. Recommendations

The study revealed many challenges, primarily focusing on poor ICT literacy and the incompetent agriculture data lifecycle in the agriculture sector in Bahrain. To overcome these challenges, an action plan must be identified with collaboration and support from both agriculture stakeholders and policymakers. Initially, this collaboration should raise awareness and usage towards adopting ICT in agriculture and transform the awareness into actual use. To encourage more usage, ICT services for agricultural production should be developed to be accessible to all users, considering all age groups.

On the other hand, ICT for agriculture diffusion using emergent technologies such as IoT and Big Data is a priority. The government should provide different incentives to encourage more citizens and residents, such as installing technologies at subsidized rates, supplying expert technical assistance, providing power supply at low costs, and monitoring network fluctuation.

Establishing a resilient infrastructure based on training programs and workshops to improve ICT literacy among farmers. This can be accomplished through the inclusion of educational institutions along with academic researchers, which will ensure the reformation of farmers' agriculture methods and beliefs and enhance their awareness of the importance of ICT in agriculture. The training programs and workshops will introduce different agriculture technologies and their benefits and build the farmer's ICT capacity and skills. They are starting with the simple usage of computers to the progressive adoption of precision farming technologies such as soil sensors and agricultural data collection and analysis. Furthermore, it will help farmers design and implement cost-effective agrarian technological solutions that suit Bahrain's environment in terms of the small area of agricultural land and the conditions of farmers.

On the other hand, data owned by governmental and non-governmental entities needs to be accessible through easy-to-use and well-documented web interfaces and APIs, especially open and public data. This will enable the integration of additional data sources to benefit both the platform and individual stakeholders.

Policymakers should consider the available regulations to enable flexible usage and adoption of agricultural technologies and devices. The Kingdom of Bahrain established the Personal Data Protection Law

(PDPL) in 2018; it was updated in 2022 by the Personal Data Protection Authority. The law aims to regulate and protect personal data for citizens and legal residents of Bahrain. Agriculture data operations, including extraction, processing, storage, and usage, must also be regulated to ensure data privacy, security, and reliability. Moreover, it is recommended that the simulation model be regulated and governed by the Ministry of Municipalities Affairs & Agriculture to provide formality over the implementation and increase trust among stakeholders, especially farmers and farm owners. This will encourage and promote data-sharing practices among different stakeholders and increase platform adoption and usage.

Through these efforts, the Kingdom of Bahrain will achieve agriculture sustainability, boost its agricultural productivity, minimize resource consumption, and increase its GDP.

A proposed work plan to improve ICT readiness and adoption among agriculture stakeholders necessitates full cooperation among various governmental, private, and educational institutions for its success (Figure 6.1) and will encompass the following tasks:

1. Organizing workshops by the Ministry of Education in collaboration with educational institutions to introduce modern technologies and their significance in the agricultural field.
2. Conducting a series of theoretical and practical training courses aimed at gradually advancing participants from the fundamentals of ICTs to the use and maintenance of modern technologies, such as remote sensing, soil sensors, and precision agriculture technologies Table 6.1, presents a sample proposal of training workshops.
3. Incorporating specific agricultural technology and data analysis modules into agricultural programs offered in educational

institutions.

4. Ensuring the readiness of farms in terms of electricity and communication networks by governmental authorities, as well as facilitating affordable internet services for farmers.
5. Allocating a supportive budget for the procurement of agricultural data collection devices and equipment, as well as data storage, processing, and analysis services, to enable stakeholders to easily benefit from them.
6. Implementing regulatory frameworks that oversee the sharing and utilisation of data, with the oversight of pertinent governmental bodies, in order to protect user rights and optimise advantages.

Table (6.1): sample workshops and training programs

Title	Estimated Duration	Description
Introductory workshop I	2 days (10 Hours)	Introduction to Computer Hardware, Basics of MS Windows, Online Browsing and Services.
Introductory workshop II	3 days (15 Hours)	Basics of MS Word and MS Excel
Introduction to Digital Tools in Agriculture	3 Days (15 hours)	this is a theoretical introduction to different applications on agriculture (Mobile and Web), sensors, data collection and analysis tools.
Introduction to Soil Sensors	2 Days (6 hours)	Installation, calibration, and connection
Introduction to remote sensing	2 days (10 hours)	Using Multispectral cameras, Using satellite imaging
Data Analysis	2 Days (8 hours)	previewing different data types collected, understanding, and comparing data.

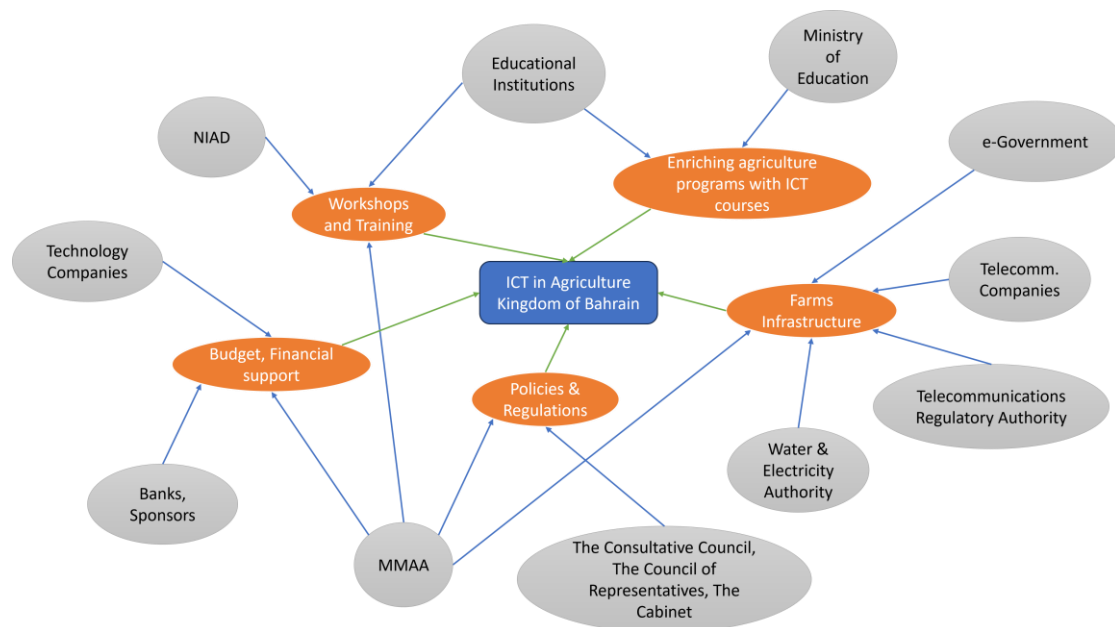


Figure (6.1): A conceptual cooperation map to improve ICT readiness and adoption.

6.4. Future Work

As stated before, the developed framework opens the door for limitless future applications; however, from the author's point of view, priority is to be given to three follow-up projects that will provide edge advantages to the study and will enhance and localize the implementation of other future applications.

Agriculture syntax in Bahrain is special, and it includes many unique Arabic keywords that require a customized dictionary to translate. Embedding these keywords in the study model will improve the platform's usability. It will upscale AI features so the current simple enquires functionality can be turned into an agriculture expert's system. This can be achieved through the collection and setting of agriculture-based Bahraini ontologies to create and embed a custom semantics dictionary into the platform.

Agriculture research and technologies extend beyond farming crops to livestock, fisheries, and agroecology. The simulation model was built to accommodate different aspects related to crop production; however, the

framework was designed to be flexible enough to include different agriculture categories in the future. This can consist of the use of wireless sensors to monitor livestock health in real time and the use of site-specific technologies for aquaculture farm management.

On the other hand, external resources are data-intensive sources that will add significant value to the collected local data. Sharing agriculture data with international platforms such as the platform for Big Data in agriculture organization (CGIAR, <https://bigdata.cgiar.org/>) will improve the simulation model's output and the forthcoming related analytical applications.

The proposed framework has the potential to advance the agriculture sector in Bahrain. Also, it can benefit other sectors through numerous applications, such as the creation of smart parks for agricultural tourism. The framework may be used as a pilot for innovation and sustainability in many sectors by utilizing ICT solutions specifically designed for the agriculture sector. Smart parks provide a chance to combine technical breakthroughs with recreational and educational activities, appealing to both visitors and enthusiasts. Therefore, the framework has a broader influence that goes beyond agriculture, promoting collaboration amongst other sectors and contributing to comprehensive socio-economic growth in Bahrain.

6.5. Limitations

Big Data is revolutionizing agriculture by providing farmers with the tools and insights needed to enhance productivity, sustainability, and profitability. However, addressing challenges related to data collection, privacy, infrastructure, and data quality is essential for realizing the full potential of Big Data in agriculture. As technology continues to advance, it is imperative that stakeholders, including governments, researchers, and

farmers, collaborate to harness the opportunities presented by Big Data while mitigating its associated challenges. The future of agriculture depends on our ability to effectively integrate and leverage Big Data to create a more resilient and sustainable food system.

The study confronts both common and local (area-specific) challenges. Regarding local obstacles, this research is the first attempt of its kind in Bahrain and was predominantly carried out during the COVID-19 epidemic and the subsequent lockdown in the country. Consequently, accessing different resources, such as farms and stakeholders, proved to be demanding, in addition to the scarcity of other resources. Conducting research on many farms was a challenging task that necessitated the assistance of third-party organizations like NIAD. Furthermore, a few of the individuals being interviewed had difficulty scheduling the day and time for the interview. Notably, one of the key stakeholders being interviewed consistently delayed the scheduling process for about three months.

Social constraints are another obstacle that needs a unique approach. Agriculture stakeholders in Bahrain, especially farmers, have their own closed community, making information collection and study participation very challenging.

Conversely, there may be limitations on financing due to the need for specialized devices and applications in the study. The equipment used for collecting experimental data was adequate. However, the availability of cloud services for simulation was restricted owing to the trial account's expiration. A significant number of Progressive Web App (PWA) features were created on a local level in order to reduce the expenses associated with cloud services, which resulted in increased time and effort. In addition, the instruments and technologies used need technical proficiency to operate well. For example, one of the devices employed

necessitated manual calibration by professionals to adapt to the specific soil conditions in Bahrain. Moreover, the utilization of particular agricultural technology, such as drones (UAVs), is limited to specific entities and necessitates several approvals to be used under the supervision of these entities in accordance with existing laws and regulations.

On the other hand, common challenges are the global limitations that affect frameworks and applications related to Big Data in agriculture, such as the ability to add new data sources dynamically. Heterogeneity in data sources (soil sensors, weather stations, legacy data, satellite remote sensing, etc.) in the matter of technology, generated data, and communication schemes creates a significant challenge for agriculture Big Data platforms to just plug and play. Meta-data management is another demanding task due to the diversity of data sources and agriculture data. Moreover, one of the main controversial issues with Big Data is the integrity of the collected data. Data lakes tend to store all kinds of data without checking, which raises reliability questions.

On the other hand, national-scale platforms tend to have API-specific challenges, such as the difference from one application to another, which tends to change over time by the application developer. Furthermore, many APIs are poorly documented and difficult to embed with external systems. Last but not least, Big Data technology is immature, and most of the available empirical studies are considered under development. Furthermore, Big Data as a technology is evolving as the research community introduces new areas of applications and improves its capabilities, features, and security domains.

References

- Adrian, A. M., Norwood, S. H., & Mask, P. L. (2005). Producers' perceptions and attitudes toward precision agriculture technologies. *Computers and Electronics in Agriculture*(48), 256–271.
- Aissi, M. E., Benjelloun, S., Loukili, Y., Lakhrissi, Y., Boushaki, A. E., Chougrad, H., & Ali, S. E. (2022). Data Lake Versus Data Warehouse Architecture: A Comparative Study. In Y. L. Saad Bennani, *WITS 2020* (pp. 201-210). Singapore: Springer.
- Ajayi, V. O. (2017). Primary sources of data and secondary sources of data. *Benue State University*, 1-6.
- Al-Ammary, J. H., & Ghanem, M. E. (2023). Information and communication technology in agriculture: awareness, readiness and adoption in the Kingdom of Bahrain. *Arab Gulf Journal of Scientific Research*.
- Alant, B. P., & Bakare, O. O. (2021). A case study of the relationship between smallholder farmers' ICT literacy levels and demographic data w.r.t. their use and adoption of ICT for weather forecasting. *Heliyon*, 7.
- Alayam. (2019, February 11). العزوف عن الاستثمار بالزراعة لانخفاض العائد... خبراء: إنتاجية القطاع الزراعي في البحرين منخفضة جداً. Retrieved from ALAYAM: <https://www.alayam.com/alayam/local/779254/News.html>
- Albilad. (2012, September 16). عمّة. "النخلة عمّتنا... والبحرين بلد المليون" Retrieved from Albilad Press: <https://www.albiladpress.com/article162591-49.html>
- albiladpress. (2020, September 21). الزراعة العضوية عاجزة عن كسب ود المزارع البحريني. Manama, Kingdom of Bahrain. Retrieved from <https://www.albiladpress.com/newspaper/4360/670320.html>
- Aldosari, F., Shunaifi, M. S., Ullah, M. A., Muddassir, M., & Noor, M. A. (2019). Farmers' perceptions regarding the use of Information and Communication Technology (ICT) in Khyber Pakhtunkhwa, Northern Pakistan. *Journal of the Saudi Society of Agricultural Sciences*, 18, 211-217.
- Ali, M., Mubeen, M., Hussain, N., Wajid, A., Farid, H. U., Awais, M., . . . Nasim, W. (2019). Role of ICT in Crop Management. In M. Hasanuzzaman, *Agronomic Crops* (pp. 637-652). Singapore: Springer.
- Alreshidi, E. (2019). Smart Sustainable Agriculture (SSA) Solution Underpinned by Internet of Things (IoT) and Artificial Intelligence (AI). *International Journal of Advanced Computer Science and Applications*, 10(5), 93-102.
- Alwatan. (2020, January 25). الزراعة انجازات متميزة في ٢٠١٩. تأهيل البحرينيين و مبادرات للنهوض بالقطاع. Retrieved February 2022, from <https://alwatannews.net/Bahrain/article/864285/%D8%A7%D9%84%D8%B2%D8%B1%D8%A7%D8%B9%D8%A9-%D8%A5%D9%86%D8%AC%D8%A7%D8%B2%D8%A7%D8%AA-%D9%85%D8%AA%D9%85%D9%8A%D8%B2%D8%A9-%D9%81%D9%8A-2019%D8%AA%D8%A3%D9%87%D9%8A%D9%84-%D8%A7%D9%84%D8%A8%D8%AD%D8%B1%D9%8A%>
- Alwatan. (2021). الوطنية لتنمية القطاع الزراعي " توقع عقد مشروع "الكشف المبكر عن سوسة" الوطنية لتنمية القطاع الزراعي " باستخدام تقنيات الفضاء النخيل الحمراء" Retrieved January 20, 2022, from <https://alwatannews.net/article/969764/Bahrain/%D8%A7%D9%84%D8%B2%D8%B1%D8%A7%D8%B9%D8%A9-%D8%A5%D9%86%D8%AC%D8%A7%D8%B2%D8%A7%D8%AA-%D9%85%D8%AA%D9%85%D9%8A%D8%B2%D8%A9-%D9%81%D9%8A-2019%D8%AA%D8%A3%D9%87%D9%8A%D9%84-%D8%A7%D9%84%D8%A8%D8%AD%D8%B1%D9%8A%>

9%88%D8%B7%D9%86%D9%8A%D8%A9-
 %D9%84%D8%AA%D9%86%D9%85%D9%8A%D8%A9-
 %D8%A7%D9%84%D9%82%D8%B7%D8%A7%D8%B9-
 %D8%A7%D9%84%D8%B2%D8%B1%D8%A7%D8%B9%D9%8A-
 %D8%AA%D9%88%D9%82%D8%B9-%D8%B9%D9%82%D8%AF-
 %D9%

- Al-Zayani, R. Y., & Karaymeh, F. (2020, February 16). *Food security in the kingdom of Bahrain: opportunities and challenges*. Retrieved from Strategy Optimization Consultancy: <https://strategyopt.com/articles/food-security-in-the-kingdom-of-bahrain-opportunities-and-challenges/>
- Amalina, F., Hashem, I. A., Azizul, Z. H., Fong, A. T., Firdaus, A., Imran, M., & Anuar, N. B. (2019). Blending Big Data Analytics: Review on Challenges and a Recent Study. *IEEE Access*, 8, 2169-3536. doi:10.1109/ACCESS.2019.2923270
- Annosi, M. C., Brunetta, F., Capo, F., & Heideveld, L. (2020). Digitalization in the agri-food industry: the relationship between technology and sustainable development. *Management Decision*, 58(8), 1737-1757.
- Aryal, J. P., Sapkota, T. B., Rahut, D. B., & Jat, M. (2020). Agricultural sustainability under emerging climatic variability: the role of climate-smart agriculture and relevant policies in India. *International Journal of Innovation and Sustainable Development*, 14(2), 219-245.
- Aubert, B. A., Schroeder, A., & Grimaudo, J. (2012). IT as enabler of sustainable farming: An empirical analysis of farmers' adoption decision of precision agriculture technology. *Decision Support Systems*, 54, 510-520.
- AWS. (n.d.). *AWS Cloud Databases*. Retrieved from Amazon Web Services: <https://aws.amazon.com/products/databases/>
- Azure. (n.d.). *Types of Databases on Azure*. Retrieved from Azure: <https://azure.microsoft.com/en-us/products/category/databases/>
- Bahrain, T. K. (2021). *Kingdom of Bahrain National Report | Achieved Progress in the Implementation of the New Urban Agenda*. Manama: Urban Agenda Platform.
- Bahrini, R., & Qaffas, A. A. (2019). Impact of Information and Communication Technology on Economic Growth: Evidence from Developing Countries. *Economies*, 7(1). doi:10.3390/economies7010021
- Bani, S. (2020). Efficient use of water for food production through sustainable crop management: Kingdom of Bahrain. *Desalination and Water Treatment*, 176, 2013-2019.
- Bansal, S. K., & Kagemann, S. (2015). Integrating Big Data: A Semantic Extract-Transform-Load Framework. *Computer*, 48(3), 42-50.
- Barh, A., & Balakrishnan, M. (2018). Smart phone applications: Role in agri-information dissemination. *Agricultural Reviews*, 82-85.
- Barnes, A., Soto, I., Eory, V., Beck, B., Balafoutis, A., Sánchez, B., . . . Gómez-Barbero, M. (2019). Exploring the adoption of precision agricultural technologies: A cross regional study of EU farmers. *Land Use Policy*, 163-174.
- BDB. (2019). *Agriculture Loans*. Retrieved February 18, 2022, from <https://www.bdb-bh.com/financial-services/agricultural-loans/>

- Bertoglio, R., Corbo, C., Renga, F. M., & Matteucci, M. (2021). The Digital Agricultural Revolution: A Bibliometric Analysis Literature Review. *IEEE Access*, 9, 134762 - 134782.
- Bhat, S. A., & Huang, N.-F. (2021). Big Data and AI Revolution in Precision Agriculture: Survey and Challenges. *IEEE Access*, 9, 110209-110222.
- Bhushan, B. (2018). Historical evolution of magnetic data storage devices and related conferences. *Microsystem Technologies*, 24, 4423–4436.
- Bifulco, I., Cirillo, S., Esposito, C., Guadagni, R., & Polese, G. (2021). An intelligent system for focused crawling from Big Data sources. *Expert Systems With Applications*, 184.
- Bilali, H. E., & Allahyari, M. S. (2018). Transition towards sustainability in agriculture and food systems: Role of information and communication technologies. *INFORMATION PROCESSING IN AGRICULTURE*, 5, 456–464.
- Bilali, H. E., Bottalico, F., Palmisano, G. O., & Capone, R. (2020). Information and Communication Technologies for Smart and Sustainable Agriculture. *30th Scientific-Experts Conference of Agriculture and Food Industry* (pp. 321–334). Sarajevo: Springer.
- Bisong, E. (2019). An Overview of Google Cloud Platform Services. In E. Bisong, *Building Machine Learning and Deep Learning Models on Google Cloud Platform* (p. 709). OTTAWA: Apress.
- Board, B. E. (2015, September 18). *ICT Fact Sheet*. Retrieved November 2021, from Economic Development Board: <https://bahrainedb.com/business-opportunities/information-communication-technology/>
- Bucci, G., Bentivoglio, D., & Finco, A. (2019). Factors affecting ICT adoption in Agriculture: a case study in Italy. *Quality - Access to Success*, 122-129.
- Ceravolo, P., Azzini, A., Angelini, M., Catarci, T., Cudré-Mauroux, P., Damiani, E., . . . Zaraket, F. (2018). Big Data Semantics. *Journal on Data Semantics*, 65–85.
- Chen, M., Mao, S., Zhang, Y., & Leung, V. C. (2014). Big Data Storage. In *Big Data* (pp. 33-49). Cham: Springer.
- Clarke, S., Wylie, G., & Zomer, H. (2013). ICT 4 the MDGs? A Perspective on ICTs' Role in Addressing Urban Poverty in the Context of the Millennium Development Goals. *Information Technologies & International Development*, 55-70.
- Clercq, M. D., Vats, A., & Biel, A. (2018). *Agriculture 4.0 The Future of Farming Technology*. Dubai: World Government Summit.
- Coble, K. H., Mishra, A. K., Ferrell, S., & Griffin, T. (2018). Big Data in Agriculture: A Challenge for the Future. *Applied Economic Perspectives and Policy*, 79–96.
- Corbellini, A., Mateos, C., Zunino, A., Godoy, D., & Schiaffino, S. (2017). Persisting big-data: The NoSQL landscape. *Information Systems*, 63, 1-23.
- Coursera. (2022, October 13). *Data Lake vs. Data Warehouse: What's the Difference?* Retrieved from Coursera : <https://www.coursera.org/articles/data-lake-vs-data-warehouse>
- Cravero, A., Pardo, S., Sepúlveda, S., & Muñoz, L. (2022). Challenges to Use Machine Learning in Agricultural Big Data: A Systematic Literature Review. *Agronomy*, 12(3).

- Davoudian, A., Chen, L., & Liu, M. (2018). A Survey on NoSQL Stores. *ACM Computing Surveys*, 51(2), 1–43.
- Delgado, J. A., Short, N. M., Roberts, D. P., & Vandenberg, B. (2019). Big Data Analysis for Sustainable Agriculture on a Geospatial Cloud Framework. *Frontiers in Sustainable Food Systems*, 3, 1-13.
- DLG. (2018, January). *Digital Agriculture - Opportunities. Risks. Acceptance*. Retrieved from Deutsche Landwirtschafts-Gesellschaft – German Agricultural Society: <https://www.dlg.org/en/agriculture/topics/a-dlg-position-paper>
- El-Khoury, G. (2016). Agriculture in Arab countries: selected indicators. *Contemporary Arab Affairs*, 9(4), 644-649.
- El-Sappagh, S. H., Hendawi, A. M., & Bastawissy, A. H. (2011). A proposed model for data warehouse ETL processes. *Journal of King Saud University - Computer and Information Sciences*, 23(2), 91-104.
- EVERT, F. K., FOUNTAS, S., JAKOVETIC, D., CRNOJEVIC, V., TRAVLOS, I., & KEMPENAAR, C. (2017). Big Data for weed control and crop protection. *Weed Research*, 57, 218-233.
- FABREGAS, R., KREMER, M., & SCHILBACH, F. (2019). Realizing the potential of digital development: The case of agricultural advice. *SCIENCE*, 366(6471).
- FAO. (2009a). *How to Feed the World in 2050*. Rome: The Food and Agriculture Organization of the United Nations.
- FAO. (2009b). *Irrigation in the Middle East region in figures*. Rome: FAO.
- FAO. (2017). *Information and Communication Technology (ICT) in Agriculture*. Rome: Food and Agriculture Organization of the United Nations.
- FAO. (2018). *Gender And Icts Mainstreaming Gender In The Use Of Information And Communication Technologies (Icts) For Agriculture And Rural Development*. Rome: The Food and Agriculture Organization of the United Nations.
- FAO, IFAD, UNICEF, WFP, & WHO. (2022). *The State of Food Security and Nutrition in the World 2022. Repurposing food and agricultural policies to make healthy diets more affordable*. Rome: FAO. doi:10.4060/cc0639en
- Farooq, M. S., Riaz, S., Abid, A., Abid, K., & Naeem, M. A. (2019). A Survey on the Role of IoT in Agriculture for the Implementation of Smart Farming. *IEEE Access*, 156237-156271.
- Fernandes, J. A., & Reite, K.-J. (2021). Conclusions and Future Vision on Big Data in Pelagic Fisheries Sustainability. In *Big Data in Bioeconomy Results from the European DataBio Project* (pp. 411-414). Cham: Springer.
- FINK, M., HÖLTL, A., & BRANDTWEINER, R. (2020). Potential Impact Of Information Systems Tackling Sdg10 On The Dimensions Of Health Care And Gender. *Sustainable Development and Planning XI* (pp. 221-234). WIT Press.
- Ge, M., Bangui, H., & Buhnova, B. (2018). Big Data for Internet of Things: A Survey. *Future Generation Computer Systems*, 87, 601-614.
- Gil, J. D., Reidsma, P., Giller, K., Todman, L., Whitmore, A., & Ittersum, M. v. (2019). Sustainable development goal 2: Improved targets and indicators for agriculture and food security. *Ambio*, 685–698.

- Gill, S. S., Chana, I., & Buyya, R. (2017). IoT Based Agriculture as a Cloud and Big Data Service: The Beginning of Digital India. *Journal of Organizational and End User Computing*, 29(4), 1-23.
- Google. (n.d.). *Cloud Bigtable*. Retrieved March 12, 2022, from Google Cloud: <https://cloud.google.com/bigtable#:~:text=Bigtable%20is%20a%20key%2Dvalue,Ads%2C%20Drive%2C%20and%20Maps>.
- Google. (n.d.). *Google Cloud databases*. Retrieved from Google Cloud: <https://cloud.google.com/products/databases>
- Gössling, S., & Hall, C. M. (2019). Sharing versus collaborative economy: how to align ICT developments and the SDGs in tourism. *Journal of Sustainable Tourism*, 74-96.
- Gulf Daily News. (2021, May 20). *New push to train farmers in Bahrain*. Retrieved from ZAWYA: <https://www.zawya.com/en/economy/new-push-to-train-farmers-in-bahrain-fysop5xf>
- Hariri, R. H., Fredericks, E. M., & Bowers, K. M. (2019). Uncertainty in big data analytics: survey, opportunities, and challenges. *Journal of Big Data*, 6. doi:10.1186/s40537-019-0206-3
- Hashem, I. A., Yaqoob, I., Anuar, N. B., Mokhtar, S., Gani, A., & Khan, S. U. (2015). The rise of “big data” on cloud computing: Review and open research issues. *Information Systems*, 47, 98 –115.
- Hox, J. J., & Boeije, H. R. (2005). Data collection, primary versus secondary. In *Encyclopedia of Social Measurement* (Vol. 1, pp. 593-599). Elsevier.
- Huang, Y., Zhong-xin, C., Tao, Y., Xiang-zhi, H., & Xing-fa, G. (2018). Agricultural remote sensing big data: Management and applications. *Journal of Integrative Agriculture*, 17(9), 1915–1931.
- ICARDA. (2007). *ICARDA in the Arabian Peninsula Ties that bind*. Aleppo, Syrian Arab Republic: International Center for Agricultural Research in the Dry Areas (ICARDA).
- IGA. (2016). *Bahrain Open Data Portal*. Retrieved November 20, 2021, from <https://www.data.gov.bh/en/ResourceCenter>
- IGA. (2018). *Bahrain. eGovernment*. Retrieved from eGovernment: <https://www.bahrain.bh>
- IGA. (2023, June 14). *iGA CE Launches Beta Version of Revamped Bahrain Open Data Portal*. Retrieved from Information & eGovernment Authority’s Corporate Website: <https://www.iga.gov.bh/article/iga-ce-launches-beta-version-of-revamped-bahrain-open-data-portal>
- Inaba, T., & Squicciarini, M. (2017). ICT: A new taxonomy based on the international patent classification. *OECD Science, Technology and Industry Working Papers*.
- Independent Evaluation Group. (2013). *The World Bank Group and the Global Food Crisis : An Evaluation of the World Bank Group Response*. Washington, DC: World Bank Group. Retrieved from <https://openknowledge.worldbank.org/handle/10986/16505>
- ITU. (2003, December 12). *World Summit on the Information Society*. Retrieved from International Telecommunication Union: <https://www.itu.int/net/wsis/docs/geneva/official/poa.html>
- ITU. (2022). *ITU DataHub*. Retrieved from International Telecommunication Union: <https://datahub.itu.int/data/>
- ITU. (2022). *WSIS Stocktaking 2022 Global Report*. Geneva: International Telecommunication Union.

- Januzaj, Y., Ajdari, J., & Selimi, B. (2015). DBMS as a Cloud service: Advantages and Disadvantages. *Procedia - Social and Behavioral Sciences*, 1851-1859.
- Keswani, B., Mohapatra, A. G., Keswani, P., Khanna, A., Gupta, D., & Rodrigues, J. (2020). Improving weather dependent zone specific irrigation control scheme in IoT and big data enabled self driven precision agriculture mechanism. *Enterprise Information Systems*, 14(9-10), 1494-1515.
- Khder, M. A. (2021). Web Scraping or Web Crawling: State of Art, Techniques, Approaches and Application. *International Journal of Advances in Soft Computing and its Applications*, 13(3), 145-168.
- Khine, P. P., & Wang, Z. S. (2018). Data lake: a new ideology in big data era. *4th Annual International Conference on Wireless Communication and Sensor Network (WCSN 2017)* (pp. 395-405). Wuhan: EDP Sciences.
- Kirsch, G., & Sullivan, P. A. (1992). Methods and Methodology in Composition Research. *Rhetoric Society Quarterly*, 22(3), 70-73.
- Klerkx, L., Jakku, E., & Labarthe, P. (2019). A review of social science on digital agriculture, smart farming and agriculture 4.0: New contributions and a future research agenda. *NJAS - Wageningen Journal of Life Sciences*, 90-91.
- Knierim, A., Kernecker, M., Erdle, K., Kraus, T., Borges, F., & Wurbs, A. (2019). Smart farming technology innovations – Insights and reflections from the German Smart-AKIS hub. *NJAS - Wageningen Journal of Life Sciences*, 90-91.
- Kostoska, O., & Kocarev, L. (2019). A Novel ICT Framework for Sustainable Development Goals. *Sustainability*, 11(7).
- Lange, A. F., & Peake, J. (2020). Precision Agriculture. In Y. J. Morton, F. v. Diggelen, J. J. Jr., B. W. Parkinson, S. Lo, G. Gao, & F. v. Y. T. Jade Morton (Ed.), *Position, Navigation, and Timing Technologies in the 21st Century: Integrated Satellite Navigation, Sensor Systems, and Civil Applications* (pp. 1735-1747). Wiley.
- Li, C., & Niu, B. (2020). Design of smart agriculture based on big data and Internet of things. *International Journal of Distributed Sensor Networks*, 16(5).
- Li, C., & Niu, B. (2020). Design of smart agriculture based on big data and Internet of things. *International Journal of Distributed Sensor Networks*, 16(5).
- Lowenberg-DeBoer, J., & Erickson, B. (2019). Setting the Record Straight on Precision Agriculture Adoption. *Agronomy Journal*, 111(4), 1552-1569.
- M.H.Padgavankar, & S.R.Gupta. (2014). Big Data Storage and Challenges. *International Journal of Computer Science and Information Technologies*, 5(2), 2218-2223.
- Magnin, C. (2016, August 19). *How big data will revolutionize the global food chain*. Retrieved from McKinsey Digital: <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/how-big-data-will-revolutionize-the-global-food-chain>
- Majumdar, J., Naraseeyappa, S., & Ankalaki, S. (2017). Analysis of agriculture data using data mining techniques: application of big data. *Journal of Big Data*, 4.
- Mali, N., & Bojewar, S. (2015). A Survey of ETL Tools. *International Journal of Computer Techniques*, 20-26.

- Mallek, H., Ghozzi, F., & Gargouri, F. (2020). Towards Extract-Transform-Load Operations in a Big Data context. *International Journal of Sociotechnology and Knowledge Development*, 12(2), 77-95.
- Marín-Ortegaa, P. M., Dmitriyevb, V., Abilovb, M., & Gómezb, J. M. (2014). ELTA: New Approach in Designing Business Intelligence Solutions in Era of Big Data. *Procedia Technology*, 16, 667-674.
- MarketsandMarkets. (2023, August). *Precision Farming Market*. Retrieved from MarketsandMarkets: <https://www.marketsandmarkets.com/Market-Reports/precision-farming-market-1243.html>
- Mauro, A. D., Greco, M., & Grimaldi, M. (2016). A formal definition of Big Data based on its essential features. *Library Review*, 65(3), 122-135.
- Mazzetto, F., Gallo, R., & Sacco, P. (2020). Reflections and Methodological Proposals to Treat the Concept of “Information Precision” in Smart Agriculture Practices. *Sensors*, 20, 317–322.
- Mehrad, A., & Zangeneh, M. H. (2019). Comparison between Qualitative and Quantitative Research Approaches: Social Sciences. *International Journal For Research In Educational Studies*, 5(7).
- Mekala, M. S., & Viswanathan, P. (2017). A novel technology for smart agriculture based on IoT with cloud computing. *International conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)* (pp. 75-82). Palladam, India: IEEE.
- Mekonnen, Y., Namuduri, S., Burton, L., Sarwat, A., & Bhansali, S. (2020). Review—Machine Learning Techniques in Wireless Sensor Network Based Precision Agriculture. *Journal of The Electrochemical Society*, 167(3).
- Milovanović, S. (2014). The Role And Potential Of Information Technology In Agricultural Improvement. *Economics of Agriculture*, 2, 471- 485.
- Minet, J., Curnel, Y., Gobin, A., Goffart, J.-P., Mélard, F., Tychon, B., . . . Defourny, P. (2017). Crowdsourcing for agricultural applications: A review of uses and opportunities for a farmsourcing approach. *Computers and Electronics in Agriculture*, 126-138.
- Mohieldin, M. (2018). *Leveraging technology to achieve the Sustainable Development Goals*. Retrieved from The World Bank: <http://blogs.worldbank.org/voices/leveraging-technology-achieve-sustainable-development-goals>
- Moreddu, C. (2016). Public-Private Partnerships for Agricultural Innovation: Lessons From Recent Experiences. *OECD Food, Agriculture and Fisheries Papers*.
- Morris, R. J., & Truskowski, B. J. (2003). The evolution of storage systems. *IBM Systems Journal*, 42(2), 205–217.
- Municipalities. (2019, September 17). *الزراعة والفاو توقعان ٤ اتفاقات شراكة في مجال الأمن الغذائي*. Retrieved December 15, 2021, from <https://www.mun.gov.bh/agri/NewsDetails.action?param=1042&lang=en>
- Murray W. Scown, M. V. (2020). Billions in Misspent EU Agricultural Subsidies Could Support the Sustainable Development Goals. *One Earth*, 3(2), 237-250.
- Nallusamy, A., Balasubramaniam, S., & Chellappan, S. K. (2015). Use Of Information And Communication Technology (ICT) To Achieve

- Information Literacy In Agriculture. *International Journal of Agricultural Extension*, 111-122.
- Naser, H. A. (2022). Technology for environmental management in Small Island Developing States: the case of Bahrain. *Current Opinion in Environmental Sustainability*.
- Neves, R. A., & Cruvinel, P. E. (2020). Model for Semantic Base Structuring of Digital Data to Support Agricultural Management. *14th International Conference on Semantic Computing* (pp. 337-340). California: IEEE.
- Newman, I., & Ridenour, C. (1998). *Qualitative-Quantitative Research Methodology: Exploring the Interactive Continuum*. Educational Leadership Faculty Publications.
- Ngo, V. M., Le-Khac, N.-A., & Kechadi, M.-T. (2019). Designing and Implementing Data Warehouse for Agricultural Big Data. In S. S.-J. Keke Chen, *Big Data – BigData 2019* (pp. 1-17). California: Springer, Cham.
- Nguyen, V.-Q., Ngoc, N., & Kim, K. (2017). Design of a Platform for Collecting and Analyzing Agricultural Big Data. *Journal of Digital Contents Society*, 18(1), 149-158.
- Nhemachena, C., Matchaya, G., Nhemachena, C. R., Karuaihe, S., Muchara, B., & Nhlengethwa, S. (2018). Measuring Baseline Agriculture-Related Sustainable Development Goals Index for Southern Africa. *Sustainability*, 10(3).
- NIAD. (2019). *About Niad*. Retrieved October 15, 2021, from <http://bigdata.com.bh/about-niad/>
- Niculescu, V. (2020). On the Impact of High Performance Computing in Big Data Analytics for Medicine. *Applied Medical Informatics*, 9-18.
- Niu, J. (2016). Organisation and description of datasets. *Archives and Manuscripts*.
- Ofoeda, J., Boateng, R., & Effah, J. (2019). Application Programming Interface (API) Research: A Review of the Past to Inform the Future. *International Journal of Enterprise Information Systems (IJEIS)*, 15(3), 76-95.
- Osinga, S. A., Paudel, D., Mouzakitis, S. A., & Athanasiadis, I. N. (2022). Big data in agriculture: Between opportunity and solution. *Agricultural Systems*, 195.
- Oussous, A., Benjelloun, F.-Z., Lahcen, A. A., & Belfkih, S. (2018). Big Data technologies: A survey. *Journal of King Saud University - Computer and Information Sciences*, 30(4), 431-448.
- P.S., Gopal, M., Chintala, & Renta, B. (2020). Big Data Challenges and Opportunities in Agriculture. *International Journal of Agricultural and Environmental Information Systems*, 11(1), 48-66.
- Pallavi, S., Mallapur, J. D., & Bendigeri, K. Y. (2017). Remote sensing and controlling of greenhouse agriculture parameters based on IoT. *2017 International Conference on Big Data, IoT and Data Science (BIG)* (pp. 44-48). Pune: IEEE.
- Panwar, A. (2020). Data Lake Architecture: A New Repository for Data Engineer. *International Journal of Organizational and Collective Intelligence*, 10(1).
- Parke, C. (2015, December). Impact of Technology on Agriculture and Food Production. Research Gate. doi:10.13140/RG.2.1.1608.8400

- Parra-López, C., Reina-Usuga, L., Carmona-Torres, C., Sayadi, S., & Klerkx, L. (2021). Digital transformation of the agrifood system: Quantifying the conditioning factors to inform policy planning in the olive sector. *Land Use Policy*.
- Patel, S., & Iliyas, S. (2014). Impact Of Information Technology In Agriculture Sector. *International Journal of Food, Agriculture and Veterinary Sciences*, 4(2), 17-22.
- Planning, D. o. (2009). *State of the Environment in the Kingdom of Bahrain*. Manama: Public Commission for the Protection of Marine Resources, Environment and Wildlife.
- Priyadarshini, P., & Abhilash, P. C. (2020). Policy recommendations for enabling transition towards sustainable agriculture in India. *Land Use Policy*, 96.
- Purnomo, S. H., & Lee, Y.-H. (2010). An Assessment of Readiness and Barriers towards ICT Programme Implementation: Perceptions of Agricultural Extension Officers in Indonesia. *International Journal of Education and Development using Information and Communication Technology*, 19-36.
- Raksha, Rao, I. S., Meera, S. N., & Kumar, V. (2015). e-Readiness in Agricultural Extension System. *Indian Research Journal Of Education Extention*, 77-81.
- Ravindra, H., Sehgal, M., & Narasimhamurthy, H. B. (2020). Use of Information and Communications Technology (ICT) Tools in Increasing Income of Tribal Farmers of Karnataka. *Asian Journal of Agricultural Extension, Economics & Sociology*, 138(9), 168-178.
- Ray, P. P. (2017). Internet of things for smart agriculture: Technologies, practices and future direction. *Journal of Ambient Intelligence and Smart Environments*, 395–420.
- Ruder, S.-L. (2019). A Digital Agricultural Revolution: Ontario Grain Farmer Perceptions of Digital Farming and Big Data. *Thesis in Master of Environmental Studies*. Waterloo, Ontario, Canada: University of Waterloo.
- Saiz-Rubio, V., & Rovira-Más, F. (2020). From Smart Farming towards Agriculture 5.0: A Review on Crop Data Management. *Agronomy*, 20(10).
- Salloum, S. A., Khan, R., & Shaalan, K. (2020). A Survey of Semantic Analysis Approaches. *The International Conference on Artificial Intelligence and Computer Vision* (pp. 61-70). Cairo: Springer.
- Sambrekar, K., Rajpurohit, V. S., & Joshi, J. (2018). A Proposed Technique for Conversion of Unstructured Agro-Data to Semi-Structured or Structured Data. *Fourth International Conference on Computing Communication Control and Automation* (pp. 1-5). Pune: IEEE.
- Saranya, P., & Asha, P. (2019). Survey on Big Data Analytics in Health Care. *2019 International Conference on Smart Systems and Inventive Technology (ICSSIT)* (pp. 46-51). Tirunelveli, India: IEEE.
- Sarker, M. N., Islam, M. S., Murmu, H., & Rozario, E. (2020). Role of Big Data On Digital Farming. *International Journal of Scientific & Technology Research*, 1-11.
- Sayruamyat, S., & Nadee, W. (2020). Acceptance and readiness of Thai farmers toward digital technology. In Y.-D. Zhang, J. K. Mandal, C. So-In, & N.

- V. Thakur, *Smart Trends in Computing and Communications* (pp. 75–82). Springer Nature Singapore.
- Scown, M. W., Brady, M. V., & Nicholas, K. A. (2020). Billions in Misspent EU Agricultural Subsidies Could Support the Sustainable Development Goals. *One Earth*, 3(2), 237-250.
- Shebeeb, B. (2006). Productivity growth: Bahrain agriculture and fisheries sector. *Emirates Journal of Food and Agriculture*, 18(2), 29-46.
- SIDDIQA, A., KARIM, A., & GANI, A. (2017). Big data storage technologies: a survey. *Frontiers of Information Technology & Electronic Engineering*, 18(8), 1040-1070.
- Singh, A. S., & Masuku, M. (2014). Sampling Techniques & Determination Of Sample Size In Applied Statistics Research: An Overview. *International Journal of Economics, Commerce and Management*, 2(11), 1-22.
- Singh, R., & Singh, G. S. (2017). Traditional agriculture: a climate-smart approach for sustainable food production. *Energy, Ecology and Environment*, 2, 296-316.
- Södergård, C., Mildorf, T., Berre, A. J., Tsalgatidou, A., & Charvát, K. (2021). Big Data Technologies in DataBio. In C. Södergård, T. Mildorf, E. Habyarimana, A. J. Berre, J. A. Fernandes, & C. Zinke-Wehlmann, *Big Data in Bioeconomy Results from the European DataBio Project* (pp. 3-16). Cham: Springer.
- Sorri, K., Mustafee, N., & Seppänen, M. (2022). Revisiting IoT definitions: A framework towards comprehensive use. *Technological Forecasting and Social Change*, 179. doi:10.1016/j.techfore.2022.121623
- Spanaki, K., Karafili, E., & Despoudi, S. (2021). AI applications of data sharing in agriculture 4.0: A framework for role-based data access control. *International Journal of Information Management*, 59.
- Spielman, D., Lecoutere, E., Makhija, S., & Campenhout, B. V. (2021). Information and Communications Technology (ICT) and Agricultural Extension in Developing Countries. *Annual Review of Resource Economics*, 13(1), 177-201.
- Statistics, G. D. (2015). *Millennium Development Goals of The Kingdom of Bahrain 2015*. Manama: Central Informatics Organization.
- Steiner, T. (2018). What is in a Web View: An Analysis of Progressive Web App Features When the Means of Web Access is not a Web Browser. *Companion Proceedings of the The Web Conference 2018* (pp. 789–796). Leon: International World Wide Web Conferences Steering Committee.
- Strohbach, M., Daubert, J., Ravkin, H., & Lischka, M. (2016). Big Data Storage. In J. M. Cavanillas, E. Curry, & W. Wahlster, *New Horizons for a Data-Driven Economy* (pp. 119-141). Cham: Springer.
- Subashini, P., & Fernando, S. (2017). Empowerment of Farmers Through ICT Literacy. *National Information Technology Conference (NITC)* (pp. 119-124). Colombo: IEEE.
- Tandel, S. S., & Jamadar, A. (2018). Impact of Progressive Web Apps on Web App Development. *International Journal of Innovative Research in Science, Engineering and Technology*, 9439-9444.
- Tariq, N., Asim, M., Al-Obeidat, F., Farooqi, M. Z., Baker, T., Hammoudeh, M., & Ghafir, I. (2019). The Security of Big Data in Fog-Enabled IoT Applications Including Blockchain: A Survey. *Sensors*, 19(8).

- Thompson, F. M. (1968). The Second Agricultural Revolution, 1815-1880. *The Economic History Review*, 21(1), 62-77. doi:10.2307/2592204
- Trivelli, L., Apicella, A., Chiarello, F., Rana, R., Fantoni, G., & Tarabella, A. (2019). From precision agriculture to Industry 4.0 Unveiling technological connections in the agrifood sector. *British Food Journal*, 121(8), 1730-1743. doi:https://doi.org/10.1108/BFJ-11-2018-0747
- UN. (2015, September 25). *UN Technology Facilitation Mechanism (TFM)*. Retrieved from United Nations: <https://sdgs.un.org/tfm>
- UN. (2022). *Global indicator framework for the Sustainable Development Goals and targets of the 2030 Agenda for Sustainable Development*. The United Nations. Retrieved from https://unstats.un.org/sdgs/indicators/Global%20Indicator%20Framework%20after%202022%20refinement_Eng.pdf
- Unwin, T. (2019). The Future Use of Technology in Education and Learning in the Commonwealth. *The Round Table*, 447-458.
- UPDA. (2022, March). *Land Use Statistics*. Retrieved July 14, 2022, from <https://www.upda.gov.bh/en/category/land-use-statistics>
- Valduriez, P. (2009). Parallel Database Management. In *Encyclopedia of Database Systems* (pp. 2026–2029). Springer.
- Ven, A. H., & Poole, M. S. (2017). Field Research Methods. In J. A. Baum, *The Blackwell Companion to Organizations* (pp. 867-888). Blackwell Publishers Ltd.
- Waples, R. S., & Gaggiotti, O. (2006). What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular Ecology*, 15(6), 1419-1439.
- Wong, P. P. (2011). Small island developing states. *WIREs*, 1-6.
- WorldBank. (2021a, March 16). *A Roadmap for Building the Digital Future of Food and Agriculture*. Retrieved from The World Bank: <https://www.worldbank.org/en/news/feature/2021/03/16/a-roadmap-for-building-the-digital-future-of-food-and-agriculture>
- WorldBank. (2021b, April 5). *CLIMATE-SMART AGRICULTURE*. Retrieved from The World Bank: <https://www.worldbank.org/en/topic/climate-smart-agriculture>
- Wu, J., Guo, S., Huang, H., Liu, W., & Xiang, Y. (2018). Information and Communications Technologies for Sustainable Development Goals: State-of-the-Art, Needs and Perspectives. *IEEE Communications Surveys & Tutorials*, 20(3), 2389 - 2406.
- Yang, X., Shu, L., Chen, J., Ferrag, M. A., Wu, J., Nurellari, E., & Huang, K. (2021). A Survey on Smart Agriculture: Development Modes, Technologies, and Security and Privacy Challenges. *IEEE/CAA Journal of Automatica Sinica*, 8(2), 273-302. doi:10.1109/JAS.2020.1003536
- Zhai, Z., Martínez, J. F., Beltran, V., & Martínez, N. L. (2020). Decision support systems for agriculture 4.0: Survey and challenges. *Computers and Electronics in Agriculture*, 170. doi:https://doi.org/10.1016/j.compag.2020.105256
- Zhang, A., Jakku, E., Llewellyn, R. S., & Baker, E. (2018). Surveying the Needs and Drivers for Digital Agriculture in Australia. *Farm Policy Journal*, 15(1).
- Zouari, F., Kabachi, N., Boukadi, K., & Ghedira-Guegan, C. (2021). Data Management in the Data Lake: A Systematic Mapping. 25th

- International Database Engineering & Applications Symposium* (pp. 280-284). Montreal: ACM.
- Zuppo, C. M. (2012). Defining ICT in a Boundaryless World : The Development of a Working Hierarchy. *International Journal of Managing Information Technology*, 4(3), 13-22.

Appendices

Appendix A

Online Questionnaire

استخدام تكنولوجيا المعلومات والاتصالات في المجال
الزراعي بمملكة البحرين

*مطلوب

*العمر

☐ أقل من 25 عام

☐ 25 - 35 عام

☐ 36 - 50 عام

☐ أكثر من 50 عام

*المستوى التعليمي

☐ غير متعلم

☐ المرحلة الابتدائية

☐ تعليم متوسط

☐ تعليم جامعي أو أعلى

*الدخل الشهري

☐ أقل من 250 دينار بحريني

☐ 250 - 500 دينار بحريني

☐ 501 - 1000 دينار بحريني

☐ 1001 - 2000 دينار بحريني

☐ أكثر من 2000 دينار بحريني



الجنس *

ذكر ☐

أنثى ☐

مساحة الأرض *

أقل من نصف هكتار ☐

نصف - 1 هكتار ☐

1 - 3 هكتار ☐

أكثر من 3 هكتار ☐

المساحة المزروعة *

أقل من نصف المزرعة ☐

50% - 75% من مساحة المزرعة ☐

المزرعة كاملة ☐

طريقة الزراعة المستخدمة (يمكنك اختيار أكثر من طريقة) *

صوب زجاجية ☐

زراعة بدون تربة ☐

زراعة تقليدية ☐

أخرى ☐



انواع المحاصيل التي تتم زراعتها (يمكنك اختيار أكثر من محصول) *

☐ تمور

☐ محاصيل عطرية

☐ محاصيل ورقية (المطبخ)

☐ اعشاب

☐ حمضيات

☐ فواكه

☐ خضروات

☐ اخرى

كمية الأسمدة المستخدمة سنويا (بالكيلو جرام) *

إجابتك

كمية المبيدات المستخدمة سنويا (بالكيلو جرام) *

إجابتك



كيف تقيم مستواك بالنسبة لاستخدام الادوات الخاصة بتقنية المعلومات؟ *

لا اوافق بشدة	لا اوافق	لا اعلم	اوافق	اوافق بشدة	
☐	☐	☐	☐	☐	اعرف كيف استخدم الكمبيوتر جيداً
☐	☐	☐	☐	☐	استطيع تصفح الانترنت بسهولة و بدون مساعدة
☐	☐	☐	☐	☐	انا على علم بالفرق بين الأجهزة و البرامج و التطبيقات
☐	☐	☐	☐	☐	استخدم خدمات عبر الانترنت (مثال: التسوق و الدفع الإلكتروني)

استخدم الأجهزة الذكية التالية (يمكنك اختيار أكثر من جهاز) *

☐ الهواتف الذكية (مثال: iPhone, Samsung, HTC)

☐ الألواح الذكية (مثال: iPad, Tablets)

☐ أجهزة الحاسب المحمولة (Laptop)

☐ أجهزة الحاسب المكتبية (Desktop)

☐ لا استخدم أجهزة ذكية



استخدم وسائل التواصل الاجتماعي الإلكترونية التالية (يمكنك اختيار أكثر من وسيلة) *

☐ فيسبوك (Facebook)

☐ انستغرام (Instagram)

☐ سناب شات (SnapChat)

☐ تويتر (Twitter)

☐ واتس اب (WhatsApp)

☐ أخرى

☐ لا استخدم وسائل التواصل الاجتماعي الإلكترونية

قيم معلوماتك الخاصة بتقنية المعلومات في المجال الزراعي *

لا اوافق بشدة	لا اوافق	لا اعلم	اوافق	اوافق بشدة	
☐	☐	☐	☐	☐	اعلم جيدا ما هي قواعد البيانات
☐	☐	☐	☐	☐	أنا على دراية جيدة بأجهزة استشعار التربة (Soil Sensors)
☐	☐	☐	☐	☐	أنا على دراية كاملة بالاستشعار عن بعد الزراعي (Agriculture Remote Sensing)
☐	☐	☐	☐	☐	أنا أفهم مصطلحات تقنية مثل إنترنت الأشياء (IoT) والبيانات الضخمة (Big Data)
☐	☐	☐	☐	☐	أنا على دراية جيدة بمصطلح تقنيات الزراعة الدقيقة (Precision Agriculture)



إلى أي مدى تستخدم تقنية المعلومات في إدارة أعمالك الزراعية؟ *

لا اوافق بشدة	لا اوافق	لا اعلم	اوافق	اوافق بشدة	
☐	☐	☐	☐	☐	استخدم اجهزة الحاسب الالى لادارة عمل المزرعة
☐	☐	☐	☐	☐	يوجد اشتراك خاص بالمزرعة في خدمات الانترنت
☐	☐	☐	☐	☐	احتاج الانترنت لاتمام الاعمال الخاصة بالمزرعة
☐	☐	☐	☐	☐	استخدم برامج و تطبيقات خاصة لإدارة المزرعة
☐	☐	☐	☐	☐	استخدم مجسات التربة (Soil Sensors)
☐	☐	☐	☐	☐	استخدم تكنولوجيا الاستشعار عن بعد (Remote Sensing)



الى اي مدى تتقبل استخدام الأدوات و التطبيقات الخاصة بتقنية المعلومات في مزرعتك؟ *

لا اوافق بشدة	لا اوافق	لا اعلم	اوافق	اوافق بشدة	
☐	☐	☐	☐	☐	لدي فضول شديد حول كيفية عمل التقنيات الحديثة الخاصة بالزراعة
☐	☐	☐	☐	☐	احب تجربة طرق جديدة لتطوير مزرعتي
☐	☐	☐	☐	☐	غالبًا ما أبحث عن معلومات حول المنتجات و التقنيات الجديدة في المجال الزراعي

ما مدى استعدادك لتطبيق استخدام تقنية المعلومات في مزرعتك؟ *

لا اوافق بشدة	لا اوافق	لا اعلم	اوافق	اوافق بشدة	
☐	☐	☐	☐	☐	لدي الموارد والفرص والمعرفة لاستخدام التقنيات الخاصة بالزراعة الدقيقة (Precision Agriculture)
☐	☐	☐	☐	☐	لدي إمكانية الوصول إلى الموارد التي سأحتاجها لاستخدام تقنيات الزراعة الدقيقة.



ما العوائق التي واجهتها أو يمكن مواجهتها إن اردت تطبيق تقنية المعلومات في المجال الزراعي؟ (يمكنك اختيار أكثر من إجابة) *

- ☐ نقص المعرفة بالتقنيات المستخدمة
- ☐ تكلفة تجهيز البنية التحتية
- ☐ ضعف امدادات الطاقة الكهربائية و شبكات المعلومات
- ☐ لا توجد خبرات تقنية كافية
- ☐ ارتفاع تكلفة الأجهزة و البرامج
- ☐ الافتقار أو صعوبة الوصول الى الدعم المادي
- ☐ أخرى

في اعتقادك هل استخدام تقنية المعلومات في المجال الزراعي سوف يؤدي الى: *

لا اوافق بشدة	لا اوافق	لا اعلم	اوافق	اوافق بشدة
☐	☐	☐	☐	☐
استخدام تقنيات الزراعة الدقيقة سوف يزيد من انتاجية المحصول				
☐	☐	☐	☐	☐
تقنيات الزراعة الدقيقة بإمكانها توفير المعلومات الضرورية لإتخاذ قرارات أفضل تسهم في تقليل استهلاك الموارد مثل (الماء، المبيدات، و الاسمدة)				
☐	☐	☐	☐	☐
البيانات التي توفرها تقنيات الزراعة الدقيقة ستقلل من تأثير انشطتي الزراعية على البيئة				

إرسال



Appendix B

Agriculture Data Lifecycle in Bahrain

The following interview questions will identify different Information and Communication Technologies (ICTs) used to enable communication between agriculture stakeholders (farmers, business owners, researchers, etc..). Questions will include but not limited to; 1- collection, storage, and delivery of agriculture data – 2- preparation, storage, and delivery of extension services – 3- Different ICTs platforms used to allow effective interaction between agriculture stakeholders (forums, expert systems, crop-sharing systems, etc..).

Table (B.1): Interview Card (questions and interview meta data)

Interview Card # Agriculture/Data Expert		Date/Time	Interview Location
Interviewee		Organization	
Position		Contact	
English Questions		Questions in Arabic	
1	How many farms in Bahrain? By individuals? By corporates? Based on Alayam newspaper in 2019 we have around 650 farm owners?	كم عدد المزارع في البحرين؟ مملوكة للأفراد؟ مملوكة لشركات؟ استنادًا إلى جريدة الأيام لعام ٢٠١٩، لدينا حوالي ٦٥٠ مالك مزرعة، هل هذا دقيق؟ وما هي الإحصائيات الصحيحة؟	
2	Who are the primary stakeholders of the agriculture sector in Bahrain?	من هم أصحاب المصلحة الرئيسيين في قطاع الزراعة في البحرين؟	
3	What are the major crops cultivated by season? What is the percentage of imported crops yearly?	ما هي المحاصيل الرئيسية المزروعة حسب الموسم؟ ما هي نسبة المحاصيل المستوردة سنوياً؟	
4	What is the average farm size in Bahrain? Minimum and maximum?	ما هو متوسط مساحة المزارع في البحرين؟ الحد الأدنى والحد الأقصى؟	
5	How many open farms and	كم عدد المزارع المفتوحة والمزارع المحمية في	

	protected farms in Bahrain?	البحرين؟
6	What type of agriculture extension services provided by the authority to the Bahraini farmers?	ما نوع خدمات الإرشاد الزراعي التي تقدمها الحكومة للمزارعين البحرينيين؟
7	How are the information of the extension services prepared and delivered?	كيف يتم إعداد وتقديم المعلومات الخاصة بالخدمات الإرشادية؟ كيف يتم التواصل مع المزارعين وأصحاب المصلحة الآخرين؟
8	Are there any agriculture knowledge or expert networks in Bahrain?	هل توجد أي شبكات معلوماتية زراعية أو خبراء في البحرين؟
9	If yes, what are they? Which platforms are used for communication? Who manages those networks? What information do they provide?	إذا نعم، ما هم؟ ما المنصات المستخدمة للتواصل؟ من يدير تلك الشبكات؟ ما هي المعلومات التي يقدمونها؟
10	What external communities do agriculture stakeholders in Bahrain participate in? If none, why? If yes, please list them, and how they impact the agriculture sector in Bahrain? Such as (CIGAR, FAO e-agriculture)	ما هي المجتمعات الخارجية التي يشارك فيها أصحاب المصلحة الزراعيون في البحرين؟ إذا لم يكن كذلك، فلماذا؟ إذا كانت الإجابة بنعم، يرجى ذكرهم، وكيف يؤثر ذلك على قطاع الزراعة في البحرين؟
11	Are there any projects (internal/external) that aim to collect agriculture data? If yes, what are they? What is the framework they use?	هل هناك أي مشاريع (داخلية / خارجية) تهدف إلى جمع البيانات الزراعية؟ إذا نعم، ما هم؟ ما هي التقنيات والأساليب المستخدمة؟
12	What type of agriculture data is collected in Bahrain?	ما نوع البيانات الزراعية التي يتم جمعها في البحرين؟
13	How is the data collected? Tools, methods, who (personnel, department)?	كيف يتم جمع البيانات؟ أدوات، طرق، من (الأفراد، القسم)؟

14	How often is the data collected?	كم مرة يتم جمع البيانات؟ شهريا – سنويا هل هناك فترة محددة؟
15	What is the storage mechanism for the collected data?	ما هي آلية تخزين البيانات المجمعة؟
16	Who benefits from the collected data? How do they access it? What are the access privileges?	من يستفيد من البيانات المجمعة؟ كيف يصلون إليها؟ ما هي شروط استخدام البيانات؟
17	What agriculture data do you need? What are the difficulties do you face to collect these data? What can be done to overcome those difficulties?	ما البيانات الزراعية التي تحتاجها؟ ما هي الصعوبات التي تواجهها في جمع هذه البيانات؟ ما الذي يمكن عمله للتغلب على تلك الصعوبات؟
18	Are you familiar with the expression precision agriculture and related technologies such as (soil sensors, remote sensing, and Big Data)?	هل انت على دراية بمبدأ الزراعة الدقيقة و التقنيات المحيطة بها مثل مجسات التربة، الاستشعار عن بعد، و البيانات الضخمة؟

Appendix C

Cloud Functions Used for the Simulation Model

Table (C.1): List of Cloud Functions and their features.

Cloud Function	Description
agro_add_stakeholder	a. Insert a new stakeholder record into the table 'agro_stakeholders'. b. Post data to the IGA API c. Update the 'agro_stakeholders' table (set 'active' field to 1). d. Insert a new record in 'mapping_stakeholders_types' table. e. Send a confirmation SMS and email.
agro_farms	f. Insert a new record in the 'agro_farms' table. g. Retrieve stakeholder's farms list and details.
agro_crops	a. Insert a new record in the 'agro_crops' table. b. Retrieve crops details.
agro_manual_source	a. Insert a new record into 'agro_manual_files' table. b. Insert a new record into 'agro_manual_fields' table.
agro_init_auto	a. Insert a new record in the 'agro_tech' table. b. Retrieve list of sources. c. Insert a new record into 'agro_tech_apis' and 'agro_tech_fields' tables. d. Creates a new CSV file on the Cloud Storage buckets initialized with the source fields as headers.
agro_crop_prod	e. Insert a new record in the 'agro_crop_prod' table. f. Retrieve list of crops production. g. Assign and activate a data source to an active production. h. Retrieve and visualize production details.
agro_auto_extract	a. Trigger automatically every 6 hours. b. Extract data using target APIs.

	c. Load data into CSV files
agro_web_scrap	a. Trigger automatically once every day. b. Extract data from webpages. c. Create and update CSV files with fields extracted from webpages.
agro_data_ETL	a. Extract data from structured or unstructured sources based on search criteria. b. Run transformation operations such as data cleaning, segregation, and aggregation on extracted data. c. Downloading transformed data in different formats. Loading transformed data into Google BigQuery tables for analytics and visualization.

Table (C.2): Cloud Function “agro_init_auto” used to initialise automated sources.

```

1. from google.cloud import storage
2. from flask import jsonify, request
3. import csv
4. import json
5. import requests
6.
7. def agro_init_auto(request):
8.     """
9.     Initializes the AgroBahrain system automatically based on the provided
10.    JSON data.
11.    Args:
12.        request (flask.Request): The HTTP request object.
13.    Returns:
14.        Tuple: A tuple containing JSON response, status code, and headers.
15.    """
16.    if request.method == 'OPTIONS':
17.        # Handling preflight CORS request
18.        headers = {
19.            'Access-Control-Allow-Origin': '*',
20.            'Access-Control-Allow-Methods': 'GET',
21.            'Access-Control-Allow-Headers': 'Content-Type',
22.            'Access-Control-Max-Age': '3600'
23.        }
24.        return ('', 204, headers)
25.
26.    # Headers for CORS
27.    headers = {'Access-Control-Allow-Origin': '*'}
28.
29.    # Parse JSON data from the request
30.    flask_json = request.json
31.
32.    # Extracting fields from JSON
33.    stakeholder_id = flask_json['stakeholder_id']
34.    tech_id = flask_json['tech_id']
35.    api_url = flask_json['api_url']
36.    data_field = flask_json['data_field']
37.    field_unit = flask_json['field_unit']
38.    field_desc = flask_json['field_desc']

```

```

40.
41.     # Add API and API fields
42.     payload = {
43.         'stakeholder_id': stakeholder_id,
44.         'tech_id': tech_id,
45.         'api_url': api_url,
46.         'data_field': data_field,
47.         'field_unit': field_unit,
48.         'field_desc': field_desc
49.     }
50.     api_url = 'http://my_domain/sim/add_api_fields.php'
51.     response = requests.get(api_url, params=payload)
52.     api_id = response.json()
53.
54.     # Define CSV fields
55.     csv_fields = ['temp_c', 'wind_kph', 'pressure_mb', 'humidity', 'uv',
56. 'created_at']
57.
58.     # Initialize Google Cloud Storage client
59.     storage_client = storage.Client()
60.     store_bucket = storage_client.bucket('agro_auto_extract')
61.
62.     # Define the path for the new CSV file
63.     stakeholder_folder = f"{stakeholder_id}/{tech_id}_{api_id}.csv"
64.     blob = store_bucket.blob(stakeholder_folder)
65.
66.     # Write CSV header
67.     with blob.open(mode='w') as f:
68.         dw = csv.DictWriter(f, delimiter=',', fieldnames=csv_fields)
69.         dw.writeheader()
70.
71.     # Response message
72.     csv_init = "API initialization Successful"
73.     return (jsonify(csv_init), 200, headers)

```

Table (C.3): Cloud Function “agro_auto_extract” used to automatically collect data every 6 hours from farms with active sensors.

```

1. from google.cloud import storage # Importing Google Cloud Storage library
2. from flask import jsonify, request # Importing Flask libraries
3. import csv # Importing CSV library
4. import json # Importing JSON library
5. from datetime import datetime # Importing datetime library
6. import requests # Importing requests library
7. import pytz # Importing pytz library
8. import pandas as pd # Importing pandas library
9.
10. # Define a Cloud Function named auto_extract that takes a request object
11. def auto_extract(request):
12.     # Check if the request method is OPTIONS
13.     if request.method == 'OPTIONS':
14.         # Define CORS headers for preflight request
15.         headers = {
16.             'Access-Control-Allow-Origin': '*',
17.             'Access-Control-Allow-Methods': 'GET',
18.             'Access-Control-Allow-Headers': 'Content-Type',
19.             'Access-Control-Max-Age': '3600'
20.         }
21.         # Return an empty response with status code 204 and CORS headers
22.         return ('', 204, headers)
23.
24.     # Define CORS headers for regular requests
25.     headers = {'Access-Control-Allow-Origin': '*'}
26.
27.     # Get Active Techs (return tech id)
28.     techIDs = []
29.     get_tech_url = 'http://my_URL/sim/get_tech_id.php' # URL to fetch tech
IDs

```

```

30. tech_response = requests.get(get_tech_url) # Send GET request to fetch
tech IDs
31. tech_li = tech_response.json() # Parse JSON response
32. for tech in tech_li:
33.     techIDs.append(tech["tech_id"]) # Extract tech IDs
34.
35. # Get tech APIs (endpoints and API IDs)
36. apis_ids = []
37. endpoints = []
38. json_techids = json.dumps(techIDs) # Convert tech IDs to JSON format
39. api_payload = {'techids': json_techids} # Payload for API request
40. get_apis_url = 'http://my_URL/sim/get_tech_apis.php' # URL to fetch tech
APIs
41. apis_response = requests.get(get_apis_url, params=api_payload) # Send
GET request to fetch tech APIs
42. apis_li = apis_response.json() # Parse JSON response
43.
44. # Extract API IDs and endpoints
45. api_ids = apis_li[0]
46. for api_id in api_ids:
47.     apis_ids.append(api_id[0])
48. api_endpoints = apis_li[1]
49. for api_endpoint in api_endpoints:
50.     endpoints.append(api_endpoint[0])
51.
52. # Create list of CSV files using tech ID and API ID
53. csv_files = [techid + '_' + apis_ids[i] + '.csv' for i, techid in
enumerate(techIDs)]
54.
55. # Get endpoints response and store it in list of dictionaries
56. endpoints_responses = []
57. for i, ep in enumerate(endpoints):
58.     endpoint_url = ep
59.     endpoint_response = requests.get(endpoint_url)
60.     endpoint_li = endpoint_response.json()
61.     endpoints_responses.append({'csvfile': csv_files[i], 'response':
endpoint_li})
62.
63. # Get endpoints response and store it in CSV files
64.
65. # Initialize Google Cloud Storage client
66. storage_client = storage.Client()
67. store_bucket = storage_client.get_bucket('agro_auto_extract') # Get
bucket by name
68.
69. # Get list of blobs (files) in the bucket
70. blobs = storage_client.list_blobs(store_bucket)
71.
72. csv_blobs = [blob.name for blob in blobs] # Extract blob names
73.
74. # Iterate through CSV files
75. for csv_file in csv_files:
76.     # Iterate through blobs
77.     for csv_blob in csv_blobs:
78.         # Check if blob name ends with CSV file name
79.         if csv_blob.endswith(csv_file):
80.             list_of_values = []
81.             list_of_columns = []
82.             blob = store_bucket.blob(csv_blob) # Get blob object
83.
84.             # Read CSV file
85.             with blob.open(mode='r') as f:
86.                 csvreader = pd.read_csv(f)
87.                 list_of_columns = list(csvreader.columns)
88.
89.             # Write to CSV file
90.             with blob.open(mode='w') as f:
91.                 for lc in list_of_columns:
92.                     if lc == 'created_at':
93.                         bah_time = pytz.timezone('Asia/Bahrain')

```

```

94.             c_at = datetime.now(bah_time)
95.             list_of_values.append(c_at.strftime('%Y-%m-%d
%H:%M:%S'))
96.         else:
97.             for e in endpoints_responses:
98.                 for key, value in e.items():
99.                     if key == 'response':
100.                         for k, v in value.items():
101.                             if lc == k:
102.                                 list_of_values.append(v)
103.             # Write columns and values to CSV file
104.             csvwriter = csv.writer(f)
105.             csvwriter.writerow(list_of_columns)
106.             csvwriter.writerow(list_of_values)
107.
108.     csv_update_completion = 'success' # CSV update completion status
109.
110.     # Return JSON response with status code 200 and CORS headers
111.     return (jsonify(csv_update_completion), 200, headers)
112.

```

الملخص

قطاع الزراعة على أعتاب عصر ثوري يدفعه التقدم التكنولوجي. تلك التقنيات تحدث ثورة في الزراعة، وتوفر فرصاً غير مسبوقة لتعزيز الكفاءة والاستدامة والمرونة. ومن أجل دمج هذه التقنيات بنجاح، يتعين على الحكومات والباحثين والمزارعين العمل معاً للتغلب على العقبات وضمان تغيير عادل وأخلاقي وشامل في قطاع الزراعة العالمي.

تواجه مملكة البحرين، وهي جزيرة صغيرة في قلب الخليج العربي، تحديات زراعية طبيعية مثل التربة القاسية والظروف المناخية، فضلاً عن نقص الموارد المائية. علاوة على ذلك، فإن قطاع الزراعة في البحرين لم يحتضن بعد إمكانات الثورة الرقمية بشكل كامل. كما يعتمد المزارعون في البحرين بشكل كبير على الممارسات التقليدية والتجارب السابقة، مع عدم توفر البيانات الزراعية أو صعوبة الوصول إليها.

تقدم الدراسة مقترح إطار عمل لإنشاء نظام وطني متكامل للبيانات الزراعية الضخمة في مملكة البحرين معتمدة على أربعة أهداف. أولاً، استكشاف مستوى الوعي والاستعداد الرقمي لدى المزارعين في مملكة البحرين. حيث يعد هذا البحث أمراً بالغ الأهمية لتعزيز القدرات والمهارات الرقمية للمزارعين. ثانياً، تبحث الدراسة في احتياجات الزراعة وديناميكيات العرض والطلب للبيانات الزراعية. وتتضمن تحديد مصادر البيانات المتاحة، وفهم المتطلبات المحددة لكل أصحاب المصلحة، ومعالجة التحديات المرتبطة بالحصول على البيانات الزراعية. بالإضافة إلى ذلك، تؤكد الدراسة على أهمية هذه البيانات لإنشاء هيكل المستخدم/المالك لإدارة البيانات الزراعية الضخمة في مملكة البحرين.

وبناءً على مكتسبات الهدفين الأول والثاني، تم انجاز الهدف الثالث وهو تطوير إطار عمل لتطبيقات البيانات الضخمة للزراعة في البحرين. ويعتمد الإطار على دمج أدوات البيانات الضخمة والخدمات السحابية مما يؤدي إلى تسهيل الوصول للبيانات المخزنة لمختلف أصحاب المصلحة. وأخيراً، يتضمن الهدف الرابع تطوير نموذج محاكاة بناءً على الإطار الذي تم تطويره في الهدف السابق. سيقوم هذا النموذج بتقييم الجدوى ومحاكاة العمليات والوظائف المختلفة لمنصة الزراعة الذكية.

ونظرًا لطبيعة البحث المتعددة المحاور، فقد تم استخدام مجموعة متنوعة من مناهج البحث. وتشمل المنهجيات استبيانًا مغلقًا تم توزيعه على أكثر من ١٠٠ مزارع ومستأجر مزرعة. القيام بزيارات ميدانية للمزارع لمراقبة وفهم العمليات الزراعية المتاحة وكيفية إدارتها. تم إجراء مقابلات شخصية مع مختلف أصحاب المصلحة في مجال الزراعة، وتم إجراء مراجعة واسعة النطاق للأبحاث ذات الصلة، والتي تتعلق بمفاهيم البيانات الضخمة الزراعية. وشملت الأدوات والموارد التكنولوجية المستخدمة في البحث أجهزة استشعار التربة اللاسلكية، وكاميرا متعددة الأطياف، والخدمات السحابية (Google Cloud Platform)، ومنصة استشعار عن بعد، وأدوات مختلفة كبيئة التطوير الأساسية.

كشفت الدراسة عن العقبات الملحوظة التي تواجه تطبيقات البيانات الزراعية في مملكة البحرين، مثل عدم وجود تقنيات وأدوات كافية لجمع البيانات، وعدم كفاية واجهات الوصول إلى البيانات، والقيود المختلفة المفروضة على تبادل البيانات. ومع ذلك، أظهرت النتائج حماسًا واضحًا للتكنولوجيا الزراعية بين الأفراد وخاصة الجيل الجديد من المزارعين، ومع ذلك يظهر أصحاب المصلحة في الزراعة فهمًا محدودًا واعتمادًا محدودًا لأدوات تكنولوجيا المعلومات والاتصالات لإدارة عملياتهم الزراعية اليومية.

بناءً على الأساليب المتبعة لاستكشاف البيئة الزراعية في البحرين، تم بناء إطار عمل لتلبية السمات المميزة على وجه التحديد مع مراعاة العقبات مثل الممارسات الزراعية على نطاق صغير وانخفاض كفاءة المزارعين في تكنولوجيا المعلومات والاتصالات. يتكون الإطار من مرحلتين: مرحلة أولية تركز على تعزيز قدرات تكنولوجيا المعلومات والاتصالات بين أصحاب المصلحة في المزارع ومرحلة ثانية توضح كيفية جمع البيانات الزراعية وتخزينها وتحليلها بشكل شبه تلقائي. ولإظهار خصائص الإطار المقترح تم إنشاء نموذج للمحاكاة في السحابة باستخدام خدمات Google Cloud، التي تتضمن واجهات برمجة التطبيقات المخصصة، وتطبيق الويب التقدمي (PWA)، والمعدات والتقنيات الزراعية الذكية.

يتطلب تحقيق الاستفادة الزراعية في البحرين خطة عمل بالتعاون والدعم من أصحاب المصلحة في مجال الزراعة وصانعي السياسات. في البداية، يجب أن يؤدي هذا التعاون إلى رفع مستوى الوعي والاستخدام نحو اعتماد تكنولوجيا المعلومات والاتصالات في الزراعة وتحويل الوعي إلى استخدام فعلي. ولتشجيع المزيد من الاستخدام، ينبغي تطوير خدمات تكنولوجيا المعلومات والاتصالات للإنتاج الزراعي

لنكون في متناول جميع المستخدمين، مع مراعاة جميع الفئات العمرية. علاوة على ذلك، فإن تكنولوجيا المعلومات والاتصالات لنشر الزراعة باستخدام التقنيات الناشئة مثل إنترنت الأشياء والبيانات الضخمة تمثل أولوية.

ان هذه الدراسة تمثل الخطوة الأولى في بناء عملية مفاهيمية لجمع وتخزين وتحويل البيانات الزراعية الضخمة في البحرين. وتماشياً مع السياسة الزراعية لمملكة البحرين والتي تمت صياغتها في أوائل عام ٢٠١٠، تمهد هذه الدراسة الطريق لمزيد من البحث في العديد من المجالات ذات الصلة بالزراعة.

جامعة البحرين

برنامج البيئة والتنمية المستدامة



كلية العلوم

(AGROBAHRAIN): نحو الزراعة الذكية في عصر البيانات الضخمة

مخطط بحث لاستكمال متطلبات الحصول على درجة دكتوراه الفلسفة في
البيئة والتنمية المستدامة

اعداد

محمد عصام حسين غانم
الرقم الجامعي: ١٩٩٨٢٨٩٩

اشراف

د. جفلة العماري
أستاذ مشارك
جامعة البحرين

مملكة البحرين
مايو ٢٠٢٤