Deriving Experiments from E-SECO Software Ecosystem in the Technology Transfer Process for the Livestock Domain

Jonas Gomes
Federal University of Juiz de Fora
Juiz de Fora, Brazil
jonas.gomes@estudante.ufjf.br

Vinícius Carvalho Lopes
Federal University of Goiás
Goiânia, Brazil
University of Notre Dame
South Bend, USA
vinilopes03@gmail.com

Valdemar V. Graciano Neto
Federal University of Goiás
Goiânia, Brazil
valdemarneto@ufg.br

Roberto Felício de Oliveira
State University of Goiás
Posse, Brazil
roberto.oliveira@ueg.br

Mohamad Kassab
Penn State University
Malvern, USA
muk36@psu.edu

José Maria N. David,
Regina Braga
Federal University of Juiz de Fora
Juiz de Fora, Brazil
{jose.david,regina.braga}@ufjf.edu.br

Wagner Arbex
Federal University of Juiz de Fora
Embrapa – Dairy Cattle
Juiz de Fora, Brazil
wagner.arbex@embrapa.br

ABSTRACT

The process of transferring technology from research institutes to industry involves benchmarking it in exhaustive experiments to assure it reaches the established quality criteria. This is also true for the livestock domain, in which the technologies developed to sustainably raise animals production are submitted to experiments while preserving their health and wellness. However, since such institutions often conduct several parallel innovation projects, the establishment of an infrastructure to support those experiments can be costly, repetitive, and error-prone. For that purpose, we developed E-SECO, a software ecosystem that encapsulates a life-cycle model for scientific experiments and its supporting platform and actors. The main contribution of this paper is presenting how the E-SECO architecture was successfully applied to create a livestock architecture (named e-Livestock architecture) from which two different (and independent) scientific experiments involving real systems were deployed and executed in the livestock domain. The first experiment involved a Compost Barn production system, i.e., the environment and surrounding technology where bovine milk production takes place; whilst the second experiment involved an automated monitoring environment for aviaries. Preliminary results showed the effectiveness of E-SECO to (i) abstract concepts of scientific experiments for livestock domain, (ii) support reuse and derivation of an architecture to support engineering real systems for different livestock sub-domains, and (iii) support the experiments towards a future transfer of technology to industry.

CCS CONCEPTS

• **Applied computing → Agriculture**; • **Software and its engineering → Software architectures**.

KEYWORDS

Compost barn, livestock, aviary, software ecosystem, experiments, dairy cattle.

ACM Reference Format:

Jonas Gomes, Vinícius Carvalho Lopes, Valdemar V. Graciano Neto, Roberto Felício de Oliveira, Mohamad Kassab, José Maria N. David,, Regina Braga, and Wagner Arbex. 2022. Deriving Experiments from E-SECO Software Ecosystem in the Technology Transfer Process for the Livestock Domain. In *10th IEEE/ACM International Workshop on Software Engineering for Systems-of-Systems and Software Ecosystems (SESoS'22), May 16, 2022, Pittsburgh, PA, USA.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3528229.3529386

1 INTRODUCTION

The world lives an intense population increase, which demands for the development of techniques to expand the food production in a sustainable way. In that direction, livestock plays a prominent role. Scientific advances are imperative to foster food production, promoting solutions to enable large-scale delivery while preserving sustainability and animals health and wellness. Several countries have created research institutes focused on the development of technologies to support advances on food production,

SESoS'22 , May 16, 2022, Pittsburgh, PA, USA

ACM ISBN 978-1-4503-9334-8/22/05…$15.00
https://doi.org/10.1145/3528229.3529386

such as EMBRAPA[1] (an acronym for Brazilian Agricultural Research Corporation) in Brazil, Food and Agriculture Organization of the United Nations[2] (FAO), and the Institute of Food Research, in United Kingdom[3]. Those institutions are research-oriented and produce innovative solutions to be transferred to industry later. However, before transferring any technology to industry, we need to submit it to intense experimentation, i.e., benchmarking the technology under controlled environment and variables to assure it matches quality metrics to be transformed in a commercial product. However, historically, despite the similarities between the experiment processes and procedures conducted in the multiple research projects developed in a same research institute, the experiments have been recurrently designed from scratch, consuming effort, being costly and error-prone [1]. For solving that problem, E-SECO was created. E-SECO is a software ecosystem whose underpinning platform supports the execution of scientific experiments [6].

The main contribution of this paper is to show how E-SECO has been used to derive an architecture so-named e-Livestock Architecture and, from that architecture, to derive different conforming systems and corresponding experiments for two different livestock sub-domains: Compost barn and aviaries. The compost barn experiment was carried out in a joint real project between the Federal University of Juiz de Fora and a EMPRAPA research unit in the same city (Juiz de Fora, Minas Gerais), while the aviary experiment was independently carried out in the context of a master's research project developed in the Federal University of Goiás and deployed *in loco*, in a farm in Posse, Goiás. Regardless of the geographical distance between the institutions, both experimental platforms were devired in conformance to E-SECO, gathering evidence to support us to conclude that E-SECO can be considered well-succeeded to foster reuse scientific experiments artifacts and steps towards enabling a suitable evaluation of the proposed technological solutions and a subsequent likely technology transfer to industry and commerce.

For presenting the results of our investigation, this paper is organized as follows: Section 2 presents a brief background. Section 3 presents E-SECO with details for the prior main contributions of our research, besides presenting the e-Livestock architecture derived from E-SECO for conducting experiments in livestock domain. Section 4 presents the derivation of two different experiments using E-SECO, besides comparing our results with related work. Finally, concluding remarks are presented in Section 5.

2 BACKGROUND

In the agricultural domain, scientific experiments involve interactions between researchers, including aspects such as (i) the use of a large volume of heterogeneous data that must be processed, stored, and analyzed to support the experiments and (ii) the need for supporting distributed computing resources and services. Furthermore, agriculture domain requires intense relationships between resources and applications and between geographically distributed researchers. In this context, agricultural research institutions had opened their frontiers to collaborate, similarly to what was done by software companies [3], giving rise to a new concept of development, where several software solutions, companies, and developers adhere to a common platform. This approach is known as Software Ecosystem.

A Software Ecosystem (SECO) is a set of actors who collaborate and interact with a common market. Interactions are often underpinned by a common technological platform that operates through the exchange of information, resources, and artifacts [21]. A Software Ecosystem (SECO) can be defined by its relationships with companies, users, and other stakeholders in providing services by the participating SECO systems [14, 19, 26, 29, 30]. Therefore, the architecture of a SECO must be flexible, as it can integrate with external processing platforms, which evolve independently and continuously. These relationships occur to generate greater value for SECO, which requires opening its borders where third-party applications start to connect and benefit from its services, generating value for the parties involved. Therefore, the SECO architecture must be extensible and flexible [9, 11]. Furthermore, the architecture of a SECO needs to be scalable since it should be extensible, which can lead to a sudden and unexpected growth of requests for services [10]. The participating software systems, companies, users, and other parties must then promote applications and data integration in that scenario.

Considering the agricultural domain, there are some specific requirements that must also be fulfilled by the SECO platform. For example, farmers are increasingly using sensors to support their daily activities. Based on this premise, farmers can extract information and make strategic decisions. The combination of data collection and processing can provide accurate support for decision making in agribusiness [24], as well as allowing these decisions to be auditable. Data accuracy is one of the main goals of using IoT in Digital Precision Livestock, bringing some challenges such as reproducibility, authenticity and tracking of data generated by sensors in smart agriculture. Therefore, it is necessary to carry out analysis and build quality models, in order to guarantee the reliability and quality of the data.

The use of techniques related to data provenance can help in data traceability, allowing to manage the context of data production, use, and analysis [4]. Data Provenance refers to "contextual elements that describe information in the past" [4]. Therefore, information on data provenance is essential for researchers to understand, reproduce, examine, and audit the experiment results. Data Provenance can be considered information about the parts involved in the production of an object [28] considering the activities and agents involved. However, information provenance is critical for scientific workflows to (i) support the computational experiments sharing, (ii) results interpretation and (iii) problems diagnosis, as proposed by [5, 13, 17, 32]. The data provenance helps not only in the analysis and tracking of production and animal welfare data but also in data related to the environment, sustainability, and economics [2]. By exploring the relationships between data, considering its provenance, combined with intelligent analysis techniques, a more efficient support in decision-making in agriculture 4.0 can be achieved. Next section presents E-SECO software ecosystem.

[1]https://www.embrapa.br/
[2]https://www.agriculture-xprt.com/companies/food-and-agriculture-organization-of-the-united-nations-fao-26929
[3]https://www.agriculture-xprt.com/companies/institute-of-food-research-11209

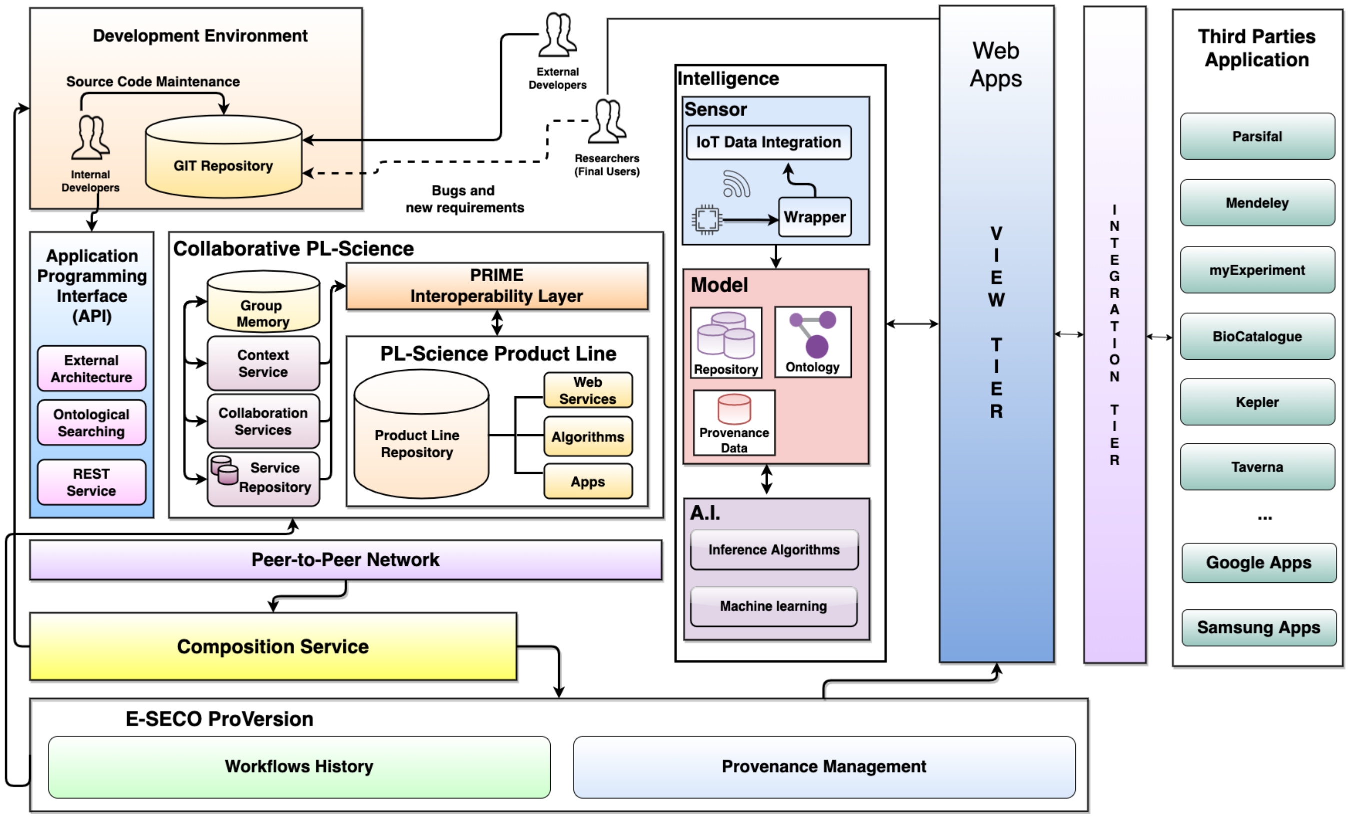

Figure 1: Overview of the E-SECO ecosystem platform highlighting the solution.

3 E-SECO: A SCIENTIFIC SOFTWARE ECOSYSTEM

E-SECO [1, 6] is a Scientific Software Ecosystem (SSECO) developed in an E-Science joint project led by the Federal University of Juiz de Fora, Brazil in conjunction with EMBRAPA. It provides a platform that allows the accomplishment of experiment steps. A scientific SECO usually manages multiple relationships with external research, considering that the scientific context requires specialized and interdisciplinary knowledge, as in E-SECO's case.

An experimentation process usually traces the following steps on E-SECO. During the problem investigation step, scientists look for similar experiments, interact with other researchers using the E-SECO platform, define their goals and break down the experiment into smaller steps. In the experiment prototyping step, scientists build a prototype by designing workflows and reusing available assets, also accessing artifacts persisted in E-SECO-related repositories. Therefore, researchers can explore the assets and reuse their components to produce new products and provide new artifacts during the experiment prototyping step. As a final step, researchers analyze and publish their results and contributions, using the collaboration support provided by the E-SECO platform.

E-SECO Technical Specifications. E-SECO supports the integration with external applications through independent components developed for this purpose. Currently, the E-SECO platform connects with the following external platforms: Mendeley[4], Parsifal[5], MyExperiment[6], BioCatalogue[7] and Kepler[8]. E-SECO was conceived following Component-and-Connectors architectural style in conformance with MVC and SOA. For space restrictions, details about E-SECO architectural components instantiation will be discussed in future versions of this work.

Figure 1 presents the E-SECO platform main services. The E-SECO `Development Environment` service comprises the `Execution Environment`, which aims to support the execution of the experiment. The `Development Process` service supports the services built on the E-SECO platform, and the `Collaborative Services` service aims to support the collaborative activities conducted during experimentation processes. The `Development` Environment service interacts with additional support layers for the experimentation process.

Given the requirements imposed by the agriculture domain, we enhanced the E-SECO platform to fulfill them. The Provenance and Context layer were specified to support provenance management. This service is responsible for capturing, storing, performing inferences, and sharing information from scientific experiments. This

[4]https://www.mendeley.com/newsfeed
[5]https://parsif.al/
[6]https://www.myexperiment.org/home
[7]https://www.biocatalogue.org/
[8]https://kepler-project.org/

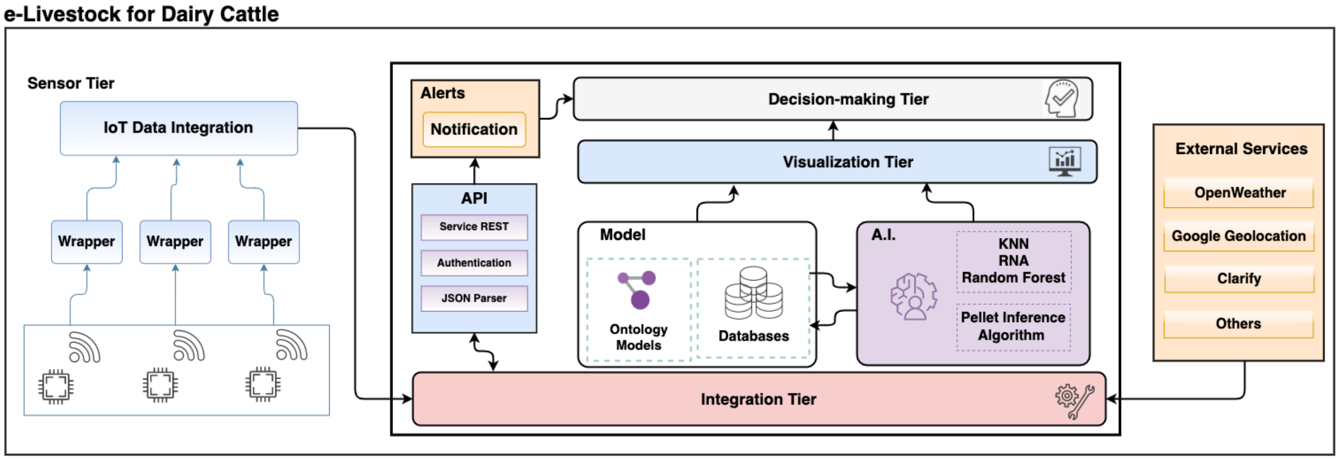

Figure 2: An overview on the e-Livestock Architecture for dairy cattle.

service is based on the PROV standard model [28] and its extensions. The Intelligence service was also specified, encompassing a Sensor, Model, and an Artificial Intelligence (AI) module. The **Sensor Module** collects IoT data generated by sensors used in the physical space, for example, devices connected to the animal's body or in the environment. This layer is designed to handle large volumes of data, ensuring availability and fault tolerance for the collected data. In addition, it also handles different data formats, commonly exported by sensors. The **Model module** addresses the integration and provenance models. By integrating external data, it is possible to derive new knowledge and store information about provenance. The model module also encompasses specific metadata such as model accuracy, average errors, algorithm type, and input dataset. We designed the **AI module** to provide new results based on data stored on models. These results can be predictions or classifications. It is possible to add new algorithms to this module according to the need and type of data collected. For example, if a new dataset from social media is added to the external services layer, we can add a natural language analysis algorithm to this module. The provenance of the data used by the model is captured during this communication between the A.I. module and the Model module. The next subsection presents details on e-Livestock Architecture, an abstract architecture created in conformance to E-SECO to enable the instantiation of systems and respective experiments to livestock subdomains.

3.1 e-Livestock Architecture for Dairy Cattle

To support dairy livestock experiments, the e-Livestock architecture for dairy cattle was designed on top of E-SECO platform, for the collection and processing of data generated by sensors attached to the animals, and to add additional environmental information related to provenance. e-Livestock for dairy cattle reuses specific services from E-SECO to support its construction, specifically, the Integration, Intelligence and Provenance management services. The architecture is structured into five main tiers, as shown in Figure 2. The Sensor Tier consists of cleaning, formatting, and transforming the data coming from sensors. After cleaning, the data is formatted and sent to be integrated with external data sources and subsequently persisted in a database. The Integration Tier manages and processes the data collected by external services, interacts with the Artificial Intelligence module, and persists data in the Model module. The Model module stores a set of metadata generated by the A.I. module so that it is possible to track the data provenance of the prediction results. The External Services tier comprises data sources external to the application. Finally, the Visualization Tier supports decision-making. The tiers and data flow in each step are described below.

Sensor Tier: That tier collects IoT data generated by sensors deployed in the farm. Sensors collect information such as temperature, humidity and neckless. In addition, the sensor tier also handles different data formats generated by internal systems on farms. These data are sent to the integration tier.

Integration Tier: That tier is responsible for processing the collected data to be integrated with information from other sources, services, and external APIs, such as context data, environmental information (temperature, humidity, and weather forecast). The main advantage of this tier is that it can centralize and aggregate external information to the researcher to enrich the decision-making process. For instance, it allows communication with other modules, such as the intelligence module, through a REST API. The A.I. module communicates with the integration tier, receiving the data already processed, then executing the most suitable intelligent algorithm for a given dataset. It sends the results, predictions, or classifications, to the API to be persisted.

External Services Tier: it represents external services, databases, historical bases, social networks, and any external data sources to add value to the data collected by the Sensor Tier. As needed, we can easily add new sources to the architecture through the Integration Tier. For example, by aggregating weather forecast data, it is possible to provide a new perspective for decision-making. This tier is also related to the data provenance. By recording data source, sensor, and data type, it is possible to track and analyze the context of decisions that used this information. Once the data is aggregated and stored, the retrieval of metadata stored in the database enables the decisions traceability.

Model Tier: This tier deals with farm data and its provenance. By integrating external data, it is possible to store the context of the data generated on the farm to extract the provenance and enable the generation of dashboards. By capturing the provenance of an animal's feed and comparing it with its milk production, we can identify the best diet to increase milk production, for example. The Model tier also stores the metadata of the models, such as model accuracy, average errors, algorithm type, and input dataset used. In this way, it is possible to extract the provenance of the results, which can be used to make future decisions. For example, with the prediction of food consumption, the researcher can estimate the expected cost of purchasing inputs and plan storage according to the probability of consumption of animals indicated by the algorithm.

Visualization Tier: The visualization tier allows the researcher to visualize the data in real-time through a panel according to a time interval. The researcher can also analyze and interpret data at different granularities. It allows users to visualize the results of the A.I. module.

3.1.1 Data Provenance. To achieve data provenance, the PROV standard model [27], used by the E-SECO platform, was extended to accurately capture the provenance of PLF-related data. Figure 3 presents a partial view of the model, which was specified as an ontological model [4] derived from the PROV-O model. Entities represent animals, agents are researchers and activities are actions performed on the smart farm. Activities can be described as insemination, milking or processing the collected data [18]. Considering the provenance of these data, it is possible to identify the data sources and the interactions that researchers and users carry out, tracking decisions related to these specific activities. As a result, it is possible to track decisions related to these specific activities.

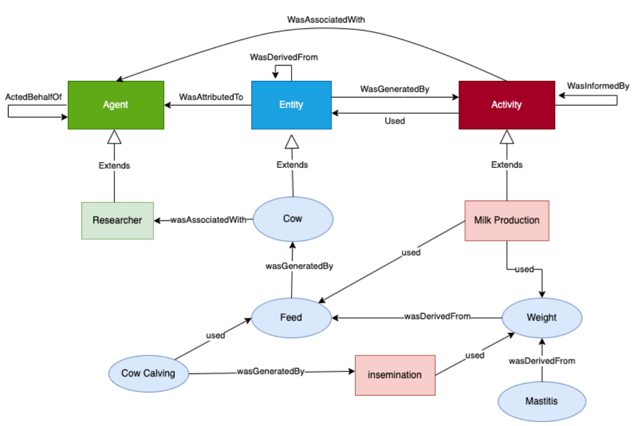

Figure 3: Provenance Model for e-Livestock.

3.1.2 Inference Mechanism. That model supports to answer questions such as: (i) Did the production of an animal that was eating correctly and had adequate weight decrease due to the occurrence of inflammation (mastitis)?; (ii) Was the animal's weight loss caused due to a change in diet that the animal did not adapt to?; (iii) Was an animal's weight loss due to a calving event?; (iv) Did the average production drop due to a temperature change?; (v) Did the temperature variation make the animal spend more energy maintaining body temperature instead of producing milk?; (vi) Did the average of mastitis cases grow, due to an increase in humidity, favoring the proliferation of environmental bacteria?; among other questions.

Considering that ontological model (Figure 3), we used object properties to implement the class relationships. We created Semantic Web Rule Language (SWRL) [20] rules to discover new associations between farm activities and animals. We specified the rules based on information provided by researchers from the Brazilian Agricultural Research Corporation – Dairy Cattle (EMBRAPA). As a result, using the declared model (explicit knowledge) with the addition of specific SWRL rules and inference mechanism (Pellet reasoner) [33], our provenance model infers the relationship of instances under the actions on the farm (implicit knowledge). These relationships are considered new knowledge produced from the processing of SWRL rules and inference engines about ontological instances (Figure 4). Rule S1 infers milked animals, S2 looks for animals with mastitis for which the milk has been discarded. Rule S3 looks for milk discards caused by mastitis. Rule S4 seeks milk production, and S5 infers milk production smaller than 12 (liters) that have mastitis.

Name	SWRL rules
S1	cow(?c) ^ milk(?m) ^ dairy_milk(?d) ^ wasMilked(?c, ?d) -> wasGeneratedBy(?m, ?d)
S2	cow(?c) ^ milk(?m) ^ wasMilked(?c, ?d) ^ wasProcessed(?m, ?p) ^ hadMastitis(?p, ?m) -> wasDiscartedBy(?m, ?p)
S3	process_milk(?p) ^ milk(?m) ^ cow(?c) ^ dairy_milk(?d) ^ wasProcessed(?m, ?p) ^ hadMastitis(?p, ?m) -> sqwrl:select(?m, ?p)
S4	dairy_milk(?d) ^ milk(?m) ^ wasGeneratedBy(?m, ?d) -> wasMilked(?m, ?d)
S5	milk(?m) ^ process_milk(?p) ^ swrlb:lessThan(?milk_value, 12) -> hadMastitis(?p, ?m)

Figure 4: SWRL Rules.

3.1.3 AI Module. The AI Module was developed to provide answers based on historical farm data. These answers can be milk production forecasts, mastitis type classification, and food consumption estimation. Based on historical data of the type of food, average consumption per animal, number of animals, and even time of year, it is possible to optimize the amount of adequate feed to be supplied to the animals. That action can avoid spending more than the necessary by buying commodities such as corn and soybeans, generating economic impacts on the farm. By preventing waste, we can contribute to a more sustainable farm.

From that prescribed component, we can add other algorithms according to the data collected. We can also use data from social media to get users' feedback on the quality of a particular farm-made product. Once the smart model is trained and ready to use, it requests input data from the API and sends the results generated by the model back to the API. The data provenance used by the model is captured during this communication between the A.I. module and the API. The next section shows two different systems derived in conformance to the E-livestock abstract architecture and the experiments conducted over it (by transition, in conformance to E-SECO).

4 A CONDUCTION OF SCIENTIFIC EXPERIMENTS DERIVED FROM E-SECO/E-LIVESTOCK ARCHITECTURE

In this section, we detail two different and independent scientific experiments derived from E-livestock architecture: one in Compost Barn domain, and another in the Aviary domain, as follows.

4.1 Compost Barn Experiment

To evaluate the e-LiveStock architecture, we collected and processed data from the Compost Barn production system [12], located at EMBRAPA - Gado de Leite experimental field using the e-LiveStock architecture derived from the E-SECO services. The Compost Barn at EMBRAPA is part of a research project with Brazilian and international institutions related to improving the dairy cattle production system. Cows are organized into lots, according to their production, and can change lots over time. Continuous monitoring allows for adjustments in the animals' living conditions, increasing production and life quality. We use a dataset obtained in the field. This dataset contains the records of weight (kg), milk production (L), and mastitis of a given animal, in addition to the averages of temperature (°C) and humidity (%) over the months in the Compost Barn. Figure 6 shows a partial view of this dataset, where the mastitis field indicates the positive (true) or negative (-) incidence of the disease in the animal. Mastitis is an inflammation of the mammary glands that reduces milk production and causes discomfort to the animal. In addition, it generates costs with medication and technical assistance for the treatment.

As data is collected, processed, and made available on the dashboards, it became possible to constantly monitor the Compost Barn environment and detect events such as: increased weight of animals; increase or decrease in milk production per batch; Compost Barn temperature and humidity variation. By storing the provenance of training and testing of intelligent models, information on average errors, model type, and input data, it was possible to analyze and discover the best predictive model for a given data set. Using smart models, such as an ontological model, researchers could make more accurate planning and producers could investigate animals that were not producing as expected and improve the animal's life quality, food, and health. The producer could then prevent diseases and ensure animal welfare.

Using the monitoring capabilities, we identified a reduction in the milk production of one of the lots and with the provenance data, it was possible to investigate the reason for discarding the milk. During two months of monitoring, there was a high humidity level in the Compost Barn (highlighted in red in Figure 6), which generated a proliferation of environmental bacteria. These bacteria caused mastitis in the animal, which consequently needed to be medicated and lost weight. Thus, the decision was made to discard the milk due to the medications. Another case was detecting the sudden increase in weight of one of the cows and its change of lot. A peak was observed in one of the lots by tracking the evolution of animal weights. With the provenance model, we identified that one of the cows in that batch had an insemination event, causing an increase in weight and, later, the decision to migrate a batch.

Furthermore, it was possible to estimate the milk production of this animal and monitor whether the projection would take place within the following months, using intelligent algorithms. The results of intelligent algorithms offered a future insight for production planning and animal management. For the context of this study, the neural network had a more accurate result, although, for other datasets, new tests need to be conducted to choose the best model. The results obtained cannot be generalized and additional studies will have to be conducted later.

As a result, the e-Livestock architecture provided decision support through data provenance with A.I. Results were displayed in graphs and analyzed by the researchers. Once a decision was made or an event considered unusual was detected, we could clarify the reason and the entire process that generated the decision through the relationships captured by the provenance model.

4.2 Continuous Monitoring and Automated Decision in Aviary Environments

A second software-intensive product was derived in conformance to the proposed E-livestock architecture and E-SECO ecosystem platform. The product consisted of a system for monitoring and automated decision in aviary environments [25]. Birds are sensitive to factors such as humidity and temperature, which pressures producers for a continuous monitoring of those variables in aviary environments. To avoid losses due to diseases or deaths caused by the unpredicted variations in those factors, a system product was conceived and deployed in conformance to the e-livestock proposed architecture, as can be seen in Figure 5.

We can observe that the system developed conforms to the e-livestock architecture, since it is also structured using the same tiers: the `Sensor Tier` involving sensors related to (i) luminosity, (ii) temperature, and (iii) air humidity linked to a LoRa module, which worked as the platform to which the sensors were connected, and the actuators, i.e., devices linked to control a lamp (to increase the temperature and luminosity when required) and a FAN (to refresh the environment, when needed). The LoRa platform with sensors and actuators was linked to a gateway so that all the data collected could be forwarded to a physically stored by a cloud/web based service in a database.

Once data was collected in `Sensor Tier`, it was transferred via the LoRa connection until a gateway, from which data could be processed by `Integration Tier` and `Decision-Making Tier`. The `Decision-Making Tier` is responsible for, according to the thresholds established (such as an inferior threshold of 30 degrees Celsius), performing automated decisions on the environment, such as turning the lamp or FAN on/off. The data collected by sensors

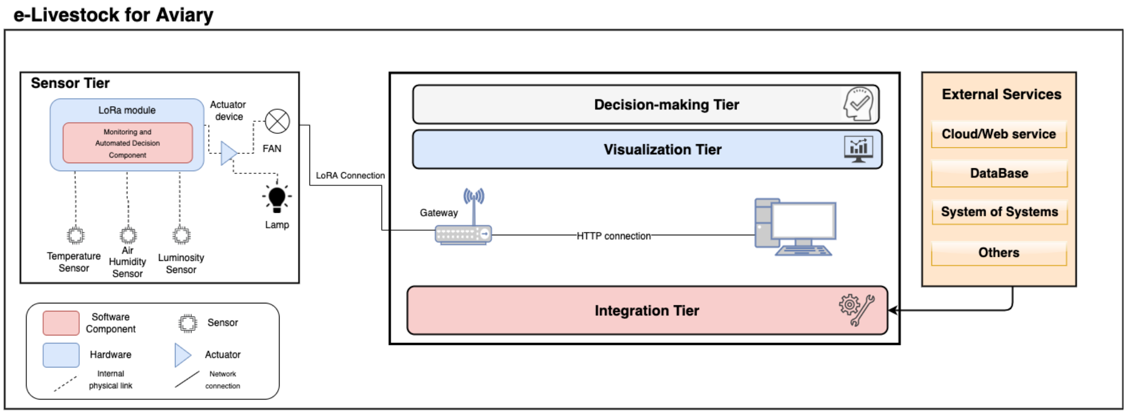

Figure 5: E-livestock architecture for Aviaries.

Month	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr
Weight	640	624	630	664	671	683	643	711
Milk	-	-	38.8	46	45.1	46	46.5	46
Mastitis	true	true	-	-	-	-	-	-
Umiditty	98	99	88	89	89	85	86	85
Temperature	26.1	28.1	24	22.4	22.6	21.7	22.7	24.8

Figure 6: Animal Dataset.

have three purposes: to (i) serve as input for automated decision *in-loco* in the `Decision-Making Tier`, (ii) be stored in a remote database and to be delivered for `External Services` for further processing and functionalities, (iii) be displayed to the farmers using the `Visualization Tier` and (iv) be analyzed and mined to search for optimization patterns, such as to pre-define triggers on temperature that exhibited better results over the system functioning eventually using the `A.I. Module` (not included in this prototype, but already predicted as a likely module).

The system was deployed in a real farm in Posse, Goiás (Brazil) and worked for three days, from November 10th, 2021 until November 13th, 2021. A new experiment started on December 6th, 2021 and runned for more time. During its operation, the system was well-succeeded to monitor the aforementioned variables for 24 hours during 13 days. The lamp was turned-on most of the time to maintain the temperature above 30 degrees Celsius. The system operated the lamp (turning it on or off) 14 times, and the FAN was not required to work since the maximum temperature set was 37 degrees Celsius and the local temperature was low due to recurrent rainfall in the farm region. 50 birds were monitored during that period. There were two losses. The farmers classified the final result as effective, since they did not need to work directly on it during the period to monitor the considered variables.

4.3 Related Work

Parrot, Lacroix and Wade [31] developed a multi-agent collaborative architecture to support dairy industry decisions. In their approach, the authors used ontology to map meanings from different domains and establish communication between agents. The authors did not propose integrating experiments, unlike our work using direct relationships to integrate data, and inference algorithms (Pellet) to discover new knowledge. These relationships and new knowledge help to improve the integration process, including new connections between information sources.

Janssen et al. [22] described the SEAMLESS (System for Environmental and Agricultural Modeling) architecture, which integrates databases from different domains, such as climatic conditions, soil, and cropping patterns. Compared to SEAMLESS, our work enables discovering new knowledge as well as integrating and sharing such knowledge among the different research centers.

Jonqueta et al. [23] present a platform called AgroPortal that receives and hosts ontologies, aligning them and enabling data reuse in agricultural software applications. However, these papers do not use these ontologies as integration models and data provenance to support reuse in the context of experiments.

Da Cruz et al. (2019) [8] discuss the use of data provenance in the agricultural context, considering problems such as the reproducibility of experiments. The article also presents "Rflow", a framework for managing workflows. Da Cruz et al. (2018) [7] present the OpenSoils program. OpenSoils architecture is driven by open provenance and aims to improve the reproducibility of experiments and deliver high-quality datasets, knowledge, and provenance-based maps.

In these studies, the support of a SECO platform is not explored, nor is it used to derive specific architectures to support specific agricultural experiments. Therefore, as far as we can ascertain, the proposals available in the literature do not discuss solutions to support the derivation of agricultural architectures in the context of an "ecosystem platform". Next section presents final remarks.

5 FINAL REMARKS

The main contribution of this paper is descriptive, i.e., our contribution was to show how E-SECO scientific ecosystem was used to derive an abstract architecture named e-livestock architecture, which was itself used as a reference to instantiate two different (and independent) systems and corresponding experiments in two livestock subdomains: compost barn and aviary. The two conducted experiments showed that E-SECO is able to support reuse and research advances on ecosystem platforms of scientific software, besides encapsulating primitives for the conduction of scientific experiments in order to support a forthcoming technology transfer from livestock research institutes to industry.

Our results are relevant since they address problems related to world food production. Our approach has potential to be replicated in food production research institutions over the world, besides being a contribution for scientific SECOs and livestock technological solutions.

Future work include supporting interoperability between E-SECO conforming systems, possibly integrating them into scientific workflow management systems, sensors, and external platforms. We also intend to invest on reinforcing and prioritizing quality attributes, such as data provenance, interoperability [15, 34] and reliability [16]. New rules can also be defined to support data enrichment in the ontology and their integration with other domain-specific ontologies to increase the capacity of knowledge extraction. Finally, we believe that it would be useful to conduct new experiments in other livestock subdomains to evaluate the support offered by e-livestock architecture in different application subdomains.

ACKNOWLEDGMENTS

We thank the researchers who participated in the evaluation of this proposal, as well as the Brazilian Agricultural Research Corporation (Embrapa/Brazil). This work was partially funded by UFJF/Brazil, CAPES/Brazil, CNPq/Brazil (grant: 311595/2019-7), and FAPEMIG/Brazil (grant: APQ-02685-17), (grant: APQ-02194-18).

REFERENCES

[1] Lenita Ambrósio, Heitor Linhares, José Maria N David, Regina Braga, Wagner Arbex, Mariana Magalhães Campos, and Rafael Capilla. 2021. Enhancing the Reuse of Scientific Experiments for Agricultural Software Ecosystems. *Journal of Grid Computing* 19, 4 (2021), 1–24.

[2] Christiane Bahlo, Peter Dahlhaus, Helen Thompson, and Mark Trotter. 2019. The role of interoperable data standards in precision livestock farming in extensive livestock systems: A review. *Computers and electronics in agriculture* 156 (2019), 459–466.

[3] Jan Bosch. 2009. From software product lines to software ecosystems.. In *SPLC*, Vol. 9. 111–119.

[4] Peter Buneman and Wang-Chiew Tan. 2019. Data provenance: What next? *ACM SIGMOD Record* 47, 3 (2019), 5–16.

[5] Bin Cao, Beth Plale, Girish Subramanian, Ed Robertson, and Yogesh Simmhan. 2009. Provenance information model of karma version 3. In *2009 Congress on Services-I*. IEEE, 348–351.

[6] Tadeu Classe, Regina Braga, José Maria N David, Fernanda Campos, and Wagner Arbex. 2017. A distributed infrastructure to support scientific experiments. *Journal of Grid Computing* 15, 4 (2017), 475–500.

[7] Sérgio Manuel Serra da Cruz, Marcos Bacis Ceddia, Renan Carvalho Tàvora Miranda, Gabriel Rizzo, Filipe Klinger, Renato Cerceau, Ricardo Mesquita, Ricardo Cerceau, Elton Carneiro Marinho, Eber Assis Schmitz, et al. 2018. Data Provenance in Agriculture. In *International Provenance and Annotation Workshop*. Springer, 257–261.

[8] Sérgio Manuel Serra da Cruz and José Antonio Pires do Nascimento. 2019. Towards integration of data-driven agronomic experiments with data provenance. *Computers and Electronics in Agriculture* 161 (2019), 14–28.

[9] Simone da Silva Amorim, Eduardo Santana de Almeida, and John D McGregor. 2013. Extensibility in ecosystem architectures: an initial study. In *Proceedings of the 2013 International Workshop on Ecosystem Architectures*. 11–15.

[10] Simone da Silva Amorim, Eduardo Santana de Almeida, and John D McGregor. 2014. Scalability of ecosystem architectures. In *ICSA*. IEEE, 49–52.

[11] Simone da Silva Amorim, John D McGregor, Eduardo Santana de Almeida, and Christina von Flach G. Chavez. 2014. Flexibility in ecosystem architectures. In *ECSA Workshops*. 1–6.

[12] Embrapa Gado de Leite. 2020. Brasil tem a primeira instalação de compost barn destinada a pesquisa. https://www.embrapa.br/busca-de-noticias/-/noticia/53360675/brasil-tem-a-primeira-instalacao-de-compost-barn-destinada-apesquisa.

[13] Daniel De Oliveira, Eduardo Ogasawara, Fernanda Baião, and Marta Mattoso. 2010. Scicumulus: A lightweight cloud middleware to explore many task computing paradigm in scientific workflows. In *2010 IEEE 3rd International Conference on Cloud Computing*. IEEE, 378–385.

[14] Rodrigo Pereira dos Santos, Cláudia Werner, Olavo Barbosa, and Carina Alves. 2012. Software Ecosystems: Trends and Impacts on Software Engineering. In *26th SBES*. IEEE, Natal, Brazil, 206–210. https://doi.org/10.1109/SBES.2012.24

[15] Juliana Fernandes, Valdemar Vicente Graciano Neto, and Rodrigo Pereira dos Santos. 2021. An Approach Based on Conceptual Modeling to Understand Factors that Influence Interoperability in Systems-of-Information Systems. In *XX SBQS*. ACM, 34:1–34:10. https://doi.org/10.1145/3493244.3493275

[16] Francisco Henrique Ferreira, Elisa Yumi Nakagawa, and Rodrigo Pereira dos Santos. 2021. Reliability in Software-intensive Systems: Challenges, Solutions, and Future Perspectives. In *47th SEAA*. IEEE, Palermo, Italy, 54–61. https://doi.org/10.1109/SEAA53835.2021.00016

[17] Yolanda Gil, Ewa Deelman, Mark Ellisman, Thomas Fahringer, Geoffrey Fox, Dennis Gannon, Carole Goble, Miron Livny, Luc Moreau, and Jim Myers. 2007. Examining the challenges of scientific workflows. *Computer* 40, 12 (2007), 24–32.

[18] Jonas S Gomes, José Maria N David, Regina Braga, Wagner Arbex, Bryan Barbosa, Wneiton Luiz Gomes, and Leonardo M Gravina Fonseca. 2021. e-LivestockProv: An Architecture based on Provenance to Manage Traceability in Precision Livestock Farming. In *Anais do I Workshop de Práticas de Ciência Aberta para Engenharia de Software*. SBC, 43–48.

[19] Valdemar Vicente Graciano Neto, Rodrigo Pereira dos Santos, Davi Viana, and Renata Araujo. 2020. Towards a Conceptual Model to Understand Software Ecosystems Emerging from Systems-of-Information Systems. In *Software Ecosystems, Sustainability and Human Values in the Social Web*, Rodrigo Pereira dos Santos, Cristiano Maciel, and José Viterbo (Eds.). Springer, 1–20.

[20] Ian Horrocks, Peter F Patel-Schneider, Harold Boley, Said Tabet, Benjamin Grosof, Mike Dean, et al. 2004. SWRL: A semantic web rule language combining OWL and RuleML. *W3C Member submission* 21, 79 (2004), 1–31.

[21] Slinger Jansen, Anthony Finkelstein, and Sjaak Brinkkemper. 2009. A sense of community: A research agenda for software ecosystems. In *2009 31st ICSE-Companion*. IEEE, 187–190.

[22] Sander Janssen, Erling Andersen, Ioannis N Athanasiadis, and Martin K van Ittersum. 2009. A database for integrated assessment of European agricultural systems. *Environmental Science & Policy* 12, 5 (2009), 573–587.

[23] Clement Jonquet, Anne Toulet, Elizabeth Arnaud, Sophie Aubin, Esther Dzale Yeumo, Vincent Emonet, John Graybeal, Marie-Angélique Laporte, Mark A Musen, Valeria Pesce, et al. 2018. AgroPortal: A vocabulary and ontology repository for agronomy. *Computers and Electronics in Agriculture* 144 (2018), 126–143.

[24] GS Karthick, M Sridhar, and PB Pankajavalli. 2020. Internet of things in animal healthcare (IoTAH): review of recent advancements in architecture, sensing technologies and real-time monitoring. *SN Computer Science* 1, 5 (2020), 1–16.

[25] Vinícius C Lopes, Roberto Felício de Oliveira, and Valdemar Vicente Graciano Neto. 2021. Towards an IoT-Based Architecture for Monitoring andAutomated Decision-Making in an Aviary Environment. In *Anais do XIII Congresso Brasileiro de Agroinformática*. SBC, 320–328.

[26] Konstantinos Manikas. 2016. Revisiting software ecosystems research: A longitudinal literature study. *Journal of Systems and Software* 117 (2016), 84–103.

[27] Paolo Missier, Khalid Belhajjame, and James Cheney. 2013. The W3C PROV family of specifications for modelling provenance metadata. In *16th EDBT*. 773–776.

[28] Luc Moreau and Paolo Missier. 2013. PROV-DM: The prov data model. W3C recommendation. *World Wide Web Consortium* (2013).

[29] Valdemar Vicente Graciano Neto, Fábio Basso, Rodrigo Pereira dos Santos, Noor Hasrina Bakar, Mohamad Kassab, Cláudia Werner, Toacy Cavalcante de Oliveira, and Elisa Yumi Nakagawa. 2019. Model-driven engineering ecosystems. In *Proceedings of the 7th SESoS and 13th WDES 2019, Montreal, QC, Canada, May 28, 2019*. IEEE / ACM, 58–61. https://doi.org/10.1109/SESoS/WDES.2019.00016

[30] Valdemar Vicente Graciano Neto, Rodrigo Pereira dos Santos, and Renata Mendes de Araujo. 2017. New Challenges in the Social Web: Towards Systems-of-Information Systems Ecosystems. In *VIII WAIHCWS*, Cristiano Maciel, José Viterbo, and Rodrigo Pereira dos Santos (Eds.), Vol. 2039. CEUR-WS.org, 1–12. http://ceur-ws.org/Vol-2039/paper01.pdf

[31] Lael Parrott, René Lacroix, and Kevin M Wade. 2003. Design considerations for the implementation of multi-agent systems in the dairy industry. *Computers and electronics in agriculture* 38, 2 (2003), 79–98.

[32] Yogesh L Simmhan, Beth Plale, Dennis Gannon, and Suresh Marru. 2006. Performance evaluation of the karma provenance framework for scientific workflows. In *International Provenance and Annotation Workshop*. Springer, 222–236.

[33] Evren Sirin, Bijan Parsia, Bernardo Cuenca Grau, Aditya Kalyanpur, and Yarden Katz. 2007. Pellet: A practical owl-dl reasoner. *Journal of Web Semantics* 5, 2 (2007), 51–53.

[34] Pedro Henrique Dias Valle, Lina Garcés, and Elisa Yumi Nakagawa. 2021. Architectural strategies for interoperability of software-intensive systems: practitioners' perspective. In *SAC '21*, Chih-Cheng Hung, Jiman Hong, Alessio Bechini, and Eunjee Song (Eds.). ACM, 1399–1408. https://doi.org/10.1145/3412841.3442015