World Organisation for Animal Health
Founded as OIE

TECHNICAL ITEM II

Advancement in the Veterinary Services through Digitalisation (Data management, Veterinary Information Systems, Big Data, Meta Language, Artificial Intelligence)

Prof Beatriz Martinez Lopez

18/09/2024

Technical Item II:
Advancement in the Veterinary Services through Digitalisation (Data management, Veterinary Information Systems, Big Data, Meta Language, Artificial Intelligence)

Prof Beatriz Martinez Lopez

Director of the Center for Animal Disease Modelling and Surveillance

Summary

The massive amount of animal health data daily produced and Big Data analytics can constitute a revolution in how we approach disease control and epidemiology. Advances in genomics, sensors and information technologies allow a more detailed and precise characterisation of animal health. In this report we discuss the value that digitalisation can bring to Veterinary Services and encourage to advance towards Precision Veterinary Epidemiology as a concept that takes advantage of the multilevel animal health related data to better understand disease dynamics in a population and design more cost-effective systems for surveillance, early detection and rapid control of livestock diseases. We explore the sources of information, how it works and the main challenges to translate this concept into practice.

We believe digitalisation not only can significantly contribute to the improvement of livestock efficiency and sustainability with smaller environmental footprint, but also to boost vertical and horizontal advances in the way we prevent and manage livestock diseases both locally and globally. However, in order to make this a reality, critical advancements and changes in the way we collect, standardise, integrate, share, and use data are required. There is also a need to build interdisciplinary teams of computer scientist, engineers and/or data scientists that work collaboratively with veterinarians and other domain experts and are focused on animal health issues. Certainly, there are some challenges, but there is also a tremendous opportunity by veterinary services to apply these available computational tools to significantly improve animal health.

Introduction

Data collection, manipulation, and analysis are essential to inform timely interventions and design surveillance and control programs as cost-effectively as possible. Animal health data collection has been traditionally limited to gathering and monitoring some specific indicators believed to be informative to assess specific problems. However, in the last decades, the generation, collection, and the capability of data digitalisation, storage and analysis have skyrocketed and have widen the possibilities. Thus, nowadays, large amounts of multi-scale data are daily produced and collected by producers (such as trade and production records, animal wearable sensors), veterinarians and diagnostic laboratories (animal health records), or management and environmental monitoring systems (temperature, relative humidity, wind, etc.), among others; and they can be utilized to improve animal health management. These vast amounts of data constitute the so-called "Big Data", a massive group of extremely large, complex, and diverse data that cannot be processed by traditional means[1,2].

Certainly, the use of this Big Data can revolutionise how we approach animal health management similarly to how the collection of on-farm reproductive parameters or mortality records aided to increase productivity and efficiency in livestock in the past. We can understand animal health problems in more precise and transversal ways and shift to more proactive and customised approaches. However, the (Big) data management and the application of Big Data analytics is not straightforward and has to face several hurdles, traditionally defined by the so-called three V's: 1) volume - reflecting the increasing size of these dataset, 2) velocity - the rapid and often real-time rate at which the data is updating, and 3) variety - the different types of data collected across sources and spatiotemporal scales[3]. Moreover, other two important V's can be added: veracity and value, to reflect the need for data integrity/reliability and actionable results from the analysis[4,5]. Furthermore, additional V's up to seven[6], ten[7] or even 42[8] have also been suggested to more extensively describe the characteristics and properties of Big Data. However, overcoming these difficulties is nowadays more feasible with the availability of computational resources that facilitate data capture, storing and processing large volumes of data, cloud computing, advancements in automation and machine learning and development of analytical tools and integrated platforms. Taken this into account, the interest in Big Data analytics has increased exponentially to better manage animal health. However, Veterinary Services in many parts of the world lack the capacity to collect data in a digital and standardised way and they may lack also the necessary computer infrastructure and specialised personnel to properly handle such data. This is definitely an important barrier, but it is something that could be relatively easily addressed. In fact, creating this infrastructure to collect and handle these data (e.g. smart phones, tablets, laptops, servers, etc.) and bringing few engineers, programmers, computer or data scientists to support Veterinary Services won't be very costly and could be an incredible cost-effective investment. We have a scarcity of veterinarians worldwide, but with help of these other professionals we could make better use of the "veterinarian time" as they may facilitate and accelerate the collection, access and visualisation to relevant data so veterinarians can fully focus on evaluating such information and make better animal health decisions.

In this report we aim to provide an overview of:

1) the evolution and **current status of digitalisation in animal health**

2) the **considerations regarding data**: data types, data sources, data standards, data quality, etc. needed to better prevent and control diseases in food-producing animals

3) what we believe are the five top "**Animal Health digitalisation challenges**" to accelerate the digital revolution in animal health.

1. Current status of digitalisation in animal health

The applications and challenges of digitalisation in animal health and related Big Data analytics have been recently covered by several reviews and opinion papers. Some are discussing the opportunities and challenges of the Internet of Animal Health Things (IoAHT)[9], the value of using next generation sequencing (NGS), which allows DNA/RNA sequencing and mutation detection in a short period of time, and associated Big Data for precision medicine in horses and other animal species[10], their value for animal welfare[11,12] or to improve decision making in animal health[13,5]. However, Big Data in animal health and veterinary epidemiology has been severely infra-utilised in the practice[14]. Data are still presented and used separately, mostly with poor to non-existent communication between the different segments involved, and there are big challenges in all the steps of the chain, from data collection to the data analysis and production of comprehensive results. Solving these issues would be extremely beneficial for disease prevention and control.

A number of "precision" disciplines have surged trying to take advantage of the possibilities of this digitalisation and the Big Data analytics. This is the case of Precision livestock farming (PLF), which involves the inter-connection of a wide range of real-time technologies to monitoring target or key livestock parameters and thus, to identify and report relevant events that farms need to consider in order to optimise different aspects of livestock production, and this has happened with relatively success[10,15].

Likewise, we can export the same principles for animal health and disease control taking advantage of tools such as machine learning (ML) and other artificial intelligence (AI) methodologies. This approach could contribute to early detect disease events and improve animal health management through a better understanding of the dynamics of an infection and the early identification and detection of patterns and factors related with disease by combining traditional methods with Big Data analytics. This could be extremely useful to aid Veterinary Services to better design and implement preventive and control programs. In fact, the concept of precision epidemiology or "precision public health" has been already explored in human medicine[16] defining the latter as "*providing the right intervention to the right population at the right time*"[17]. In animal health, precision veterinary epidemiology could be achieved by integrating the very diverse multi-level data already generated in livestock (i.e., animal diagnostics, management, biosecurity, production indices, trade, economics, etc.). Using these high-resolution data, we can get further insights of the epidemiological problem and support more timely selection of customised interventions to specific groups of animals, farms or productions systems. Therefore, Precision veterinary epidemiology comes not to substitute "classic" epidemiology, but to take advantage of how the digital technologies and new available analytical tools allow more proactive, timely and customised assessments of population health, disease risk and on-farm vulnerabilities. We believe that this approach will provide a lot of value for livestock producers, private veterinarians and Veterinary Services allowing for more accurate and timely information and decision guidance (Figure 1).

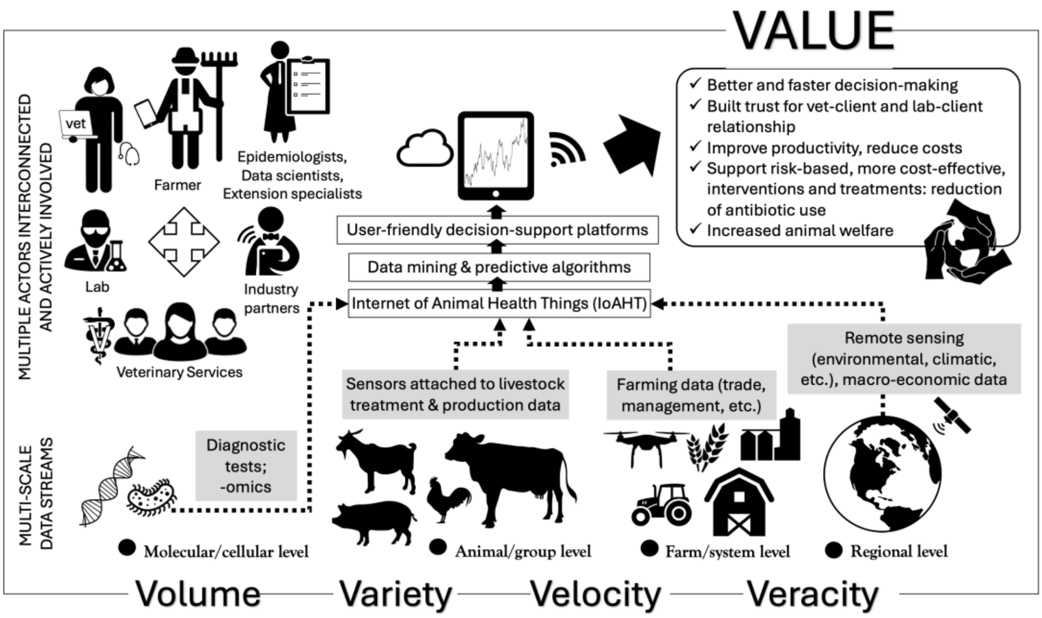

Figure 1. Scheme of how digitalisation can facilitate the integration of multiple data streams and the communication across stakeholders to generate value for animal health.

2. Data considerations

The wide variety of animal health related data is one of the biggest strengths (but also the main challenge) to fully achieve digitalisation by Veterinary Services. The collection of animal and farm information (i.e., digitalisation of agriculture) has substantially specialised and diversified over the years, particularly in large scale production systems of middle- and high-income countries. Today, it is easy to obtain information within herds, animals and at the molecular level, but also at larger scales, beyond herds and collectives, from a huge variety of resources that collect, store and share multiple data: automatic devices on/in-animals, on-farm sensors, records from the diverse monitoring programs carried out on the herd, laboratories, institutions, different companies among the livestock sector, mandatory reports, or programs, etc. This promises a high capacity to characterise and understand a system (their risks, weaknesses, strengths…) and thus achieving a high customisation of the interventions. However, these richness of data, data sources and data types are also the main challenge to operationally implement precision veterinary epidemiology. It is complex to integrate, map, visualise and provide timely access to stakeholders to all these data, as the data are proprietary and are collected, stored, and handled by diverse entities (i.e. production managements, diagnostic labs, veterinary clinics, pharma companies, etc.) and without following a common data structure, case definitions or data standards. Moreover, in low-middle income countries or areas with smallholders most of these data are still not collected or, if they are collected, they are collected in paper (e.g., epidemiological surveys and other questionnaires) which makes more complicated to integrate, process and use, which certainly delays the access of data and challenges real-time applications. The use of tablets or smart phones to collect information by Veterinary Services in the field can minimise data entry errors and accelerate the use of that data to support more timely decision making. However, this won't be enough. Research and policies should be directed to advance towards a more equitable and inclusive digitalisation process by facilitating the access of smallholders and Veterinary Services in low-resource settings to the digital opportunities. Below, we discuss the main data sources and types of animal health related data collected at molecular, animal, farm and regional level.

2.1. Information at molecular/cellular level

The molecular information constitutes an important pillar in the diagnosis, monitoring, and control of infectious diseases. PCR diagnosis and sequencing are routinely performed in diverse livestock segments to detect and identify specific or emerging strains circulating in a farm or to track new infections. The use of molecular information is well established in the epidemiological investigation to understand the source of an outbreak and the distribution at populational level, e.g., using phylogeography[18].

New molecular technologies such as next-generation sequencing (NGS) and gene expression analysis (i.e. process by which the information encoded in a gene is turned into end product, protein or function by the process of transcription and translation) are increasingly gaining presence due to the reduce costs and increased accessibility and facilitating the production and availability of massive genomic information for epidemiology. They have led to the emergence of different disciplines defined under the term of 'omics (e.g., proteogenomics, metabolomics, metagenomics, transcriptomics, etc.), high-throughput methodologies that can measure all the same molecules simultaneously providing a holistic view of a molecular process. The application of these technologies to the epidemiology supposes a revolution in how pathogens and hosts can be characterised. For example, we can identify specific traits related to pathogen resistance[19-21], key cell receptors in specific individuals[22,23] and the molecular pathways of the infection[24-26]; study persistence of infections, virulence, host–pathogen interactions, and pathogenicity, colonisation and survival mechanisms[27,28]; understand the genetic diversity of pathogens and interaction with microbial flora communities; evaluate the antibiotic resistance and associated resistant genes[29-30]. Omics also aid to identify biomarkers of infections or disease[31,32], which can be used to monitor collectives and early detect disease when it still goes unnoticed[33].

These new molecular technologies provide a more precise picture of the relationship between host and pathogen and may constitute a paradigm shift in the way we study biological problems. In human medicine, we have seen how these tools are already being implemented in the precision medicine to design customised treatments for the individuals considering the specific characteristics of hosts and pathogens[16]. The same principle can be extrapolated to veterinary epidemiology to customise control and eradication programs or to investigate the origin and the extension of an outbreak. Outbreaks of the same pathogen and strain may have different dynamics in different herds, and so the importance of disease control measures that considers the specific interactions between host, pathogen and environment in a given farm. Thus, the information that NGS can offer has tremendous potential for epidemiological surveillance, outbreak detection, and infection control. However, the main limitation to its implementation by Veterinary Services is the high cost and complex information management. It is still necessary to make NGS more affordable and develop systems able to more efficiently manage and analyse the huge volume of data that it produces.

2.2. Information at animal and farm level

Nowadays, sensors are used to collect data either from the environment in which the animals are placed, from the animal or from excretions of the animal (e.g., milk, saliva…). The development of sensor technologies has enormously increased in the last years and now we have the potential to measure almost every parameter on animals or farms that may be necessary for an epidemiological investigation, such as fever, animal movement, food and water intake[10,34,14,35].

These technologies have permeated at different speed in animal health depending on the sector. The dairy industry was one of the early advocates for the generation, collection and use of Big Data through devices and cyber-physical systems, such as robots or automatic milking systems, to better detect and manage reproductive problems, identify mastitis, monitor milk quality, or monitoring lameness[36,] and detect substances produced by animals: progesterone, volatile organic compounds (ketone bodies, ethanol, methanol, etc.)[14,41]. Poultry and swine industries have made also important advances in the implementation of sensor technologies, although adoption remains low mostly due to cost[11,42]. Today, we can monitor the environment (temperature, humidity and carbon dioxide and ventilation quality)[43]; or proxies of disease on real-time. Just to cite some examples, infrared thermal imaging can be used to monitor for thermal stress[44], accelerometers for activity[45], ear tags for behavior[46]. Mean body temperature and mobility readings have been used for early detection of avian influenza in broilers[47]. Microphones have also been used for monitoring coughs and detect respiratory problems in pigs[48]; photoelectric sensors to detect lameness[49,] camera systems for assessing body weight, and changes in behaviour and activity[50,51].

Although the use of sensors has exponentially expanded in PLF[42,52,53], the use of these data for animal health purposes still has to fully develop. In fact, most of these sensor technologies need to substantially expand their adoption in the livestock sector as currently they have been used as proof-of-concept or stand-alone isolated applications. In terms of veterinary epidemiology, the integration and synergic use of sensor data together with other data (i.e., health records, information about production performance, biosecurity, vaccinations, treatments, trade and other management practices on farm), would add more value and contribute to push animal health management forward.

2.3. Information at the regional level

Information in this level comprises a heterogeneous layer with data coming from very different sources. In the first place, we generate census or population data simply by collecting the information described above for each animal/farm and putting it together to characterise the general population (i.e., production system, association of farms, region, etc.). Census data aids to identify trends, assess the risk of contact with pathogens and take informed biosecurity measures and better plan for emergency response at regional level. Moreover, there are plenty of other resources that may also serve to characterise collectives. For example, the routine use of diagnostics leads to laboratories and veterinary clinics gathering a significant and increasing volume of diagnostic and disease data that can be used to evaluate population trends. However, animal and farm level data are proprietary and are not often accessible for out of the farm operators. Data sharing is scarce and complicated, because of privacy concerns and by the lack of systems properly designed to collect, integrate, and turn back information in a comprehensible manner. However, sharing of grouped or de-identified health records have been proven extremely valuable to inform producers about general trends and can be useful to increase awareness when there are changes in the patterns and epidemiology of certain diseases, particularly for those that are endemic[54]. We can also find records of notifiable disease outbreaks in national or international public repositories (e.g., WAHIS notifiable animal disease, https://wahis.woah.org).

Other types of valuable information are records of animal trade or truck movements, slaughter statistics, environmental and climatic factors, etc. These may have several applications. Network analysis of animal movements may help to measure the risk through animal trade[55,56]; spatial and temporal clustering of cases is useful to identify areas and time periods at highest risk of disease occurrence[57,58]; disease transmission models can help to simulate disease spread and identify the most cost-effective interventions[59]. Environmental, geographical, and climatological data (temperatures, precipitation, vegetation indices, environmental proxies, etc.) are known modulators of the risk of disease widely used in exploratory research studies, modelling, and disease mapping. Today, climatological data can be collected through meteorological stations, though ground-based observation stations may be limited in certain areas. Besides, visible and infrared satellite imagery has improved its spatio-temporal resolution and accessibility over the years and provides different measures such as radiation reflectance that can be transformed and used as environmental and climatic proxies to model risk of disease emergency[60]. The high-resolution and near-real time frequency of production of meteorological-related data allows to use them for monitoring, rapid assessment of at-risk areas or for prediction of disease distribution.

The new communication technologies also offer new resources to collect information. Internet data can be used to alert about the emergence of diseases before sanitary authorities may be aware[61,62]. In human medicine, different attempts have been made to use people interactions on the Internet to detect the emergence of epidemics. For example, Google Flu Trends was possibly one of the very first applications[63], and in the recent years monitoring social media has been used as an aid to surveillance, event detection, pharmacovigilance, forecasting, disease tracking and geographic identification of diseases[64,65]. Different strategies have been used for example to monitor tweets and use them to detect the emergence of a disease earlier than through official reports.

The possibilities of data from this level are huge due to the variety of data, but access and integration to all the sources is not always easy. Farm information is usually restricted to integrated production systems, farmer's associations, or collective quality improvement plans (e.g., milk control in dairy cattle); and it is done for very specific purposes, with different magnitude across livestock productions, rarely in real-time and sharing is poor. Population level data collected out of the farm also presents problems of integration and lack of standardisation. While some sources may offer large volumes of data with good periodicity, others may be rather incomplete and scarce. This heterogeneity complicates data mapping, manipulation and modelling in a timely fashion to extract value.

3. Animal Health digitalisation Challenges

The advance of digitalisation towards its broad use and implementation by Veterinary Services, needs to address what we believe are the top five Animal Health digitalisation challenges:

3.1. Facilitate the access and integration of multilevel data

The emergence, spread and expression of disease are modulated by multiple factors of very different nature. If we aim for a better understanding, and then, predictability of diseases, a holistic approach that integrates all the information potentially related with the epidemiology of diseases and puts all the actors involved together may better reflect the causal chain and their interactions. However, in practice the collection and combination of multilevel data sources is not frequent. Even within the same level, data are kept dispersed, usually collected by different companies or services, and presented in formats which are usually hard to integrate. Moreover, companies may gain ownership about the recording data and do not always facilitate the use of data out of their system. The lack of interoperability is one of the most cited challenges to the management of animal Big Data[66] and is detrimental not only to use the information, but also to explore and infer new applications. Instead, the potential outcomes generated from the interconnection of data should be understood as an additional and perhaps bigger value and this idea has to be transmitted to the final user. New services and systems may originate from this need and work as hubs that channel the diverse streams of information. Diverse initiatives have started in agriculture pursuing the standardisation and interoperability between data formats which may serve as example (OASIS, 2016; OGC, 2016; GODAN, 2013), but little work has been conducted in livestock up to date[66]. Even when combining similar data (e.g., diagnostic information), sometimes the problem is the lack of common case definitions or disparities in the data collection protocols. Despite attempts have been carried out at different levels: clinical signs[67], diagnostic data[68]; veterinary medicine standards are quite behind compared to the human medicine ones. Therefore, there is still a huge need to create animal health data collection standards or at least to provide clear case definitions, metadata and data dictionaries to be able to effectively combine, integrate and compare all these multiple sources of information and being able to extract more value from them.

Integrating information from different farms also arises privacy and security issues. Exploitation of data must not be perceived as a potential harm to the user, otherwise, farmers may be reluctant to share their information and the whole system will find its potential limited. This might have lower importance in endemic diseases, when all the farmers are willing to control or eradicate a disease. Indeed, benchmarking and elaboration of “whitelists” of free from disease farms may be a stimulus to join a program. However, if penalties or trade restrictions are applied, confidentiality may be a concern, which limit the ability to detect new outbreaks or emerging diseases. There are different ML methods that can help to overcome this issue. Papst et al.[69] have provided an overview of privacy-preserving data analytics. Among them, federated learning is a type of ML that has gained interest as a promising solution since it enables distributed data sharing and learning between devices without having to transfer raw data. Instead of sharing the farm data in a collective system, federate learning can collect and process the data locally and share the model, keeping the data secured. Though this technology has already permeated in diverse areas of the so-called “smart agriculture”, its use for veterinary applications is still in its beginning.

The problem of integration is especially important when we want to introduce omics data. We know that the interconnections between different ‘omics disciplines with data coming from other levels (e.g., feed data) are very relevant to understand the dynamics and risks of disease. However, management and integration of omics is particularly challenging due to the high volume and dimensionality of data, as well as the heterogeneity, even across omes, because of different scales or types of data (quantitative, qualitative, etc.). Different approaches have been proposed for multi-omic analysis: statistical analysis, network analysis, supervised analysis, etc.[70-72]. They are also limited by the frequent low sample size due to the cost of the omics procedures as most of the integration approaches require a large volume of data. Initiatives have been made to store and standardise animal metagenomic data[73] and there are several public repositories that contain reference pathways that can facilitate the reconstruction and prediction of the metabolic pathways[74]. However, further steps are still mandatory to advance towards a more functional integration. Therefore, we need a **multi-level data-driven solutions and decision frameworks** that can address some of those multi-scale and confidentiality challenges.

3.2. Bridge the gap between the data availability and its effective usage

Despite we already collect a lot of data and have the capabilities to collect even more, usage is often restricted to simple descriptive statistics or limited to specific aspects of animal production and pathogen diagnostics[75,76]. This means that a lot of the collected information is not used, and that we are not extracting the full potential of that data that was so expensive to collect. For example, accelerometers are conceived and primarily used to detect oestrous in dairy cattle, but it may also contribute to early identify sick animals because these are more reluctant to movement. Therefore, it is important to emphases the importance of a ***multipurpose data collection***. A holistic approach of the process of data collection and a wide variety of indicators will provide more accurate models and will increase our capability to understand and predict diseases.

However, simply collecting more data is not enough. Quantity does not imply quality, and having a lot of data does not necessarily mean that these data are representative and reliable. Thus, advances are also needed to filter which data are more valuable, informative, representative, etc. as well as which are the easiest to collect and economically viable for each purpose. The first concern affects to the quality of the collection process itself. A lot of links may intervene in the data collection process chain and, if the involved personnel or infrastructures are not ready to manage the data, loss and distortion of information may happen. Even when the right information is collected, data may present gaps depending on how they were collected. When missing information are due to a known factor (e.g., holidays, day of the week, not enough resolution, etc.), these lacks may be somehow mitigated through prediction, imputation and proxies[72,77]. However, when information is missed due to a non-random pattern or to systematic omissions, imputation is not adequate, and the situation is harder to compensate. Machine learning is especially sensitive to this issue as it relies on training or existing data as ground truth, so this can misrepresent the system and to obtain poor estimations and hinder the capacity of AI models to extract information.

In order to transform data into something informative, specific expertise, technology and analytics are required. Interdisciplinary approaches with professionals with appropriated skills and domain expertise in each link of the chain are the best way to proceed, but this also needs the requirement of the establishment of good channels of communication, sharing of information and feedback. The project “Pig Data” has provided an example on how implement Big Data by a transdisciplinary approach in livestock production[78]. Initiatives like Digital Innovation Hubs, where hubs (e.g. university) aim to support and assist parties delivering digital applications, may aid to create interdisciplinary groups as they are being used in precision farming[79]. A similar approach could be applied to improve digitalisation in animal health and facilitate the integration of precision veterinary epidemiology in the Veterinary Services.

3.3. Develop new algorithms specifically adapted to animal health

Usually, the better the quantity and quality of data inputted into a model, the better the model outputs will be (e.g. higher accuracy). Thus, early warning mitigation strategies are facilitated by incorporating multilevel data, but the multilevel structures also represent analytical challenges due to the complex structures and dimensionalities. Machine Learning algorithms are undoubtedly powerful tools and the best approach to exploit these inputs[80], as they offer advantages to manage complex, large datasets and extract value in a more semi-automatic way. When applied to a constant flow of data from the farm, these algorithms may be used to monitor the system and detect anomalies that can be indicative of a disease emergency before this can even be noted by farmers. The adoption of ML learning algorithms has rising exponentially in the last few years in different disciplines including health, but their application in animal health remains limited[81]. Guitian et al.[82] summarised different ML algorithms applied to animal health purposes mainly falling in four purposes: diagnosis, mortality and morbidity risk assessment, disease outbreak prediction and surveillance, and health policy and planning. They include very wide algorithms and applications: neural networks for image recognition, text mining to extract values from registries, classification trees to assess the decision making, prioritise samples, early detection of diseases, detect potential reservoirs, etc. The optimal ML algorithm often significantly depends on the goal and the composition of the dataset. An ideal precision veterinary epidemiology approach should integrate different algorithms and multiple analyses to better support animal health decisions. The problem is that the existing techniques are primarily developed for single level data, so ML algorithms need to be adapted to the multilevel reality in livestock. In many cases this is a real challenge since multi-level data does not hold the independence assumption among data of different levels. For example, in multi-omics, datasets present high dimensionality (large number of features for a relatively small number of samples), which makes inference and prediction particularly difficult unless large amounts of data are available, which is constrained by the cost of the analysis. Moreover, developing these algorithms to manage the massive flow of data and deliver results in real time ups the challenge.

On the other hand, it is crucial to evaluate the predictive quality of ML methods and compared with statistical approaches whenever possible[83]. Machine learning is especially focused on improving the predictive capability and tend to overfit and become rather useless if the characteristics used for training suddenly change in the real world. New research areas, such as Probabilistic Machine Learning, have emerged at the interface between the two domains[84] and can be explored in the future.

3.4. Develop operational systems to facilitate real-time data collection, integration, analysis, visualisation and sharing to support decision making

As the final goal is to provide solutions and assistance to Veterinary Services in the decision-making, the analysis needs to be integrated and simplified. We cannot expect that the final user (i.e., veterinarians, farmers, diagnosticians, etc.) has to deal with the methodologies, but we need to provide tools ready to monitor, visualize, evaluate and generate “interpretable” alerts of risks in real-time in changing scenarios, and to facilitate communication among different stakeholders. These tools are expected to be adapted to the needs and capabilities of the stakeholders, facilitate the interpretation of the results, early warn users of relevant anomalies, and propose a set of mitigation and control measures considering the sanitary and economic consequences. All these, hopefully with a low learning curve by end-users to effectively use these tools.

Different initiatives have been starting to integrate data collection sources and individual research methods into operational tools. In Europe, the DECIDE project (https://decideproject.eu) aims to develop data-driven decision-support tools for respiratory and gastro-intestinal syndromes and integrate them in existing farm management systems[85]. The project expects to develop decision-support tools to be integrated in existing farm management systems. Platforms such as the Disease BioPortal (https://bioportal.ucdavis.edu) at University of California, Davis, are already providing the means to integrate complex data structures, with real-time visualisation and analysis aiding to the decision-making. Producers using the Disease BioPortal are able to visualize and analyse different streams of data (i.e., test diagnostics, genomic information, animal movements, production records, antimicrobial resistance, etc.), and therefore link their molecular diagnostic results with animal and population level information into the platform. The information can then be converted to descriptive plots, geographic visualisation and phylogenetic trees in real-time, facilitating risk assessment as well as accelerating risk communication between labs, veterinarians and producers. Currently the Disease BioPortal program is adding ML to automatise some of the analyses and facilitate the process of dashboard generation, improving the access and experience of end users as well as incorporating new algorithms that facilitate data interpretation and highlights events such as increased incidence of disease, novel virus strains, drops in production, etc.

Despite the challenge is still substantial, the innovation and availability of methods is easier than ever. The new shareable platforms such as GitHub, which allow share code to collectively advance in the implementation of methods. There is also widely available application software that allow to experiment on the creation of interfaces and dashboards that can be useful for final users, such as Shiny, which offer the opportunity for interactive use of research-based and peer reviewed methodology across a global network of users. For example, shiny dashboards are now available for radiograph interpretation, omics visualisation, gene inference, network analysis, and more[86-90]. It is important to highlight that most of these Shiny dashboards are generated using open-source software, which contributes to the openness of the digitalisation process, although there is still a huge way to go to enhance digital transition of smallholders and Veterinary Services in low-resource settings[91].

3.5. Train next generation of data scientists as well as “animal health workers/veterinarian-data-scientists”

The next generation animal health workforce needs not only knowledge of veterinary health, but also of diverse areas that aid to process and analyse data which are simply too large to be used with conventional tools. Future veterinary epidemiologists will also need to be familiar with ML, parallel processing, information technology systems and incorporate technical skills, such as computer programming, which have not been a traditional part of their training. To address these issues there are mainly two lines of work: adapting existent curricula to integrate the necessary knowledge for these new demands and building transdisciplinary teams with different skills and backgrounds. **Innovative curriculum development** can be addressed by adapting already existing machine learning/data science curriculum but oriented to animal health. Therefore, formation needs to join members versed in the veterinarian needs, but also Computer Science experts who can provide solid, practical, and accessible instructions on machine learning techniques, toolboxes, and knowledge; and so, bridging the gap between the two areas. The participation of diverse faculty members with diverse areas of expertise allows to offer a more well-rounded formation to the students[92]. In this regard, different initiatives have been emerging in recent years to achieve a transdisciplinary data science, for example, by using co-teaching schemes (https://datascience.duke.edu/). Other approach has been implemented in Cornell named

Discipline-based Educational Researcher program, where an instructional designer is embedded in a department to help determine which active learning methods are most appropriate for that department's focus (https://cder.as.cornell.edu). Cooperation of students from different areas and disciplines around solving open-ended practical problems could also contribute to develop skills for the future practice. This approach enhances critical thinking skills, develops literature review ability, and encourages ongoing learning, often within a team environment. Similar programs are implemented in specialised graduate programs, but they could also be implemented in the DVM curriculum.

Building an interdisciplinary team that works collaboratively is essential to the success of digitalisation in the Veterinary Services and professionals should be prepared to work in this environment. Good principles and best practices have been identified in the literature, including good communications, regular meetings, appropriate skill mix and domain expertise, clear vision and well-defined goals, good leadership and management, good team culture, training and development opportunities, flexibility, etc.[93]

Besides technical skills, communication will also be an important asset for professionals in precision veterinary epidemiology. Part of the success depends on how end users use the product, so despite the complexity of the data collection, integration, mapping and analysis, the outputs need to be easy to access, understand and use. Human-computer interaction experts are key in this process to guide the design of user-friendly software and its use by end-users. From the operational point of view, co-design and involvement of end-users is critical to the success of the application as well as to communicate its value. Similarly, follow ups to receive feed-back or to assess and communicate the benefits from a practical, economical and sanitary point of view presenting measurable and palpable goals is also very important.

Conclusion

We believe digitalisation of Veterinary Services can substantially enhance animal health, food security and food safety if the digitalisation process becomes equable, accessible and ethical for all types and sizes of stakeholders. It may also contribute to substantially increase the sustainability, animal welfare and profitability in livestock operations. Digitalisation requires real-time integration to diverse multi-scale sources of data as well as the user-friendly and secure access and visualisation of such information by end-users (e.g. using dashboards) to timely inform decision-making. We believe that by addressing the five Animal Health digitalisation challenges, there will be a significant increase in animal health, animal welfare, farm productivity and revenue in the livestock industry. Although we are focusing in this report in the area of livestock health and food safety, as those are the main areas usually covered by Veterinary Services, we certainly believe that the digital transformation and the use of precision epidemiology, it is a holistic, One Health, approach that could be effectively used to better address population level problems in any public health and animal health setting (e.g. from the study of the impact of environmental exposures, infectious diseases or chronic diseases in humans to the evaluation of epidemiological issues affecting companion animals or wildlife conservation). Authorities should prioritise to develop policies, provide recommendations and invest in digitalisation (including data standardisation and interconnectivity) because not only it will help Veterinary Services to do better, faster and more effectively their job but also it will bring numerous social and economic benefits.

Bibliographic References

1. Diebold, F. X. (2012). *"On the Origin(s) and Development of the Term "Big Data."*

2. Favaretto, M., de Clercq, E., Schneble, C. O., & Elger, B. S. (2020). What is your definition of Big Data? Researchers' understanding of the phenomenon of the decade. *PLoS ONE*, *15*(2). https://doi.org/10.1371/journal.pone.0228987

3. Laney, D. (2001). 3-D data management: controlling data volume, velocity and variety. In *META Group Research Note, February 6*.

4. McCue, M. E., & McCoy, A. M. (2017). The scope of Big Data in one medicine: Unprecedented opportunities and challenges. *Frontiers in Veterinary Science*, *4*(NOV), 1–23. https://doi.org/10.3389/fvets.2017.00194

5. VanderWaal, K., Morrison, R. B., Neuhauser, C., Vilalta, C., & Perez, A. M. (2017). Translating big data into smart data for veterinary epidemiology. *Frontiers in Veterinary Science*, *4*(JUL), 1–7. https://doi.org/10.3389/fvets.2017.00110

6. Andreu-Perez, J., Poon, C. C. Y., Merrifield, R. D., Wong, S. T. C., & Yang, G. Z. (2015). Big Data for Health. *IEEE Journal of Biomedical and Health Informatics*, *19*(4), 1193–1208. https://doi.org/10.1109/JBHI.2015.2450362

7. Nigania, J. (2018). *The 10 V's of Big Data*. http://houseofbots.com/news-detail/2819-1-the-10-v%27s-of-big-data

8. Shafer, T. (2017). *The 42 V's of Big Data and Data Science*. https://www.elderresearch.com/blog/42-v-of-big-data

9. Smith, D., Lyle, S., Berry, A., & Manning, N. (2015). *Internet of Animal Health Things (IoAHT) Opportunities and Challenges*.

10. McCue, M. E., & McCoy, A. M. (2019). Harnessing big data for equine health. *Equine Veterinary Journal*, *51*(4), 429–432. https://doi.org/10.1111/evj.13080

11. Benjamin, M., & Yik, S. (2019). Precision livestock farming in swine welfare: A review for swine practitioners. *Animals*, *9*(4), 1–21. https://doi.org/10.3390/ani9040133

12. Sassi, N. Ben, Averós, X., & Estevez, I. (2016). Technology and poultry welfare. *Animals*, *6*(10), 1–21. https://doi.org/10.3390/ani6100062

13. Piñeiro, C., Morales, J., Rodríguez, M., Aparicio, M., Manzanilla, E. G., & Koketsu, Y. (2019). Big (pig) data and the internet of the swine things: A new paradigm in the industry. *Animal Frontiers*, *9*(2), 6–15. https://doi.org/10.1093/af/vfz002

14. Knight, C. H. (2020). Review: Sensor techniques in ruminants: More than fitness trackers. *Animal*, *14*(S1), S187–S195. https://doi.org/10.1017/S1751731119003276

15. Kleen, J. L., & Guatteo, R. (2023). Precision Livestock Farming: What Does It Contain and What Are the Perspectives? In *Animals* (Vol. 13, Issue 5). MDPI. https://doi.org/10.3390/ani13050779

16. Ladner, J. T., Grubaugh, N. D., Pybus, O. G., & Andersen, K. G. (2019). Precision epidemiology for infectious disease control. *Nature Medicine*, *25*(2), 206–211. https://doi.org/10.1038/s41591-019-0345-2

17. Khoury, M. J., Iademarco, M. F., Riley, W. T., Sciences, P., Service, P. H., Corps, C., Precision, T., & Initiative, M. (2016). Precision Public Health for the Era of Precision Medicine - ClinicalKey. *Am J Prev Med*, *50*(3), 398–401. https://doi.org/10.1016/j.amepre.2015.08.031.Precision

18. Van Borm, S., Belák, S., Freimanis, G., Fusaro, A., Granberg, F., Höper, D., King, D. P., Monne, I., Orton, R., & Rosseel, T. (2014). Next-generation sequencing in veterinary medicine: How can the massive amount of information arising from high-throughput technologies improve diagnosis, control, and management of infectious diseases? *Methods in Molecular Biology*, *1247*, 415–436. https://doi.org/10.1007/978-1-4939-2004-4_30

19. Bishop, S. C., & Woolliams, J. A. (2014). Genomics and disease resistance studies in livestock. *Livestock Science*, *166*(1), 190–198. https://doi.org/10.1016/j.livsci.2014.04.034

20. Islam, M. A., Rony, S. A., Rahman, M. B., Cinar, M. U., Villena, J., Uddin, M. J., & Kitazawa, H. (2020). Improvement of disease resistance in livestock: Application of immunogenomics and CRISPR/Cas9 technology. In *Animals* (Vol. 10, Issue 12, pp. 1–20). MDPI AG. https://doi.org/10.3390/ani10122236

21. Nietfeld, F., Höltig, D., Willems, H., Valentin-Weigand, P., Wurmser, C., Waldmann, K. H., Fries, R., & Reiner, G. (2020). Candidate genes and gene markers for the resistance to porcine pleuropneumonia. *Mammalian Genome*, *31*(1–2), 54–67. https://doi.org/10.1007/s00335-019-09825-0

22. Lagumdzic, E., Pernold, C., Viano, M., Olgiati, S., Schmitt, M. W., Mair, K. H., & Saalmüller, A. (2022). Transcriptome Profiling of Porcine Naïve, Intermediate and Terminally Differentiated CD8+ T Cells. *Frontiers in Immunology*, *13*. https://doi.org/10.3389/fimmu.2022.849922

23. Pluta, A., Taxis, T. M., van der Meer, F., Shrestha, S., Qualley, D., Coussens, P., Rola-Łuszczak, M., Ryło, A., Sakhawat, A., Mamanova, S., & Kuźmak, J. (2023). An immunoinformatics study reveals a new BoLA-DR-restricted CD4+ T cell epitopes on the Gag protein of bovine leukemia virus. *Scientific Reports*, *13*(1). https://doi.org/10.1038/s41598-023-48899-4

24. Bagnicka, E., Kawecka-Grochocka, E., Pawlina-Tyszko, K., Zalewska, M., Kapusta, A., Kościuczuk, E., Marczak, S., & Ząbek, T. (2021). MicroRNA expression profile in bovine mammary gland parenchyma infected by coagulase-positive or coagulase-negative staphylococci. *Veterinary Research*, *52*(1). https://doi.org/10.1186/s13567-021-00912-2

25. Islam, M. A., Große-Brinkhaus, C., Pröll, M. J., Uddin, M. J., Rony, S. A., Tesfaye, D., Tholen, E., Hölker, M., Schellander, K., & Neuhoff, C. (2016). Deciphering transcriptome profiles of peripheral blood mononuclear cells in response to PRRSV vaccination in pigs. *BMC Genomics*, *17*(1). https://doi.org/10.1186/s12864-016-2849-1

26. Kreitmaier, P., Katsoula, G., & Zeggini, E. (2023). Insights from multi-omics integration in complex disease primary tissues. In *Trends in Genetics* (Vol. 39, Issue 1, pp. 46–58). Elsevier Ltd. https://doi.org/10.1016/j.tig.2022.08.005

27. Belák, S., Karlsson, O. E., Leijon, M., & Granberg, F. (2013). December 2013 in issue 32-3 of the Scientific and Technical Review. (1) Department of Virology. In *Rev. sci. tech. Off. int. Epiz* (Vol. 32, Issue 3).

28. Hernández-Triana, L. M., Garza-Hernández, J. A., Ortega Morales, A. I., Prosser, S. W. J., Hebert, P. D. N., Nikolova, N. I., Barrero, E., de Luna-Santillana, E. de J., González-Alvarez, V. H., Mendez-López, R., Chan-Chable, R. J., Fooks, A. R., & Rodríguez-Pérez, M. A. (2021). An Integrated Molecular Approach to Untangling Host–Vector–Pathogen Interactions in Mosquitoes (Diptera: Culicidae) From Sylvan Communities in Mexico. *Frontiers in Veterinary Science*, *7*. https://doi.org/10.3389/fvets.2020.564791

29. Hickman, R. A., Leangapichart, T., Lunha, K., Jiwakanon, J., Angkititrakul, S., Magnusson, U., Sunde, M., & Järhult, J. D. (2021). Exploring the Antibiotic Resistance Burden in Livestock, Livestock Handlers and Their Non-Livestock Handling Contacts: A One Health Perspective. *Frontiers in Microbiology*, *12*. https://doi.org/10.3389/fmicb.2021.651461

30. Oniciuc, E. A., Likotrafiti, E., Alvarez-Molina, A., Prieto, M., Santos, J. A., & Alvarez-Ordóñez, A. (2018). The present and future of whole genome sequencing (WGS) and whole metagenome sequencing (WMS) for surveillance of antimicrobial resistant microorganisms and antimicrobial resistance genes across the food chain. In *Genes* (Vol. 9, Issue 5). MDPI AG. https://doi.org/10.3390/genes9050268

31. Gaudette, S., Smart, L., Woodward, A. P., Sharp, C. R., Hughes, D., Bailey, S. R., Dandrieux, J. R. S., Santos, L., & Boller, M. (2023). Biomarkers of endothelial activation and inflammation in dogs with organ dysfunction secondary to sepsis. *Frontiers in Veterinary Science*, *10*. https://doi.org/10.3389/fvets.2023.1127099

32. Perera, T. R. W., Skerrett-Byrne, D. A., Gibb, Z., Nixon, B., & Swegen, A. (2022). The Future of Biomarkers in Veterinary Medicine: Emerging Approaches and Associated Challenges. In *Animals* (Vol. 12, Issue 17). MDPI. https://doi.org/10.3390/ani12172194

33. Koene, M. G. J., Mulder, H. A., Stockhofe-Zurwieden, N., Kruijt, L., & Smits, M. A. (2012). Serum protein profiles as potential biomarkers for infectious disease status in pigs. *BMC Veterinary Research*, *8*. https://doi.org/10.1186/1746-6148-8-32

34. Caja, G., Castro-Costa, A., & Knight, C. H. (2016). Engineering to support wellbeing of dairy animals. *Journal of Dairy Research*, *83*(2), 136–147. https://doi.org/10.1017/S0022029916000261

35. Michie, C., Andonovic, I., Davison, C., Hamilton, A., Tachtatzis, C., Jonsson, N., Duthie, C. A., Bowen, J., & Gilroy, M. (2020). The Internet of Things enhancing animal welfare and farm operational efficiency. *Journal of Dairy Research*, *87*(S1), 20–27. https://doi.org/10.1017/S0022029920000680

36. Blewey, J. (2010). Precision dairy farming: advanced analysis solutions for future profitability. *Proceedings of the First North American Conference on Precision Dairy Management*.

37. Connolly, A. (2019). The new Digital World of Dairy Farming - Bridging the data gap. *Proceedings of the II International Precision Dairy Farming Conference*, 1–8.

38. Crowe, M. A., Hostens, M., & Opsomer, G. (2018). Reproductive management in dairy cows - The future. *Irish Veterinary Journal*, *71*(1), 1–13. https://doi.org/10.1186/s13620-017-0112-y

39. Hudson, C., Kaler, J., & Down, P. (2018). Use of big data in cattle practice. *In Practice*, *983*.

40. Thorup, V. M., Nielsen, B. L., Robert, P. E., Giger-Reverdin, S., Konka, J., Michie, C., & Friggens, N. C. (2016). Lameness affects cow feeding but not rumination behavior as characterized from sensor data. *Frontiers in Veterinary Science*, *3*(MAY), 1–11. https://doi.org/10.3389/fvets.2016.00037

41. Džermeikaitė, K., Bačėninaitė, D., & Antanaitis, R. (2023). Innovations in Cattle Farming: Application of Innovative Technologies and Sensors in the Diagnosis of Diseases. In *Animals* (Vol. 13, Issue 5). MDPI. https://doi.org/10.3390/ani13050780

42. Berckmans, D. (2017). General introduction to precision livestock farming. *Animal Frontiers*, *7*(1), 6–11. https://doi.org/10.2527/af.2017.0102

43. Bustamante, E., Guijarro, E., García-Diego, F. J., Balasch, S., Hospitaler, A., & Torres, A. G. (2012). Multisensor system for isotemporal measurements to assess indoor climatic conditions in poultry farms. *Sensors*, *12*(5), 5752–5774. https://doi.org/10.3390/s120505752

44. da Fonseca, F. N., Abe, J. M., de Alencar Nääs, I., da Silva Cordeiro, A. F., do Amaral, F. V., & Ungaro, H. C. (2020). Automatic prediction of stress in piglets (Sus Scrofa) using infrared skin temperature. *Computers and Electronics in Agriculture*, *168*. https://doi.org/10.1016/j.compag.2019.105148

45. Cornou, C., Lundbye-Christensen, S., & Kristensen, A. R. (2011). Modelling and monitoring sows' activity types in farrowing house using acceleration data. *Computers and Electronics in Agriculture*, *76*(2), 316–324. https://doi.org/10.1016/j.compag.2011.02.010

46. Pandey, S., Kalwa, U., Kong, T., Guo, B., Gauger, P. C., Peters, D. J., & Yoon, K. J. (2021). Behavioral monitoring tool for pig farmers: Ear tag sensors, machine intelligence, and technology adoption roadmap. *Animals*, *11*(9). https://doi.org/10.3390/ani11092665

47. Okada, H., Itoh, T., Suzuki, K., & Tsukamoto, K. (2009). Wireless sensor system for detection of avian influenza outbreak farms at an early stage. *Proceedings of the 2009 IEEE Sensors Conference*.

48. da Silva, J. P., de Alencar Nääs, I., Abe, J. M., & da Silva Cordeiro, A. F. (2019). Classification of piglet (Sus Scrofa) stress conditions using vocalization pattern and applying paraconsistent logic Eт. *Computers and Electronics in Agriculture*, *166*. https://doi.org/10.1016/j.compag.2019.105020

49. Besteiro, R., Arango, T., Rodríguez, M. R., Fernández, M. D., & Velo, R. (2018). Estimation of patterns in weaned piglets' activity using spectral analysis. *Biosystems Engineering*, *173*, 85–92. https://doi.org/10.1016/j.biosystemseng.2017.06.014

50. Nasirahmadi, A., Richter, U., Hensel, O., Edwards, S., & Sturm, B. (2015). Using machine vision for investigation of changes in pig group lying patterns. *Computers and Electronics in Agriculture*, *119*, 184–190. https://doi.org/10.1016/j.compag.2015.10.023

51. White, R. P., Schofield, C. P., Green, D. M., Parsons, D. J., & Whittemore, C. T. (2004). The effectiveness of a visual image analysis (VIA) system for monitoring the performance of growing/finishing pigs. *Animal Science*, *78*(3), 409–418. https://doi.org/DOI: 10.1017/S1357729800058811

52. Guarino, M., Norton, T., Berckmans, D., Vranken, E., & Berckmans, D. (2017). A blueprint for developing and applying precision livestock farming tools: A key output of the EU-PLF project. *Animal Frontiers*, *7*(1), 12–17. https://doi.org/10.2527/af.2017.0103

53. Wolfert, S., Ge, L., Verdouw, C., & Bogaardt, M. J. (2017). Big Data in Smart Farming – A review. In *Agricultural Systems* (Vol. 153, pp. 69–80). Elsevier Ltd. https://doi.org/10.1016/j.agsy.2017.01.023

54. Trevisan G, Linhares LCM, Crim B, Dubey P, Schwartz KJ, Burrough ER, Main RG, Sundberg P, Thurn M, Lages PTF, Corzo CA, Torrison J, Henningson J, Herrman E, Hanzlicek GA, Raghavan R, Marthaler D, Greseth J, Clement T, Christopher-Hennings J, Linhares DCL. (2019). Macroepidemiological aspects of porcine reproductive and respiratory syndrome virus detection by major United States veterinary diagnostic laboratories over time, age group, and specimen. PLoS One. 14(10):e0223544. doi: 10.1371/journal.pone.0223544.

55. Martínez-López, B., Perez, A. M., & Sánchez-Vizcaíno, J. M. (2009). Social network analysis. Review of general concepts and use in preventive veterinary medicine. *Transboundary and Emerging Diseases*, *56*(4), 109–120. https://doi.org/10.1111/j.1865-1682.2009.01073.x

56. O'Hara KC, Beltrán-Alcrudo D, Hovari M, Tabakovski B, Martínez-López B. (2022). Network analysis of live pig movements in North Macedonia: Pathways for disease spread. Front Vet Sci.; 9:922412. doi: 10.3389/fvets.2022.922412.

57. Gao, W., Jiang, X., Hu, Z., Wang, Q., Shi, Y., Tian, X., Qiao, M., Zhang, J., Li, Y., & Li, X. (2023). Epidemiological investigation, determination of related factors, and spatial-temporal cluster analysis of wild type pseudorabies virus seroprevalence in China during 2022. *Frontiers in Veterinary Science*, *10*. https://doi.org/10.3389/fvets.2023.1298434

58. O'Hara, K. C., Pires, A. F. A., & Martínez-López, B. (2021). Evaluating the association between climatic factors and sheep condemnations in the United States using cluster analysis and spatio-temporal modeling. *Preventive Veterinary Medicine*, *191*. https://doi.org/10.1016/j.prevetmed.2021.105342

59. Baron, J. N., Aznar, M. N., Monterubbianesi, M., & Martínez-López, B. (2020). Application of network analysis and cluster analysis for better prevention and control of swine diseases in Argentina. *PLoS ONE*, *15*(6). https://doi.org/10.1371/journal.pone.0234489

60. Mazzucato M, Marchetti G, Barbujani M, Mulatti P, Fornasiero D, Casarotto C, Scolamacchia F, Manca G, Ferrè N. (2023). An integrated system for the management of environmental data to support veterinary epidemiology. Front Vet Sci. 10:1069979. doi: 10.3389/fvets.2023.1069979.

61. Arsevska, E., Roche, M., Hendrikx, P., Chavernac, D., Falala, S., Lancelot, R., & Dufour, B. (2016). Identification of terms for detecting early signals of emerging infectious disease outbreaks on the web. *Computers and Electronics in Agriculture*, *123*, 104–115. https://doi.org/10.1016/j.compag.2016.02.010

62. Tizzani, M., Muñoz-Gómez, V., de Nardi, M., Paolotti, D., Muñoz, O., Ceschi, P., Viltrop, A., & Capua, I. (2021). Integrating digital and field surveillance as complementary efforts to manage epidemic diseases of livestock: African swine fever as a case study. *PLoS ONE*, *16*(12 December). https://doi.org/10.1371/journal.pone.0252972

63. Kandula, S., & Shaman, J. (2019). Reappraising the utility of Google Flu Trends. *PLoS Computational Biology*, *15*(8). https://doi.org/10.1371/journal.pcbi.1007258

64. Prieto, V. M., Matos, S., Álvarez, M., Cacheda, F., & Oliveira, J. L. (2014). Twitter: A good place to detect health conditions. *PLoS ONE*, *9*(1). https://doi.org/10.1371/journal.pone.0086191

65. Yousefinaghani, S., Dara, R., Poljak, Z., Bernardo, T. M., & Sharif, S. (2019). The Assessment of Twitter's Potential for Outbreak Detection: Avian Influenza Case Study. *Scientific Reports*, *9*(1). https://doi.org/10.1038/s41598-019-54388-4

66. Bahlo, C., Dahlhaus, P., Thompson, H., & Trotter, M. (2019). The role of interoperable data standards in precision livestock farming in extensive livestock systems: A review. *Computers and Electronics in Agriculture*, *156*(August 2018), 459–466. https://doi.org/10.1016/j.compag.2018.12.007

67. Willette, M., Ponder, J. B., McRuer, D. L., & Clark, E. E. (2013). Wildlife health monitoring systems in North America: from sentinel species to public policy. In A. Aguirre, R. S. Ostfel, & P. Daszak (Eds.), *Conservation Medicine: Applied Cases of Ecological Health* (pp. 552–562). Oxford University Press.

68. Main, R. (2016). Supporting VDL messaging and data analysis through data standardization. *Allen D. Leman Swine Conference*.

69. Papst, F., Saukh, O., Romer, K., Grandl, F., Jakovljevic, I., Steininger, F., Mayerhofer, M., Duda, J., & Egger-Danner, C. (2019, October 22). Embracing Opportunities of livestock big data integration with privacy constraints. *ACM International Conference Proceeding Series*. https://doi.org/10.1145/3365871.3365900

70. Chetty, A., & Blekhman, R. (2024). Multi-omic approaches for host-microbiome data integration. In *Gut Microbes* (Vol. 16, Issue 1). Taylor and Francis Ltd. https://doi.org/10.1080/19490976.2023.2297860

71. Fiocchi, C. (2023). Omics and Multi-Omics in IBD: No Integration, No Breakthroughs. In *International Journal of Molecular Sciences* (Vol. 24, Issue 19). Multidisciplinary Digital Publishing Institute (MDPI). https://doi.org/10.3390/ijms241914912

72. Picard, M., Scott-Boyer, M. P., Bodein, A., Périn, O., & Droit, A. (2021). Integration strategies of multi-omics data for machine learning analysis. In *Computational and Structural Biotechnology Journal* (Vol. 19, pp. 3735–3746). Elsevier B.V. https://doi.org/10.1016/j.csbj.2021.06.030

73. Hu, R., Yao, R., Li, L., Xu, Y., Lei, B., Tang, G., Liang, H., Lei, Y., Li, C., Li, X., Liu, K., Wang, L., Zhang, Y., Wang, Y., Cui, Y., Dai, J., Ni, W., Zhou, P., Yu, B., & Hu, S. (2022). A database of animal metagenomes. *Scientific Data*, *9*(1). https://doi.org/10.1038/s41597-022-01444-w

74. Shah, H. A., Liu, J., Yang, Z., & Feng, J. (2021). Review of Machine Learning Methods for the Prediction and Reconstruction of Metabolic Pathways. In *Frontiers in Molecular Biosciences* (Vol. 8). Frontiers Media S.A. https://doi.org/10.3389/fmolb.2021.634141

75. Grzesiak W and Zaborski D (2012). Examples of the Use of Data Mining Methods in Animal Breeding. Data Mining Applications in Engineering and Medicine. InTech. Available at: http://dx.doi.org/10.5772/50893.

76. Williams ML, Mac Parthaláin N, Brewer P, James WPJ, Rose MT. (2016). A novel behavioral model of the pasture-based dairy cow from GPS data using data mining and machine learning techniques. J Dairy Sci. 99(3):2063-2075. doi: 10.3168/jds.2015-10254.

77. Vahabi, N., & Michailidis, G. (2022). Unsupervised Multi-Omics Data Integration Methods: A Comprehensive Review. In *Frontiers in Genetics* (Vol. 13). Frontiers Media S.A. https://doi.org/10.3389/fgene.2022.854752

78. Faverjon, C., Bernstein, A., Grütter, R., Nathues, C., Nathues, H., Sarasua, C., Sterchi, M., Vargas, M. E., & Berezowski, J. (2019). A transdisciplinary approach supporting the implementation of a big data project in livestock production: An example from the swiss pig production industry. *Frontiers in Veterinary Science*, *6*(JUL), 1–11. https://doi.org/10.3389/fvets.2019.00215

79. van der Weerdt, C. A., Vonder, M. R., & Berkers, F. (2019). Accelarating adoption of precision dairy farming through digital innovation hubs. *Proceedings of the Second International Precision Dairy Farming Conference*, 8–11.

80. Herland, M., Khoshgoftaar, T. M., & Wald, R. (2014). A review of data mining using big data in health informatics. *Journal of Big Data*, *1*, 2. https://doi.org/10.1021/ci0255782

81. Ezanno, P., Picault, S., Beaunée, G., Bailly, X., Muñoz, F., Duboz, R., Monod, H., & Guégan, J. F. (2021). Research perspectives on animal health in the era of artificial intelligence. In *Veterinary Research* (Vol. 52, Issue 1). BioMed Central Ltd. https://doi.org/10.1186/s13567-021-00902-4

82. Guitian, J., Arnold, M., Chang, Y., & Snary, E. L. (2023). Applications of machine learning in animal and veterinary public health surveillance. *OIE Revue Scientifique et Technique*, *42*, 230–241. https://doi.org/10.20506/RST.42.3366

83. Hepworth, P. J., Nefedov, A. V., Muchnik, I. B., & Morgan, K. L. (2012). Broiler chickens can benefit from machine learning: Support vector machine analysis of observational epidemiological data. *Journal of the Royal Society Interface*, *9*(73), 1934–1942. https://doi.org/10.1098/rsif.2011.0852

84. Ghahramani, Z. (2019). *Probabilistic modelling, machine learning, and the information revolution*. Citeseer.

85. van Schaik, G., Hostens, M., Faverjon, C., Jensen, D. B., Kristensen, A. R., Ezanno, P., Frössling, J., Dórea, F., Jensen, B.-B., Carmo, L. P., Steeneveld, W., Rushton, J., Gilbert, W., Bearth, A., Siegrist, M., Kaler, J., Ripperger, J., Siehler, J., de Wit, S., … Nielen, M. (2023). The DECIDE project: from surveillance data to decision-support for farmers and veterinarians. *Open Research Europe*, *3*, 82. https://doi.org/10.12688/open-reseurope.15988.1

86. Badgeley, M. A., Liu, M., Glicksberg, B. S., Shervey, M., Zech, J., Shameer, K., Lehar, J., Oermann, E. K., McConnell, M. V., Snyder, T. M., & Dudley, J. T. (2019). CANDI: An R package and Shiny app for annotating radiographs and evaluating computer-aided diagnosis. *Bioinformatics*, *35*(9), 1610–1612. https://doi.org/10.1093/bioinformatics/bty855

87. Class, C. A., Ha, M. J., Baladandayuthapani, V., & Do, K. A. (2018). IDINGO-integrative differential network analysis in genomics with Shiny application. *Bioinformatics*, *34*(7), 1243–1245. https://doi.org/10.1093/bioinformatics/btx750

88. Fawcett, L. (2018). Using Interactive Shiny Applications to Facilitate Research-Informed Learning and Teaching. *Journal of Statistics Education*, *26*(1), 2–16. https://doi.org/10.1080/10691898.2018.1436999

89. Schultheis, H., Kuenne, C., Preussner, J., Wiegandt, R., Fust, A., Bentsen, M., & Looso, M. (2019). WIlsON: Web-based Interactive Omics VisualizatioN. *Bioinformatics*, *35*(6), 1055–1057. https://doi.org/10.1093/bioinformatics/bty711

90. Vinogradova, S., Saksena, S. D., Ward, H. N., Vigneau, S., & Gimelbrant, A. A. (2019). MaGIC: A machine learning tool set and web application for monoallelic gene inference from chromatin. *BMC Bioinformatics*, *20*(1), 1–5. https://doi.org/10.1186/s12859-019-2679-7

91. Lioutas E.D., Charatsari C., De Rosa, M., (2021). Digitalization of agriculture: A way to solve the food problem or a trolley dilemma?. *Technology in Society*, 67, 101744. https://doi.org/10.1016/j.techsoc.2021.101744

92. National Academies of Sciences Engineering. (2017). Data Science Education in the Future. In *Envisioning the data science discipline: the undergraduate perspective: interim report ahead* (pp. 19–25). The National Academic Press.

93. Nancarrow, S. A., Booth, A., Ariss, S., Smith, T., Enderby, P., & Roots, A. (2013). Ten principles of good interdisciplinary teamwork. *Human Resources for Health*, *11*(1), 1. https://doi.org/10.1186/1478-4491-11-19