

## A study on Digital Agriculture Data Curation Service Plan for Digital Agriculture

Hyunjo Lee\*, Han-Jin Cho\*\*, Cheol-Joo Chae\*\*\*

\*Researcher, Dept. of Computer Engineering, Jeonbuk National University, Jeonju-si, Korea

\*\*Professor, Dept. of Energy IT Engineering, Far East University, Eumseong-gun, Korea

\*\*\*Professor, Dept. of General Education, Korea National College of Agriculture and Fisheries, Jeonju-si, Korea

### [Abstract]

In this paper, we propose a service method that can provide insight into multi-source agricultural data, way to cluster environmental factor which supports data analysis according to time flow, and curate crop environmental factors. The proposed curation service consists of four steps: collection, preprocessing, storage, and analysis. First, in the collection step, the service system collects and organizes multi-source agricultural data by using an OpenAPI-based web crawler. Second, in the preprocessing step, the system performs data smoothing to reduce the data measurement errors. Here, we adopt the smoothing method for each type of facility in consideration of the error rate according to facility characteristics such as greenhouses and open fields. Third, in the storage step, an agricultural data integration schema and Hadoop HDFS-based storage structure are proposed for large-scale agricultural data. Finally, in the analysis step, the service system performs DTW-based time series classification in consideration of the characteristics of agricultural digital data. Through the DTW-based classification, the accuracy of prediction results is improved by reflecting the characteristics of time series data without any loss. As a future work, we plan to implement the proposed service method and apply it to the smart farm greenhouse for testing and verification.

▶ **Key words:** Digital agriculture, Agricultural data, Curation, Clustering, Agricultural environment data

### [요 약]

본 논문에서는 다출처 농업 데이터를 통찰할 수 있는 지식체계를 마련하고, 시간 흐름을 가지는 환경인자 분석 정보를 클러스터링 할 수 있는, 농작물 환경 인자 큐레이션 서비스 방법을 제안한다. 제안하는 큐레이션 서비스는 크게 수집, 전처리, 저장, 분석의 네 단계로 구성된다. 첫째, 수집 단계에서는 OpenAPI 기반의 웹크롤러를 이용하여 다출처 농업 데이터에 대한 수집 및 정리를 수행한다. 둘째, 전처리 단계에서는 데이터 측정 오차를 감소시키기 위해 데이터 평활화를 수행한다. 이때 온실, 노지 등의 시설 특성에 따른 오차율을 고려하여 시설 유형별 평활화 방법을 적용한다. 셋째, 저장단계에서는 대용량 농업 데이터 관리를 위해, 농업 데이터 통합 스키마 및 Hadoop HDFS 기반의 저장 구조를 제안한다. 마지막으로 분석 단계에서는 농업 디지털 데이터의 시계열 특성을 고려한 DTW 기반의 시계열 분류를 수행한다. DTW 기반 시계열 분류를 통해 시계열 데이터의 특성을 손실 없이 반영하여 예측 결과 정확도를 향상시킨다. 향후 연구로는 제안한 서비스 방법을 구현하여 스마트팜 온실에 적용하고, 테스트 및 검증을 수행할 예정이다.

▶ **주제어:** 디지털 농업, 농업 데이터, 큐레이션, 클러스터링, 농업 환경 데이터

- First Author: Hyunjo Lee, Corresponding Author: Cheol-Joo Chae
- \*Hyunjo Lee (o2near@gmail.com), Dept. of Computer Engineering, Jeonbuk National University
- \*\*Han-Jin Cho (hanjincho@kdu.ac.kr), Dept. of Energy IT Engineering, Far East University
- \*\*\*Cheol-Joo Chae (chae.cheoljoo@gmail.com), Dept. of General Education, Korea National College of Agriculture and Fisheries
- Received: 2022. 01. 28, Revised: 2022. 02. 15, Accepted: 2022. 02. 15.

## I. Introduction

농업의 방식이 기존의 정밀농업이나 스마트농업에서 생산, 유통, 소비 등 농업활동의 전과정에서 데이터를 적극 활용하는 디지털 농업으로 변화하고 있다. 또한 농업 데이터의 활용방식도 수집된 데이터를 인간이 분석·모델화하여 작업을 자동화하는 방식에서 수집된 빅데이터를 사람이 아닌 AI가 분석하여 의사결정을 내리는 방식으로 변화하고 있다[1, 2]. 농업 분야 선두 그룹인 WUR(Wageningen University & Research)의 정보기술 그룹에서는 사물인터넷관련 프로젝트를, 스마트팜 그룹에서는 인공지능 관련 프로젝트를 각각 수행하고 있으며 빅데이터와 인공지능을 활용하여 필드, 온실, 축산 분야에 직접 이용할 수 있는 혁신 연구를 진행하고 있다. 한편, 농작물 환경인자 분석·예측 기술 분야에서는 단일 온실 내 환경 정보를 모델링하여 예측하고 검증하는 연구가 진행되고 있으며 주로 PCA 분석을 통해 데이터를 전처리하고 BPNN, NNAR(Neural Network Automatic Regression) 등의 알고리즘을 이용하여 단일 온실 내 환경을 제어하는 연구가 진행되고 있다[3]. 그러나 기존 방법으로는 온실 내 환경 제어 및 관리가 어렵기 때문에, 이를 완화하기 위해 빅데이터 기반의 인공지능 알고리즘을 이용하여 환경 정보의 연관관계를 분석하여 제어하는 연구가 진행되고 있다[4-7].

농업 빅데이터는 다양한 출처에서 특성에 따라 정보가 생성되고 있으나 그에 대한 수집 기술은 아직까지 통계, 정형화된 데이터를 이용한 서비스 분야에서만 활용되고 있다. 이는 농업에서 생산·유통·소비까지 전반적인 분야에 걸쳐 데이터를 수집하여 활용하는데 한계가 있기 때문이다. 또한 농업 데이터가 다양한 출처로부터 다양한 형태로 생성되기 때문에, 이를 통합된 데이터 형식으로 수집·처리하기 어려운 문제가 존재한다. 이러한 문제를 해결하기 위해 본 논문에서는 디지털 농업으로의 전환을 지원하고, 디지털 농업 데이터를 통찰할 수 있는 지식 체계를 제공하는 사용자 큐레이션 방안에 대해 제안한다.

## II. Digital Agriculture Technology

디지털 농업에서는 기존의 정밀 농업이나 스마트 농업 보다 생산, 유통, 소비 등 농업 활동의 전과정에서 데이터를 적극 활용한다. FAO(The Food and Agriculture Organization), CEMA(European Agricultural Machinery), EU 등에서 디지털 농업이라는 용어를 사용

하고 있으며, 디지털 농업이란 결국 농업에 관련된 데이터를 수집, 저장, 관리, 결합, 분석, 공유하여 새로운 가치를 창출하는 것을 의미한다.

국내에서는 2016년부터 한국형 스마트 팜 기술 개발을 위해 외부 환경에 대응 가능한 인공 지능형 온실을 개발하고 있다[8]. 한국형 스마트 팜 개발에서는 스마트팜 기술을 1~3세대 모델로 분류하고 있으며, 현재 1세대 기초 기반 기술 개발이 완료된 상태로써 빅데이터를 활용한 작물 생육 모델링 기술 개발에 집중하고 있는 실정이다. 일본의 Kubota社は NTT社와 함께 농작물의 생육상태를 빅데이터로 분석하여, 수확량이나 농약살포 등의 작업 상태 파악 가능한 시스템을 개발하고 있다. NTT社は GPS, 수확량 예측정보 제공을 담당하고 있으며, Kubota社は 농업용 로봇을 이용한 생육상태 분석 및 비료 또는 농약을 살포 가능한 기술의 상용화를 수행하고 있다[9]. 미국 Google社は 토양, 수분, 작물 건강에 대한 데이터를 수집해 종자, 비료, 농약 살포에 도움을 주는 의사결정 기술을 상용화하고 있다. Monsanto社は Climate Corporation을 통해 미국 250만개 지역의 주요 기후정보 데이터, 수확량 데이터, 토양 데이터를 기반으로 지능형 기술을 상용화하고 있다[10]. 네덜란드 Priva社は 오랜 기간 축적해온 작물 재배 데이터로부터 작물의 품종별로 생산성을 최대화할 수 있는 최적의 생육 조건을 분석하여 재배 기간 동안 적절한 환경이 지속적으로 유지되도록 온실 환경을 자동으로 관리하는 시스템 기술을 보유하고 온실 내의 작업 효율성을 높일 수 있는 기술을 상용화하고 있다. Fig. 1은 Climate Corporation과 Priva社에서 제공하는 디지털 농업 기술의 예를 보여주고 있다.



Fig. 1. Example of Digital Agriculture Technology

농축산물 생육 및 환경 데이터란, 농축산업 생산·유통·소비의 전반적인 과정에서 센서 또는 사람에 의해 수집 가능한

데이터 중 농축산물의 생육·생장에 영향을 미치는 환경 데이터 및 생육·생장 상태 데이터를 의미한다. 농축산물 생육 환경 데이터는 센서를 이용해서 측정 가능한 공통적인 환경 정보를 수집하며, 온도(°C), 습도(%), 이산화탄소농도(ppm), 광량(Wm-2), 토양수분함량(vol), 토양pH(pH), 수액EC(dS) 등을 포함한다. 농축산물 생육 상태 데이터는 품목의 외관을 수동으로 측정하여 수집하며 대상에 따라 측정하는 기준이 다르다. 예를 들어 딸기의 경우, 화방(개), 초장(cm), 엽수(개), 엽면적(cm<sup>2</sup>) 등을 측정한다. 센서로 자동 수집되는 환경 데이터와 달리, 생육 상태 데이터의 경우 사람이 수동으로 수집해야하기 때문에 획득이 어렵다. 농축산물 생육 및 환경 데이터는 환경과 품목의 생육 관계를 정량화한 데이터이기 때문에, 이를 분석함으로써 생산량 증가 및 품질 향상에 기여할 수 있다. 그러나 기존 농축산물 생육 및 환경 데이터 수집 및 활용 시스템은 환경 및 생육 상태 데이터의 시계열 특성을 고려하지 않기 때문에, 분석 결과의 정확도 및 활용도가 저하되는 문제점이 존재한다[11, 12].

### III. Digital Agriculture Data Platform

#### 1. Smart Farm Korea Service

스마트팜 코리아는 비닐하우스, 유리온실, 축사 등에 ICT 기술을 접목하여 원격·자동으로 작물과 가축의 생육 환경을 적정하게 유지·관리할 수 있도록 농가에 서비스를 제공하는 플랫폼이다. 이를 위해 작물의 생육 정보와 시설의 환경 정보에 대한 데이터를 기반으로 최적의 생육 환경을 조성 및 관리하고 축적된 정보를 바탕으로 분석 및 컨설팅을 지원하는 서비스를 스마트팜 농가에 지원한다. 시설에 따라 시설원예(딸기, 토마토, 파프리카 등), 축산(한우, 양돈, 양계 등), 노지(짜리고추, 대파, 마늘 등)에서 환경 정보, 시설 제어 정보, 그리고 주요 생육 정보를 수집하며, OpenAPI, 파일 데이터, 시각화 등 다양한 방식으로 데이터를 제공하고 있다[13].

데이터 수집량 추세, 스마트팜 보급현황, 품질·소득·생산성 향상 추이 등은 그래프로 시각화하여 정보를 제공한다. 예를 들어 데이터 수집량의 경우, 2021년 10월 기준 시설 원예 112,909 건(단위: 천), 노지작물 23,276 건(단위: 천), 축산 27,925 건(단위: 백)의 데이터를 수집하였음을 알 수 있다. 시설원예 작기 데이터, 수분센서 오차분석, 드론 생육조사 등의 데이터는 공공 데이터 법에 따라 누구나 이용 가능하고 자유롭게 활용이 가능한 파일 데이터로 제공된다. 또한, 원예 작물 이미지 정보 및 분야별 작물의 환경정보,

시설제어 정보, 그리고 생육 정보의 경우, OpenAPI로 제공된다. OpenAPI 데이터의 경우, 활용 신청 후 허가를 받아야 하며, 활용 후 그 결과를 등록해야 한다. Table 1은 시설원예에서 수집되는 작물의 환경 정보를 나타낸다. 환경정보는 크게 시설 내부의 온도, 습도, 조도, CO<sub>2</sub>농도, 광량 등의 내부 환경과, 지습, 토양의 PH, 관수량 등의 토양환경, 강우, 풍향, 풍속 등의 외부환경, 그리고 영양액 공급 관련 정보인 양액 정보로 구성되어 있으며, 센서를 통해 시설에서 자동 수집되어 제공된다.



Fig. 2. Smart Farm Korea Platform Service

시설의 제어정보는 위치별 환풍기 동작시간, 창문 개폐 조절 시간 등 환경 정보 모니터링 중 생육 환경 최적화를 위해 시설 내 장치가 동작한 정보를 자동으로 수집하여 제공한다. 주요 생육정보는 작물별로 생육변화를 확인하기 위해 수집하는 특징 정보들로 조사원이 직접 해당 스마트팜 농가에 방문하여 수집하며, 일반적으로 1주일에 1회 측정하여 수집한다. 시설원예 작물로는 토마토, 파프리카, 딸기 등이 대표적이다.

Table 1. Environmental information on facility horticulture provided by Smart Farm Korea

Class	Sect	Unit
Internal environment	intHumidity	%
	intLightlv	lux
	IntCO2	ppm
	intTemp	degree(°C)
	intSolarradiation	W/m-2s
	intQuantityoflight	μmol/m-2·s
	intRelativehumidity	%
Soil properties	intAbsolutehumidity	%
	soilHumidity	%
	soilEC	dS/m
	soilPH	ph
	soilAmountofwater	liter
	soilTemp	degree(°C)

External environment	extSolarradiation	w/m <sup>2</sup>
	extRainfall	mm
	extTemp	degree(°C)
	extHumidity	%
	extWinddirection	degree(°)
	extWindspeed	m/s
Nutrient information	extRainfalldetect	Y/N
	nutrientSupplyPH	ph
	nutrientOutputPressure	%
	nutrientTotalSupply	l
	nutrientDrainageEC	dS/m
	nutrientDrainagePH	ph
	nutrientSupplyEC	dS/m
	nutrientMediumtemp	degree(°C)
	nutrientMediumweight	kg
	nutrientSupplyrate	cc
	nutrientTotaldrainage	l
	nutrientNumberofsupply	number
	nutrientMoisture	g
	nutrientSoiltemp	degree(°C)
nutrientDrainagerate	cc	
nutrientSoilhumidity	%	

생육환경 데이터는(키워드 ‘생육’을 통해 검색) 파일 데이터 11건, OpenAPI 13건 만이 존재한다. 전체 공개 데이터의 수가 파일 데이터 48,900건, OpenAPI 7,940건, 그리드 8,753건 인 것을 감안하면 매우 적은 수의 디지털 농업 데이터만이 제공되고 있다. 이는 스마트온실, 스마트노지 및 스마트축산 관련 농가 비율이 전체 관련 농가의 50%이하이며, 스마트팜 농가의 60%정도만이 실제 데이터 수집-제공을 수행하기 때문이다[14, 15].

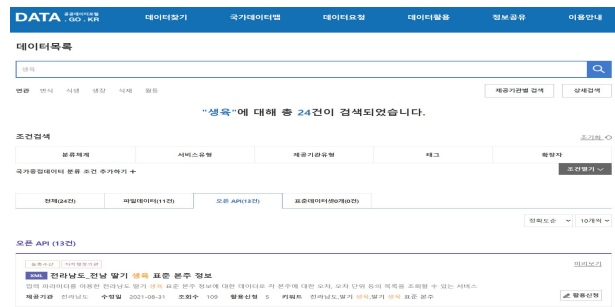


Fig. 3. Public data portal platform service

2. Public Data Portal Service

공공 데이터 포털의 농림축산 분류의 데이터는 파일 데이터 2,626건, OpenAPI 814건, 그리드 159건이며, 농가 수, 작물 재배 현황, 품종 검색 이미지 데이터 등이 다양한 데이터를 포함한다. 그러나 디지털 농업을 위한 핵심 데이터인

3. Rural Development Administration

농촌진흥청에서는 공공데이터의 제공 및 이용활성화에 관한 법률에 따라 농업관련 디지털 데이터를 제공하고 있다.

Table 2. Description of pest information service

Class	Service name	Description
Pest search service	Pest search	A information search service for pest which can occur in crops
	Detailed information provision on pest	A service of providing detailed information and image of pest
	Pathogen search	A information search service for pathogen on crops
	Detailed information provision on pathogen	A service of providing detailed information and image of pathogen
	Injurious insect search	A search service for injurious insect of crops
	Detailed information provision on Injurious insect	A service of providing detailed information and image of injurious insect
	Insect search	A search service for various kinds of insects
	Detailed information provision on insect	A service of providing detailed information and image of insect
	Weed plant search	A search service for weed plants which parasites on crops
	Detailed information provision on weed plant	A service of providing detailed information and image of weed plant
	Image search on crops with main category	A image search service for crops with the main category
	Image search on crops with sub category	A image search service for crops with the sub category
	Image search on crops with child category	A image search service for crops with the child category
	Natural enemy search	A information search service for natural enemy
Detailed information provision on natural enemy	A service of providing detailed information and image of natural enemy	
Pest forecast service	Integrated search	A integrated search service for all the information on above
	Pest forecast map	A GIS based map service for providing pest forecast information
	Pest forecast on the selected area	A information provision service for pest outbreak forecast on the selected area according to the time series
	Comparison of pest forecast and pest prediction	A service of comparing the pest forecast information and pest prediction information
Pest prediction service	Decision making support for controlling rice blast disease	A service of providing helpful information for decision making on rice blast disease control
	Pest prediction search	A search service for pest prediction information
	Detailed search for pest prediction information (cities, province)	A onetime service of providing pest prediction information for cities and provinces
	Detailed search for pest prediction information (city, county, district)	A onetime service of providing detailed pest prediction information for cities, counties and districts

제공하는 주요 데이터는 스마트팜 우수농가 공개용 데이터 (완숙토마토, 딸기, 파프리카), 농업 관련 전문서적 데이터, 농업경영관련 데이터, 그리고 병해충 관리 정보 데이터이다. 스마트팜 우수농가 공개용 데이터는 스마트팜 우수농가로 선정된 농가의 환경정보, 시설제어정보, 주요 생육정보를 공공데이터 포털을 통해 제공하며, 전체적인 데이터의 구성은 스마트팜 코리아의 OpenAPI를 통해 제공되는 것과 동일하다. 국가농작물병해충관리시스템에서는 병해충에 대한 예찰, 예측, 상담 서비스 및 병해충 정보 데이터를 제공하는 시스템이다. 병해충 정보 데이터는 OpenAPI 형태로 제공한다. Table 2는 병해충 정보 데이터를 보여준다[16].

#### IV. Proposed Method

디지털 농업에서 농작물 생육 환경을 제어하는 가장 기본적인 방법으로 센서를 이용한 실측 데이터를 기반으로 목표치를 설정하여 농작물의 생육 환경을 조절하고 있다. 그러므로 시간기반의 지역·품종에 따른 주요 농작물 환경인자(온도, 습도, 조도, CO<sub>2</sub> 등)를 수집하여 분류하는 것이 가장 중요한 요소이다. 이러한 농작물 환경인자는 다양한 플랫폼을 통해 생성·유통되고 있으나 비정형 데이터의 경우 수집 후 분석에 바로 사용하기에는 부적절한 형태를 이루고 있다. 또한 일부 데이터의 경우 부정확한 재현성 및 반복적인 문제가 존재하기 때문에 다출처 농업 데이터에서 수집한 정보를 구조화하여 연관관계를 분석 후 클러스터링을 통해 시계열 기반의 농작물 환경인자 정보를 분류할 수 있는 방안을 Fig.4와 같이 제안한다.

다양한 온라인 플랫폼에서 환경정보, 제어정보, 생산정보, 지역정보, 병충해 정보 등의 농업 데이터 수집을 위해 각 플랫폼에서 제공하는 OpenAPI와 Python을 이용하여 웹 크롤러를 개발하여 데이터를 추출한 후, 추출된 데이터는 Hadoop HDFS를 이용하여 각각의 데이터별로 구분하여 저장한다. Fig.5은 본 논문에서 제안하는 농작물 환경인자 정보 데이터베이스 스키마이다.

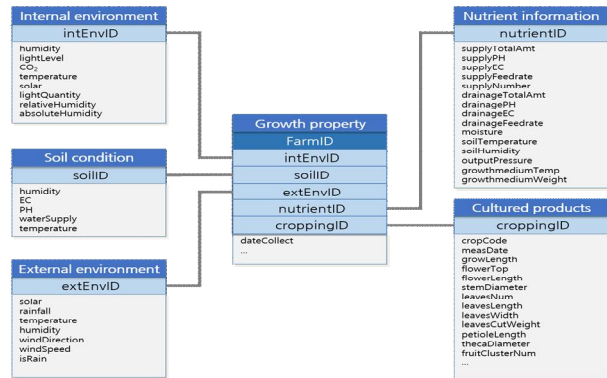


Fig. 5. Crop environmental factor information database schema

디지털 농업에서 농작물의 생산성 및 품질에 직접적인 영향을 미치는 농작물 환경인자들을 계획적으로 관리할 수 있다는 장점 때문에 사회적 요구가 증가하고 있다. 이러한 농작물 환경인자들은 서로 복잡한 상관관계를 형성하고 있으며, 데이터의 이질성이 높기 때문에 기존의 계측 데이터를 비교하는 방식으로는 한계가 있다. 기존 SVM(Support Vector Machine) 방법보다는 다중회귀 통계모형이 성능이 높고, 인공 신경망 기반의 딥러닝 모델들이 기존 과정기반 모델보다는 성능이 우수하다는 연구결과가 발표되었지만 기존의 방법들은 제한된 데이터를 사용하고 있기 때문에 실제 적용에는 어려움이 있다[17]. 또한 인공 신경망을 이용한 연구로 농작물 환경인자 정보를 PCA(Principal Component Analysis)로 전처리 후 BPNN(Back-Propagation Neural Network)을 이용하는 선행연구가 있었으나 회귀모델에 비해 성능이 우수하지만 실제 농작물 환경인자 예측 문제는 해결하지 못하였다[18].

그러므로 본 논문에서는 시계열 데이터 분석·예측에 유용하다고 알려진 RNN(Recurrent Neural Network) 기반의 딥러닝 알고리즘을 이용하여 반대 역학 모델을 학습하고 예측할 수 있는 방안을 Fig.6과 같이 제안한다. 농작물 환경인자의 시계열 특성을 반영하기 위해 DTW(Dynamic time wrapping)시계열 분류기법을 이용하여 수집된 데이터를 재분류한다. DTW 기반 시계열 데이터로 분류하기 위해 길이가

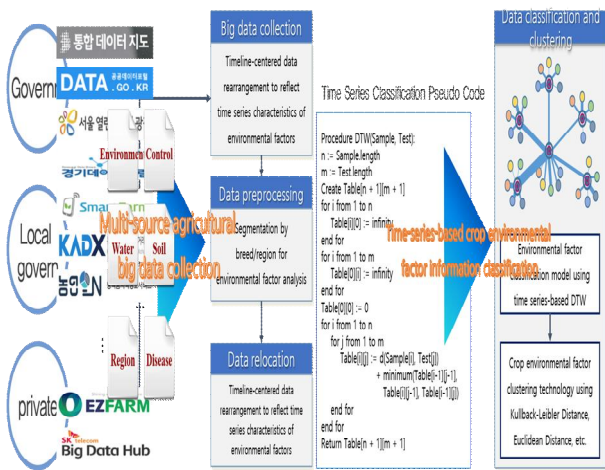


Fig. 4. Multi-source agricultural big data collection and processing and data classification and clustering flow for collecting environmental factors of crops

n개인 시계열 데이터 E와 길이가 m개인 시계열 데이터 C를 정의하고  $DTW(E,C) = \min \sum_{k=1}^m w_k$  와 같이 E로부터 C로 매핑 되는 최소경로로 계산하여 데이터를 재분류한다. 농작물 환경인자 클러스터링하기 위해 Kullback-Leibler Distance, Euclidean Distance 등의 모델을 사용하여 시간 지연과 잡음에 강한 방법을 채택하여 클러스터링 한다.

데이터 분석을 위해서는 정확한 데이터 값이 필요하지만 농작물 환경인자 데이터의 경우 대부분 센서로부터 획득된 데이터 값이기 때문에 데이터 전처리 과정이 필요하다. 농작물 환경인자 데이터의 전처리를 위해 데이터 평활화 과정을 수행하며, 시계열 특성을 반영하여 온실 데이터의 경우 이중지수 평활법(double exponential smoothing)을 이용하며 필드 데이터의 경우 Winters 계절지수평활법(Winters Seasonal Exponential Smoothing)을 사용할 수 있다.

전송한다. 스마트 팜 센서는 환경 데이터를 전달하기 위해 유선의 경우에는 IEEE802.3, PLC, RS-232C 등을 사용하고 무선의 경우에는 IEEE802.15.4, IEEE802.15.4e, IEEE802.11n 등을 사용한다. 사용자 어플리케이션에서 스마트팜에서의 환경정보를 수집하면 농작물 환경인자 정보를 분류하고 시계열성을 고려한 RNN, LSTM, ARIMA 모델을 이용하여 예측 정보를 제공한다. 그리고 최적 농작물 환경인자 큐레이션을 위해 사용자에게 농작물 품종에 따른 최적 생육 주기를 초기·중기·말기로 구분하여 온도, 습도, CO<sub>2</sub>, EC 등 변화에 따른 생육 특징, 토양 조건 등의 기준 정보를 제공하여 의사 결정을 내리는데 도움을 줄 수 있다.

### V. Conclusions

농업 빅데이터는 다양한 출처에서 각각의 특성에 맞추어 생성되고 있으나, 그에 대한 수집 기술은 아직까지 통계 또는 정형화된 데이터를 기반 한 서비스에만 활용되고 있다. 이러한 이유는 농업에서 생산·유통·소비까지 넓은 분야에서의 데이터 수집으로 인해 활용에 한계가 있으며, 다양한 출처로부터의 다양한 형태의 데이터를 수집·처리하기 어렵기 때문이다. 또한 농업에서 수집하는 환경인자 정보는 시계열 특성을 가지고 있기 때문에, 이를 고려하지 않는 기존의 연구방식에서는 타임라인을 사용하여 환경 인자를 분석할 경우 수집된 시간 이외에는 농작물 환경정보의 연관관계 분석·예측이 어려운 문제점이 존재한다. 즉, 기존 연구들은 특정 시간에서의 정보와 전·후 시간 관계의 일관성에 대한 고려가 부족하다 할 수 있다. 이를 해결하기 위해 본 논문에서는 다출처 농업 데이터를 통찰할 수 있는 지식체계를 마련하고, 시간 흐름을 가지는 환경인자의 특성을 고려한 클러스터링을 통해 농작물 환경 인자를 큐레이션 할 수 있는 서비스 방법에 대해 제안하였다. 향후 연구로는, 첫째, 제안하는 서비스 방법을 기반으로 최적의 농작물 환경인자 큐레이션 기술 적용을 위한 시나리오를 작성하고, 둘째 시나리오를 바탕으로 분석·예측 정보를 제공하는 프로토타입을 개발하고, 마지막으로 개발한 시스템을 스마트팜 온실에 적용하여 테스트 및 검증을 수행할 예정이다.

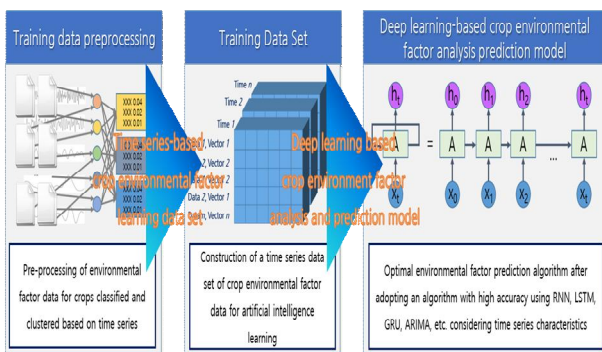


Fig. 6. Time series-based deep learning data set construction and optimal crop environmental factor analysis and prediction algorithm development flow

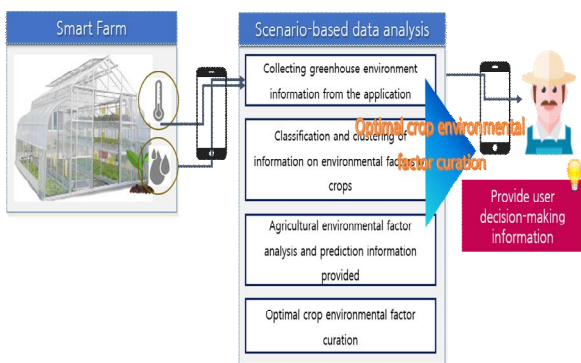


Fig. 7. Proposed of Agricultural Data Curation Service

사용자의 대내외 농업 환경 대응을 위해 지역·품종별 환경인자 정보들을 클러스터링하여 제공하고 사용자 의사결정에 도움을 주기 위해 시계열 상 환경인자 정보의 분석·예측 효율성을 향상시키기 위한 방법을 Fig.7과 같이 큐레이션 할 수 있다. 스마트 팜에 부착되어 있는 센서들은 유·무선 네트워크를 이용하여 환경 데이터를 획득하여 게이트웨이로

### ACKNOWLEDGEMENT

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT)(NRF-2021R1F1A1063042)

## REFERENCES

- [1] J. Park, D. Suh and J. Lee, "The future of agriculture, digital agriculture," Korea Rural Economic Institute Publishing, 2021.
- [2] W. S. Lee, K. Son, D. Jun, and Y. Shin, "Big Data Activation Plan for Digital Transformation of Agriculture and Rural," KIPS Trans. Softw. and Data Eng., Vol.9, No.8, pp.235-242, July 2020. DOI: 10.3745/KTSDE.2020.9.8.235
- [3] S. L. Patil, H. J. Tantau and V. M. Salokhe, "Modelling of tropical greenhouse temperature by auto regressive and neural network models," Biosystems engineering, Vol. 99, No. 3, pp 423-431, March 2008. DOI: 10.1016/j.biosystemseng.2007.11.009
- [4] D. F. Nettleton, D. Katsantonis, A. Kalaitzidis, N. Sarafijanovic-Djukic, P. Puigdollers, and R. Confalonieri, "Predicting rice blast disease: machine learning versus process-based models," BMC bioinformatics, Vol. 20, No. 1, pp 1-16, October 2019. DOI: 10.1186/s12859-019-3065-1
- [5] J. H. Kim, J. S. Kim, and Y. S. Jhang, "A Design and Implementation for a Realtime Monitoring and Controlling System in the Stockyard," Journal of The Korea Society of Computer and Information, Vol. 14, No. 10, pp 167-174, October 2008.
- [6] Y. W. Lee et al., "Implementation of Facility Management System for Plant Factory," Journal of The Korea Society of Computer and Information, Vol. 16, No. 2, pp, 141-152, Feb. 2011. DOI: 10.9708/jksci.2011.16.2.141
- [7] B. K. Son, "Mounting Detection System of Hanwoo Based on Deep Learning," Journal of next-generation convergence technology association, Vol. 5, No. 1, pp 49-56, Feb. 2021. DOI : 10.33097/JNCTA.2021.05.01.49
- [8] M. W. Han, G. Son, Y. S. Ryu, S. H. Lee and H. G. Lee, "Trends in Artificial Intelligence R & D Trends in Major Countries and Prospect of Agriculture Application," Ministry of Agriculture, Forestry and Fisheries Publishing, 2016.
- [9] K. H. Choi, "Trends in intelligent information technology in the era of the 4th industrial revolution and the direction of agricultural R&D," Korea Institute of Planning and Evaluation for Technology in Food, Agriculture, and Forestry Publishing, 2017.
- [10] H. J. Lee, "The Fourth Industrial Revolution and the Future of Agriculture: Smart Farms and the Sharing Economy," Korea Rural Economic Institute Publishing, 2017.
- [11] K. Y. Lee, J. J. Jeong, S. H. Kim, C. M. Lim, and S. S. Lee, "A Study on the Design of Data Collection System for Growing Environment of Crops," Journal of Korea Information Electronic Communication Technology, Vol.11, No.6, pp 764-771, Dec. 2018. DOI: 10.17661/jkiect.2018.11.6.764
- [12] J. Y. Lee, C. B. Moon, and B. M. Kim, "Big Data Model for Analyzing Plant Growth Environment Informations and Biometric Information," Journal of the Korea Industrial Information Systems Research, Vol.25 No.6, pp 15-23, Dec. 2020. DOI: 10.9723/jksis.2020.25.6.015
- [13] Smart Farm Korea. <https://www.smartfarmkorea.net>
- [14] Publicdata Portal. <https://www.data.go.kr/>
- [15] Smart farm status survey and performance analysis. Korea Agency of Education, Promotion and Information Service in Food, Agriculture, Forestry and Fisheries Publishing, 2021.
- [16] National Crop Pest Management System. <https://ncpms.rda.go.kr>
- [17] R. Kaundal, A. S. Kapoor, and G. P. Raghava, "Machine learning techniques in disease forecasting: a case study on rice blast prediction," BMC bioinformatics, Vol. 7, No. 1, pp 1-16, Nov. 2006., DOI: 10.1186/1471-2105-7-485
- [18] F. He, and C. Ma, "Modeling greenhouse air humidity by means of artificial neural network and principal component analysis," Computers and Electronics in Agriculture, Vol. 71, pp S19-S23, April 2010. DOI: 10.1016/j.compag.2009.07.011

## Authors



Hyunjo Lee is received the B.S., M.S., and Ph.D, degrees in Computer Engineering from Chonbuk National University in 2006, 2008, and 2014, respectively. From 2015 to 2016 he was a senior member of the research

staff in the KISTI, Daejeon, Korea, where he did research on HPC system. Since then, he is a researcher and adjunct instructor in the Chonbuk National University. His research interests include spatial network database, parallel query processing, and encrypted query processing.



Han-Jin Cho received the B.S., M.S. and Ph.D. in Computer Engineering from Hannam University, Korea in 1997, 1999 and 2002. Since then, he has been a Full Professor with the Department of Energy IT

Engineering, Far East University, Korea. He was awarded a certificate from the Ministry of Knowledge Economy, Korea in 2012. He is interested in Bigdata, AI and Information Service.



Cheol-Joo Chae received his Ph.D. degree in computer engineering from Hannam University, Daejeon, Korea, in 2009. From 2009 to 2013 he was a senior member of the research staff in the ETRI, Daejeon,

Korea, where he did research on network communication system. From 2013 to 2016 he was a senior member of the research staff in the KISTI, Daejeon, Korea, where he did research on information system. Since then, he has been a Professor with Korea National College of Agriculture and Fisheries. He is interested in Bigdata, AI and Information Service.