Informatics progress of the Global Burden of Animal Diseases towards One Health data

K. Raymond (1, 2), N. BenSassi (1, 3), G.T. Patterson (1, 3), B. Huntington (1, 4, 5), J. Rushton (1, 4), D.A. Stacey (1, 2) & T.M. Bernardo* (1, 3)

(1) Global Burden of Animal Diseases programme, World Organisation for Animal Health, 12 rue de Prony, 75017 Paris, France and University of Liverpool, 146 Brownlow Hill, Liverpool, L3 5RF, United Kingdom

(2) School of Computer Science, University of Guelph, Reynolds Building, 474 Gordon Street, Guelph, ON N1G 2W1, Canada

(3) Department of Population Medicine, Ontario Veterinary College, University of Guelph, 50 Stone Road East, Guelph, ON N1G 2W1, Canada

(4) Department of Livestock and One Health, Institute of Infection, Veterinary and Ecological Sciences, University of Liverpool, 146 Brownlow Hill, Liverpool, L3 5RF, United Kingdom

(5) Pengwern Animal Health Ltd, 259 Wallasey Village, Wallasey, CH45 3LR, United Kingdom

*Corresponding author: bernardt@uoguelph.ca

Summary

The Global Burden of Animal Diseases (GBADs) will provide data-driven evidence on which policymakers can evaluate options, base decisions, and measure the success of animal health and welfare interventions. GBADs' Informatics theme is developing a transparent and efficient end-to-end process for identifying, analysing, visualising and sharing data to calculate livestock disease burdens, and derive models and dashboards. Together with other global burdens (human health, crop loss, foodborne diseases) they can constitute One Health

data, required to address cross-cutting issues such as antimicrobial resistance and climate change.

Starting with open data available from international organisations (who are undergoing their own digital transformation), efforts to achieve an accurate estimate of livestock numbers served as a pilot, revealing difficulties in finding, accessing and reconciling data from different sources over time. Ontologies and graph databases are being developed to bridge data silos and improve the findability and interoperability of data, which can be visualised by dashboards. Data stories, a documentation website and a data governance handbook explain GBADs data, which is now available through an application programming interface. Sharing data quality assessments builds trust in the data, encouraging its application for livestock and One Health. Animal welfare data presents a particular challenge, as much of it is held privately and deliberations are ongoing as to which data are most relevant. Accurate livestock numbers are an essential input to calculating biomass, which subsequently feeds into calculations of antimicrobial use and climate change determinations. GBADs data also factors into at least eight sustainable development goals.

Keywords

Dashboards – Data access – Data quality – Data strategy – Data visualisations – Global Burden of Animal Diseases – Interoperability – Livestock – One Health data.

Introduction

Policymakers need a sound foundation to align resources with the needs of the populations they serve, taking past and future trends into account. Just as the Global Burden of Disease (GBD) has provided metrics to guide investments in human health, the Global Burden of Animal Diseases (GBADs) aims to provide data-driven evidence about animal health and welfare trends and costs on which to evaluate options, base decisions and measure the success of interventions [1]. Together with other global burdens (crop loss [2], foodborne diseases [3]) they can

constitute One Health data, required to address cross-cutting issues such as antimicrobial resistance (AMR) and climate change.

The COVID-19 pandemic accelerated and highlighted the importance of the digital transformation of the health sector and exemplified the enhanced collaboration that can ensue with digital resources [4]. International organisations, including the World Organisation for Animal Health (WOAH, founded as OIE), the World Health Organization and the Food and Agriculture Organization of the United Nations (FAO) are bolstering their data management programmes, automating access to their data by developing application programming interfaces (APIs), and providing dashboards for visualisation. Since 1990, the GBD has grown to include over 7,000 collaborating researchers in over 156 countries [5]. GBADs Informatics theme was established at the University of Guelph with multidisciplinary expertise (epidemiology, computer and data science, animal welfare). It interacts with other GBADs themes and external collaborators to ensure responsible data management and a proactively designed, responsive information architecture to manage and visualise data, and which will accommodate innovation and growth.

The objective of this paper is to describe the progress of the Informatics theme, which is developing data resources in synchrony with, and dependent on, advances by other GBADs themes [1], including modellers and economists. At this point in time, Informatics is only using existing data from official sources, curating available datasets and identifying variables of interest to the modellers, who will validate and publish their results separately. The greater challenge is finding the needed data, rather than choosing among an abundance of sources. By making data visible and accessible to users, they can help to identify errors and provide input on further needs for data resources, which collectively form the GBADs Knowledge Engine.

The initial objective of the Informatics theme is to make existing data easy to find, visualise and access for calculation of animal health burdens and related metrics (Table I). By sharing our assessment of, and measures taken to improve data quality, we hope to build trust in

the data, thereby encouraging development and dissemination of burden calculations and models within the animal health community, as well as their uptake for wider issues in which animal health plays a part.

Calculation of GBADs will depend on having accurate estimates for various counts and costs of livestock production and diseases. One might expect that a basic parameter such as livestock population numbers that were obtained from official sources would be reasonably easy to interpret and re-use. It has proven, however, to be difficult to determine, as the numbers and categorisation of animals are not consistent across different organisations, or over time. For example, sheep and goats are sometimes combined and cattle may or may not include buffaloes or be divided into beef and dairy (Table II). Accurate livestock numbers are an essential input to calculating measures of productivity, as well as biomass, which subsequently feeds into calculations of antimicrobial use (AMU) and climate change determinations.

Data is on a spectrum from closed to shared to open [11]. Starting with open data available from international organisations such as WOAH, FAO and the World Bank, as well as data from national governments, we are embarking on 'the road to reproducibility', whereby the data is transparent, interoperable, annotated, audited for quality and meets standards for privacy. The utility of this data is demonstrated by developing data products such as dashboards; minimum viable products developed in an agile manner and made publicly available to garner feedback which then drives further development [12]. Pilot and case studies developed to date are elucidated below.

Global Burden of Animal Diseases data strategy

Recognising the significant efforts by FAO, WOAH, and other institutions in gathering global data on livestock over the decades and the call for an integrated international data system [13], GBADs is working to produce curated data useful for livestock modelling. Instead of each user having to find, clean, and interpret data on their own, GBADs strives to provide a curated source of data available for all to visualise, analyse, or use in the development of models and other data

products. In order to build trust and confidence in the data, it is necessary to document where data comes from and how it has been manipulated. GBADs aims to create a community around shared, curated, and re-usable data on a single platform. The development of all data products (Table I), including metadata and provenance, is guided by four foundational principles: Findability, Accessibility, Interoperability, and Reusability, more commonly known as FAIR principles [14, 15].

Livestock population data collection and analysis

The GBADs data strategy began with the collection and analysis of livestock population numbers by country, species, and production system. Achieving accuracy of this data was both difficult and essential, as it serves as an input for subsequent calculations, such as biomass.

Starting with data from WOAH, Food and Agriculture Organization of the United Nations Corporate Statistical Database (FAOSTAT), and statistical office of the European Union (Eurostat) for proof of concept, data were downloaded via the web portal or provided by the data source (in the case of WOAH data). While FAOSTAT and Eurostat data were available via APIs, use of APIs was difficult due to insufficient technical documentation on how to use APIs to gather data, and internal software updates resulting in deprecated or unreliable APIs. As Ethiopia is the national pilot study for GBADs, data from the Ethiopian Central Statistical Agency (EthCSA) were obtained from publicly available portable document format, more commonly known as pdf reports. Web scraping programmes (Table I) were coded to convert pdf tables to a digitally accessible format (comma-separated values files, more commonly known as csv files). Collected data was then stored in cloud-hosted database tables. An API was designed to pull data from the database, allowing data to be accessed by both humans and machines.

Process flow for ingestion of model outputs

Our collaborating modellers create and validate models (e.g. biomass) using data from the GBADs API as inputs. Once the output datasets are

provided to Informatics, datasets are ingested into database tables, and made available through the API and dashboards. These datasets can then be used as inputs for subsequent models. All models were coded and made available in GitHub repositories (Table I). Information about data inputs, data provenance, methodology, and source code is documented in metadata.

Exploratory analysis of data values and species categorisations

Data quality was explored by identifying outliers in population numbers. Outliers were defined as numerical values that were much larger or smaller than possible given the temporal trends of species populations and were identified through examination of data visualisations (Figure 1). Agreeability between data sources that report populations for the same species and country was assessed by determining whether reported numbers were the same. Data reports (Table I) with visualisations of data over time, and across different countries and species were used to identify outliers and compare the data. When erroneous data values were identified, they were reported to the host organisation and in some cases they directed us to change the values and these changes were documented. Furthermore, data values using differing units (1,000 heads/number of heads, kilogram/gram [kg/g]) were harmonised to ensure consistency across curated datasets.

Exploring interoperability involved a comparative analysis of how species are categorised between and within data sources, and over time. The results of the analysis (Table II and Table III) revealed different categorisations of species between data sources, and within data sources over time. When available, controlled vocabularies and metadata were referenced to understand the semantics (i.e. contextual meaning) of a term. However, metadata were often non-existent or did not adhere to vocabularies leading to ambiguity in the understanding of what the data categories represent.

Interoperability, ontologies, and graph databases

Achieving semantic interoperability (i.e. 'the unambiguous access and interpretation of data by different stakeholders' [17]) is a challenge, as described in the previous subsection. For data to be re-used and combined, data must be reported using consistent categorisations and be interpreted in the way that was intended. Ontologies are being developed to map similar concepts in data categories based on the observed behaviour of how data are reported.

Graph databases (GraphDBs) are being developed for storage and management of data resources that change over time, between and within different sources, and across geographic regions. By using the GraphDB to connect interrelated data, similar terms can be identified and connected thereby bridging data silos and improving the findability and interoperability of data.

Data provenance

Data provenance provides information about the origin of the data, where it has flowed, changes that have been made to the data, and how it has been used. Detailed data provenance and lineage is important in two scenarios:

1) There may be different versions, updates, or changes to datasets that are used in GBADs estimates.

2) Processed datasets are obtained from GBADs data portals and APIs where they are used in calculations that produce more datasets.

To ensure that estimates and data flows are reproducible, reusable, and transparent, data provenance and lineage is documented in a GraphDB.

Data governance and private data

GBADs data flows, best practices for data use and sharing, data sharing principles, metadata standards, and data governance documentation are

communicated through an online Data Governance Handbook (Table I).

Our current work has focused on Open Data (e.g. data that do not contain personal identifiable or sensitive information); any available licensing information about the data is recorded in the metadata. In anticipation of controlled governance of private data (such as animal welfare data), we have developed secure login systems for the GBADs data portal to support secure use and dissemination of sensitive data. Data agreements and licences with private data holders will govern how sensitive data can be used, by whom, and for which purposes.

Accomplishments and future plans

Once the basic data accessibility issues have been dealt with, tools can be built to display, visualise, and make data available to users. This is being done through the development and deployment of dashboards, data quality reports, data stories, an extensive documentation website, and a GBADs API.

Dashboards

It is essential to provide users with an easy-to-use site that will help them understand the available data and to access it and its metadata. A dashboard is 'a visual display of data used to monitor conditions and/or facilitate understanding' [18]. GBADs Informatics is continually developing and deploying dashboards for animal populations (Figure 1), biomass calculations, total economic value of livestock agriculture, and comparison of data sources. Dashboards include tabs for visualisations, maps, data table downloads, and metadata. The findability and interoperability of data is improved through our metadata's use of a machine-actionable formats such as JavaScript object notation for linked data, more commonly known as json-ld, using metadata terms from the Dublin Core Metadata Initiative and schema.org [19, 20]. When external metadata were non-existent or did not adhere to standards, information was gathered from grey literature such as data collection protocols, data manuals, or the website or reports. In some cases, information about the data and data licensing

information was obtained from an established contact point or expert. Thus, each dashboard gives full access to the metadata for all data displayed or used in calculations or models. All dashboards can be accessed through the GBADs Knowledge Engine at http://gbadske.org.

Reports

Data reports are being developed to explore whether different data sources reporting similar metrics (i.e. national population by livestock species) agree in terms of numeric value, and whether data can be understood, documented, and modelled longitudinally using numbers on their own and via information in available metadata. One example of this phenomenon is animal population numbers collected by different agencies. The question we should be asking of all models and calculations is what were the criteria for selecting a particular data source and what impact does this choice have on the model? To aid modellers in assessing the multiple data sources, reports have been generated that analyse each data source and compare them to similar ones. These analyses are under development with global data sources such as FAOSTAT and WOAH and with data from national sources. Current reports include a health and mortality analysis of data from livestock reports from EthCSA, a comparison of population numbers from different data sources, and an analysis of species categorisation in population data.

Data stories

A series of data stories are being developed to illustrate the potential insights that can be gleaned from GBADs data and to stimulate outside ideas on using GBADs' data portal to add value to their work. These stories will highlight the utility of livestock data made more accessible via GBADs and bring together data from different initiatives to address key topics synergistically, providing fresh insight and ways of viewing them. The exercise of developing these stories also helps us identify interoperability challenges and differences in key methodologies between sectors and cross-check the narrative of data from different sources. For example, new estimates of human disease burdens from GBD in Ethiopia [21] can be compared with comparable livestock

disease data (recently made more accessible by GBADs) to identify trends and relationships between human and animal health at the regional level. The first data story on animal mortality in Ethiopia is available at http://gbadske.org.

Global Burden of Animal Diseases application programming interface

The GBADs API allows access to all datasets stored by GBADs in our database tables, as well as access to other external APIs. The first type of access is standard for APIs, but has features that make it easy to use for many purposes:

1) fields to return can be selected so that the user can return as much or as little of the data record as they need;

2) sophisticated queries are allowed including joins between tables and ordering of the records returned;

3) since there are differences in species naming conventions between datasets, generic 'superclasses' of species can be used and these are transferred into the appropriate name used by the dataset.

This last feature allows users to request 'All Sheep' and get back the records corresponding to 'Sheep', 'Adult sheep', and 'Lambs' in the WOAH population dataset while getting 'Sheep' records from the FAOSTAT population dataset (Table QCL).

If the dataset to be searched is not stored by GBADs then the appropriate external API request is constructed, executed, and the results are filtered to correspond to GBADs API standards which include field selection, all units are a specific singular unit (e.g. head instead of 1,000 head for animal counts, $US for certain economic indicators, etc.). This design gives the user a seamless view of all datasets available through the GBADs data portal. The GBADs Informatics website (http://gbadske.org) provides directions to the API web interface and documentation.

Documentation site

All documentation for the data portal and Knowledge Engine reside in a single repository (https://gbadskedoc.org). This website (powered by Docusaurus [22]) provides a portal to user and developer documentation, tutorials and other learning materials, relevant papers and presentations, and access to GBADs Informatics YouTube channel (presentations and tutorials). There is also a blog feature where GBADs Informatics team members provide insights into the work of the team and the products being produced. The website has been designed to be a dynamic resource for all GBADs stakeholders, and is linked to the main GBADs site (https://animalhealthmetrics.org) which provides more information about GBADs' overall programme and vision.

Global Burden of Animal Diseases in the context of One Health data

The data collated, insights produced, and resources developed by GBADs integrate across the One Health community. Direct and indirect human health impacts include dietary impacts, emerging disease risks, foodborne disease risks, and zoonotic disease transmission. GBADs outputs could also be integrated with crop loss data [2], climate data, better AMU and AMR monitoring and evaluation, and societal outcomes like gender equity and livelihood impact. In addition to its economic value, new driving forces are impacting livestock production, including the global sustainable development goals (SDGs) [23]. Sustainability involves not only environmental impact, food safety and security, but also the challenge of assuring animal welfare (AW) [24].

Animal welfare data

WOAH defines AW as 'the physical and mental state of an animal in relation to the conditions in which it lives and dies'. In addition to defining the concept, a set of AW guidelines were developed for farm animals [25]. The aim is to prevent unnecessary suffering, safeguard animal health, and improve food safety and quality. Given the various cultural, religious and political backgrounds of WOAH's member states' along with the ethical, social and economic components shaping

AW [26], closing the gap between AW policy and its practical application remains a challenge. The economic impact of AW depends on how the cost of implementing welfare measures compares to its impact on production costs [27], which often affects compliance with AW guidelines. For instance, WOAH does not yet have guidelines for laying hens given the complex welfare issues deriving from each housing system and the economic costs of switching housing systems for this species [28].

Furthermore, most AW guidelines are based on the five freedoms [29]. Depending on the production system, some of these freedoms can be easily fulfilled like the provision of *ad libitum* feed and water and enough space in intensive dairy productions. Intensive production, however, does not prevent animals from experiencing some forms of pain, injury and disease. On the other hand, although extensive systems do offer AW opportunities (like expressing natural behaviours), chronic welfare issues still persist (like chronic thirst, and exposure to harsh climates) [30]. This is to say that introducing a few changes to a production system or completely substituting it can easily result in exchanging some welfare issues for others. This not only results in difficulties prioritising AW issues, but also in identifying the right data to best reflect welfare impairment.

GBADs takes a holistic overview of animal disease when estimating economic losses. Indeed, along with communicable and non-communicable diseases, GBADs considers the lack of feed and water, the presence of injuries, accidents, and predation as part of disease. Given the intersection of the five freedoms with GBADs disease list, this holistic approach can provide a clearer understanding of the key aspects impeding good AW based on the production system. As GBADs AW methodology is developing, identifying, and harmonising the data from open and private sources is ongoing. There are different ways to measure AW including animal-based and resource-based outcomes, along with data on slaughterhouse performance. Data can be collected by trained individuals using assessment protocols, by collecting body fluids and secretions, or through precision livestock farming techniques [31, 32, 33]. However, such data is considered

confidential and must meet data privacy policies, which differ among regions. The question of interoperability and how to combine data from different sources to provide clear insights of the burden of AW remains a challenge [34]. A scaling system of the welfare opportunities available in each production system would assist policy makers to better address welfare concerns.

One Health data for the sustainable development goals

GBADs outputs could improve research and policy in adjacent sectors to contribute to SDGs [35] (Figure 2):

Human dietary health

Most estimates of national food or nutrient supplies begin with production, utilisation, and trade data on agricultural commodities from FAOSTAT [36, 37, 38]. GBADs will improve the accuracy and interoperability of animal-sourced food production numbers and trade patterns, which would strengthen approaches to estimating global nutrient supplies. GBADs is also developing a globally applicable livestock production classification system. Greater detail on livestock breed and other characteristics may increase accuracy of nutrient availability metrics from animal-sourced foods.

Zoonoses and disease prioritisation

Prioritisation exercises help stakeholders methodically rank zoonotic diseases by threat to their designated area or country [39]. Ranking of diseases confers priority for those with a high score, although there is often insufficient data to support these rankings. GBADs will increase access to data on production and disease, clearing some data gaps and highlighting others. This will support impactful disease prioritisation and effective policy design.

The human health theme of GBADs is also studying zoonoses (such as brucellosis) which are not yet directly estimated in the GBD. Through this theme, GBADs will highlight the need to develop consistent metrics for these diseases to uniformly assess their impacts on livestock and human health.

Antimicrobial resistance

AMR is a complex One Health problem, as it involves humans, livestock, and companion animals. This leads to challenges in data collection and interoperability. AMU drives resistance, therefore it is important to measure AMU to understand AMR. WOAH is leading efforts to systematically collect AMU data from member countries [40]. In parallel, GBADs will estimate the cost of antimicrobials as a component of livestock expenditures, as well as the cost of resistance in livestock and livestock production. This work will complement recent work by GBD regarding AMR in humans [41] to better define the total burden of AMR across humans and animals.

Climate health

Differences in livestock population or biomass estimates, or in descriptions and estimates of production systems, can contribute to large variations in attribution of greenhouse gas emissions to livestock [42]. Accurate numbers are paramount to developing effective climate change mitigation strategies and for encouraging confidence in and support for these measures [43]. As the climate emergency escalates, it will be increasingly important to decide how to allocate crops and cropland to human consumption versus animal feed. GBADs estimates of livestock production and classification of production system typologies can be used in tandem with data from the Global Burden of Crop Loss [2] and climate data from the National Oceanic and Atmospheric Administration, the National Aeronautics and Space Administration, and other organisations to support decisions on land-use and other environmentally relevant concerns.

Livelihoods and gender equity

GBADs data on livestock health loss and attribution, production loss and expenditure, and total economic value of livestock will increase understanding of who in a community benefits from the production and sale of livestock commodities. GBADs aims to assess which community members are least able to weather livestock health shocks and understand how such shocks economically impact different people.

Mobilising this information could improve the livelihoods of producers on the margins and support gender equity in livestock production globally.

Conclusions

GBADs Informatics theme has used advanced analytics to progress towards an end-to-end solution that seamlessly integrates and unlocks the potential of available animal health data. We hope to stimulate discussion and inspire collaboration by liberating livestock data and disseminating GBADs knowledge products through a variety of media. Data from GBADs and other global burdens (human health, crop loss, foodborne diseases) can be combined with other major data initiatives to constitute One Health data. Thus, helping to measure and meet SDGs surrounding poverty, hunger, fair wages, inequalities, and in developing a sustainable future.

Acknowledgements

Led by WOAH and the University of Liverpool, GBADs achieved USD$7m funding from the Bill and Melinda Gates Foundation and the United Kingdom's Foreign Commonwealth and Development Office between September 2020 to December 2022 and initiated activities on 1st January 2021.

References

[1] Rushton J., Huntington B. […] & Mesenhowski S. (2021). – Roll-out of the Global Burden of Animal Diseases programme. *Lancet*, **397** (10279), 1045–1046. https://doi.org/10.1016/S0140-6736(21)00189-6

[2] Finegold C., Ried J., Denby K. & Gurr S. (2019). – Global burden of crop loss. *Gates Open Res.*, **3**, 1599 (poster). https://doi.org/10.21955/gatesopenres.1116448.1

[3] Havelaar A.H., Kirk M.D. […] & Devleesschauwer B. (2015). – World Health Organization global estimates and regional comparisons of the burden of foodborne disease in 2010. *PLoS Med.*, **12** (12), e1001923. https://doi.org/10.1371/journal.pmed.1001923

[4] Bernardo T., Sobkowich K.E., Forrest R.O., Stewart L.S., D'Agostino M., Gutierrez E.P. & Gillis D. (2021). – Collaborating in the time of COVID-19: the scope and scale of innovative responses to a global pandemic. *JMIR Public Health Surveill.*, **7** (2), e25935. https://doi.org/10.2196/25935

[5] Institute for Health Metrics and Evaluation (IHME) (2022). – About GBD. IHME, Seattle, United States of America. Available at: https://www.healthdata.org/gbd/about (accessed on 15 June 2022).

[6] Rongen G., Pieter't Hart C.M., Leontaris G. & Morales-Nápoles M. (2022). – Anduryl. Available at: https://github.com/grongen/anduryl (accessed on 21 June 2022).

[7] World Organisation for Animal Health (WOAH) (2018). – Notification procedure for annual report. WOAH, Paris, France.

[8] Food and Agriculture Organization of the United Nations Corporate Statistical Database (FAOSTAT) (2022). – FAOSTAT crops and livestock products: metadata. FAO, Rome, Italy. Available at: https://www.fao.org/faostat/en/#data/QCL/metadata (accessed on 14 June 2022).

[9] Food and Agriculture Organization of the United Nations Corporate Statistical Database (FAOSTAT) (2022). – FAOSTAT enteric fermentation: metadata. FAO, Rome, Italy. Available at: https://www.fao.org/faostat/en/#data/GE/metadata (accessed on 14 June 2022).

[10] Statistical office of the European Union (Eurostat) (2023). – Data Browser: bovine population: annual data. European Union, Brussels, Belgium. Available at: https://ec.europa.eu/eurostat/databrowser/view/APRO_MT_LSCATL/default/table?lang=en (accessed on 23 March 2023).

[11] Open Data Institute (ODI) (2022). – The data spectrum. ODI, London, United Kingdom. Available at: https://theodi.org/about-the-odi/the-data-spectrum (accessed on 15 June 2022).

[12] Global Burden of Animal Diseases (GBADs) Informatics (2022). – The GBADs Knowledge Engine. GBADs Informatics, Guelph, Canada. Available at: http://gbadske.org (accessed on 15 June 2022).

[13] The World Bank (2021). – Creating an integrated national data system. *In* World Development Report 2021. The World Bank Group, Washington, DC, United States of America, 303–326. Available at: https://openknowledge.worldbank.org/bitstream/handle/10986/35218/9781464816000_Ch09.pdf (accessed on 18 September 2022).

[14] Wilkinson M.D., Dumontier M. […] & Mons B. (2016). – The FAIR guiding principles for scientific data management and stewardship. *Sci. Data*, **3**, 160018. https://doi.org/10.1038/sdata.2016.18

[15] Stacey D., Wulff K., Chikhalla N. & Bernardo T. (2022). – From FAIR to FAIRS: data security by design for the global burden of animal diseases. *Agron. J.*, **114** (5), 2693–2699. https://doi.org/10.1002/agj2.21017

[16] World Organisation for Animal Health (WOAH) (2022). – WAHIS: World Animal Health Information System database. WOAH, Paris, France. Available at: https://wahis.woah.org/#/home (accessed on 1 February 2022).

[17] Nagowah S.D., Ben Sta H. & Gobin-Rahimbux B.A. (2018). – An overview of semantic interoperability ontologies and frameworks for IoT. *In* Proc. 2018 Sixth International Conference on Enterprise Systems (ES), 1–2 October, Limassol, Cyprus. IEEE, New York, United States of America, 82–89. https://doi.org/10.1109/ES.2018.00020

[18] Wexler S., Shaffer J. & Cotgreave A. (2017). – The big book of dashboards: visualizing your data using real-world business scenarios. 1st Ed. Wiley, Hoboken, United States of America, 448 pp.

[19] Dublin Core Metadata Initiative (DCMI) Usage Board (2020). – DCMI metadata terms. Available at: https://www.dublincore.org/specifications/dublin-core/dcmi-terms (accessed on 15 June 2022).

[20] Schema.org (2022). – Welcome to Schema.org. Available at: https://schema.org (accessed on 15 June 2022).

[21] GBD 2019 Ethiopia Subnational-Level Disease Burden Initiative Collaborators (2022). – Progress in health among regions of Ethiopia, 1990–2019: a subnational country analysis for the Global Burden of Disease Study 2019. *Lancet*, **399** (10332), 1322–1335. https://doi.org/10.1016/S0140-6736(21)02868-3

[22] Meta Open Source (2022). – Docusaurus. Meta Open Source, Menlo Park, United States of America. Available at: https://docusaurus.io (accessed on 17 June 2022).

[23] Barbier E.B. & Burgess J.C. (2017). – The sustainable development goals and the systems approach to sustainability. *Economics*, **11** (1), 1–22. https://doi.org/10.5018/economics-ejournal.ja.2017-28

[24] Keeling L., Tunón H., Olmos Antillón G., Berg C., Jones M., Stuardo L., Swanson J., Wallenbeck A., Winckler C. & Blokhuis H. (2019). – Animal welfare and the United Nations sustainable development goals. *Front. Vet. Sci.*, **6**, 336. https://doi.org/10.3389/fvets.2019.00336

[25] World Organisation for Animal Health (WOAH) (2022). – Section 7. Animal welfare. *In* Terrestrial Animal Health Code. WOAH, Paris, France. Available at: https://www.woah.org/en/what-we-do/standards/codes-and-manuals/terrestrial-code-online-access/?id=169&L=1&htmfile=titre_1.7.htm (accessed on 19 September 2022).

[26] McBride E.A. & Baugh S. (2022). – Animal welfare in context: historical, scientific, ethical, moral and one welfare perspectives. *In* Human/animal relationships in transformation (A. Vitale & S. Pollo, eds). Palgrave Macmillan, Cham, Switzerland, 119–147. https://doi.org/10.1007/978-3-030-85277-1_7

[27] Fernandes J.N., Hemsworth P.H., Coleman G.J. & Tilbrook A.J. (2021). – Costs and benefits of improving farm animal welfare. *Agriculture*, **11** (2), 104. https://doi.org/10.3390/agriculture11020104

[28] Matthews W.A. & Sumner D.A. (2015). – Effects of housing system on the costs of commercial egg production. *Poult. Sci.*, **94** (3), 552–557. https://doi.org/10.3382/ps/peu011

[29] Brambell F.W.R. (1965). – Report of the technical committee to enquire into the welfare of animals kept under intensive livestock husbandry systems. Her Majesty's Stationery Office, London, United Kingdom, 84 pp.

[30] Temple D. & Manteca X. (2020). – Animal welfare in extensive production systems is still an area of concern. *Front. Sustain. Food Syst.*, **4**, 545902. https://doi.org/10.3389/fsufs.2020.545902

[31] Buller H., Blokhuis H., Lokhorst K., Silberberg M. & Veissier I. (2020). – Animal welfare management in a digital world. *Animals*, **10** (10), 1779. https://doi.org/10.3390/ani10101779

[32] Stygar A.H., Gómez Y., Berteselli G.V., Dalla Costa E., Canali E., Niemi J.K., Llonch P. & Pastell M. (2021). – A systematic review on commercially available and validated sensor technologies for welfare assessment of dairy cattle. *Front. Vet. Sci.*, **8**, 634338. https://doi.org/10.3389/fvets.2021.634338

[33] Gómez Y., Stygar A.H., Boumans I.J.M.M., Bokkers E.A.M., Pedersen L.J., Niemi J.K., Pastell M., Manteca X. & Llonch P. (2021). – A systematic review on validated precision livestock farming technologies for pig production and its potential to assess animal welfare. *Front. Vet. Sci.*, **8**, 660565. https://doi.org/10.3389/fvets.2021.660565

[34] 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. *Comput. Electron. Agric.*, **156**, 459–466. https://doi.org/10.1016/j.compag.2018.12.007

[35] United Nations (2015). – Transforming our world: the 2030 agenda for sustainable development. United Nations, New York, United States of America, 41 pp. Available at: https://sdgs.un.org/sites/default/files/publications/21252030%20Agenda%20for%20Sustainable%20Development%20web.pdf (accessed on 16 June 2022).

[36] Smith M.R., Micha R., Golden C.D., Mozaffarian D. & Myers S.S. (2016). – Global Expanded Nutrient Supply (GENuS) model: a new method for estimating the global dietary supply of nutrients. *PLoS One*, **11** (1), e0146976. https://doi.org/10.1371/journal.pone.0146976

[37] Schmidhuber J., Sur P., Fay K., Huntley B., Salama J., Lee A., Cornaby L., Horino M., Murray C. & Afshin A. (2018). – The Global Nutrient Database: availability of macronutrients and micronutrients in 195 countries from 1980 to 2013. *Lancet Planet. Health*, **2** (8), e353–e368. https://doi.org/10.1016/S2542-5196(18)30170-0

[38] Beal T., Massiot E., Arsenault J.E., Smith M.R. & Hijmans R.J. (2017). – Global trends in dietary micronutrient supplies and estimated prevalence of inadequate intakes. *PLoS One*, **12** (4), e0175554. https://doi.org/10.1371/journal.pone.0175554

[39] Rist C.L., Arriola C.S. & Rubin C. (2014). – Prioritizing zoonoses: a proposed One Health tool for collaborative decision-making. *PLoS One*, **9** (10), e109986. https://doi.org/10.1371/journal.pone.0109986

[40] World Organisation for Animal Health (OIE) (2021). – OIE annual report on antimicrobial agents intended for use in animals: better understanding of the global situation. 5th report. OIE, Paris, France, 134 pp. Available at: https://www.woah.org/app/uploads/2021/05/a-fifth-annual-report-amr.pdf (accessed on 16 June 2022).

[41] Antimicrobial Resistance Collaborators (2022). – Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *Lancet*, **399** (10325), 629–655. https://doi.org/10.1016/S0140-6736(21)02724-0

[42] Herrero M., Gerber P. […] & McAllister T.A. (2011). – Livestock and greenhouse gas emissions: the importance of getting the numbers right. *Anim. Feed Sci. Technol.*, **166–167**, 779–782. https://doi.org/10.1016/j.anifeedsci.2011.04.083

[43] Salmon G.R., MacLeod M., Claxton J.R., Pica Ciamarra U., Robinson T., Duncan A. & Peters A.R. (2020). – Exploring the landscape of livestock 'Facts'. *Glob. Food Sec.*, **25**, 100329. https://doi.org/10.1016/j.gfs.2019.100329

Table I

Global Burden of Animal Diseases Informatics outputs as of 25 June 2022

Links and access information for dashboards, documentation, and other data products are available via the GBADs Informatics Knowledge Engine website (http://gbadske.org)

Disclaimer: the Informatics theme is responsible for the curation, visualisation, and dissemination of the data whereas the other themes are responsible for validating the models they provide

Product type	Product	Description
Datasets produced by models	Output datasets	Output datasets are produced by models that are created and validated by collaborating modellers. The outputs are ingested and stored in database tables and made accessible via APIs
Dashboards	Livestock populations	Livestock population numbers displayed by country from three different data sources (Eurostat, FAOSTAT, WOAH)
	Biomass	Biomass calculations from GBADs displayed by year, species, and country. Multiple data inputs/estimation methods available
	Total economic value	Global total economic value of livestock and livestock products, multiple denominators available
	Ethiopian data sandbox	Access and visualise data scraped from the Ethiopian Central Statistical Agency. A case study for national and regional level analysis for GBADs
	Dashboard modules	Framework for rapid development and deployment of new dashboards. Tutorials are available to quickly on-board new developers
Reports and data products	Data reports	Rapidly generatable pdf reports for quick comparison of key data points over time. Topics in progress include national livestock mortality trends, livestock genetics, and data source comparisons

	Data stories	A medium to demonstrate the utility of data provided through GBADs, highlight avenues for future work and collaboration, and inspire creative thinking
Results dissemination	Informatics site	Location to host results and documentation content, including tutorials, news updates and highlights, blog posts and explainers, publications and presentations, and videos
	External presentations and publications	List of and links to presentations and publications produced by GBADs Informatics
	Style guide	Using best practices for dashboard design and colour theory, the GBADs style guides provide a consistent colour scheme and framework for data visualisations, presentations, and dashboards. The guide ensures that outputs can be identified as GBADs' outputs
	Amazon S3 buckets	Secure storage of metadata, and provenance, data cleaning, and model methodology documents. Available upon request
Documentation	GitHub repositories	Collection of publicly available code to reproduce GBADs methodology
	Tutorials	Written and video tutorials to replicate major processes and products in the Knowledge Engine
	Data governance handbook	A living document which acts as a manifesto for the intended use of data in GBADs while also providing guidance and documentation on best practices for data management for GBADs stakeholders
Data acquisition and access	Web scraping scripts	Programmes that transform data in pdfs tables and reports to a more usable and accessible format (csv, json). These scripts are available via the GBADs GitHub
	Database tables	Database tables store data collected and disseminated by GBADs. The GBADs API allows users to access the data stored in these tables
	APIs	Publicly accessible tool to access all data used and created by GBADs, including data previously not available in accessible formats
	Metadata	GBADs metadata and data provenance information are available via GBADs dashboards. Links to source

		metadata (as provided by the data source) are also provided when applicable
	User authentication system	In anticipation of private data, this system authenticates users via a log-in system to keep private data secure and accessible to only those who have permission
Interoperability tools	Graph databases	Under development for storage and management of data resources that change over time, between and within different sources, and across geographic regions. These also serve to document data provenance and lineage information
	Ontology	Livestock production and disease ontology under development to facilitate interoperability of livestock data and models from disparate sources
	Data quality assessment tools	Programmes that compare how species were reported by different countries and data sources to measure their level of agreement and identify potential data entry errors or anomalies
Miscellaneous	Horizon scanning report	Report examining determinants of success and failure in large data aggregation initiatives. Critical for developing a data strategy and future planning
	Anduryl accessibility improvements	The open-source expert elicitation tool, Anduryl [6], was adapted to be accessible through RStudio. Adapted version available via GBADs GitHub

API(s): application programming interface(s)
csv: comma-separated values
Eurostat: statistical office of the European Union
FAOSTAT: Food and Agriculture Organization of the United Nations Corporate Statistical Database
GBADs: Global Burden of Animal Diseases
json: JavaScript object notation
pdf(s): portable document format(s)
WOAH: World Organisation for Animal Health

Table II

Differences in cattle, buffalo and bovine species categorisations according to major data sources

Data source	WOAH[(a)]	FAOSTAT	FAOSTAT	Eurostat
Dataset	**WAHIS**	**Live animals: QCL[(b)]**	**FAO tier 1: GE[(c)]**	**Number of bovine animals[(d)]**
Categories	Cattle	Cattle	Cattle	Live bovine animals
	Adult beef cattle (2+ years)	Cattle and buffalo	Cattle, dairy	Bovine animals (less than 1 year old)
	Adult dairy cattle (2+ years)		Cattle, non-dairy	Bovine animals (less than 1 year old) for slaughter
	Male and female cattle (1–2 years)			Bovine animals (less than 1 year old) not for slaughter
	Calves (<1 year)			Bovine animals (1 to less than 2 years old)
	Buffaloes			

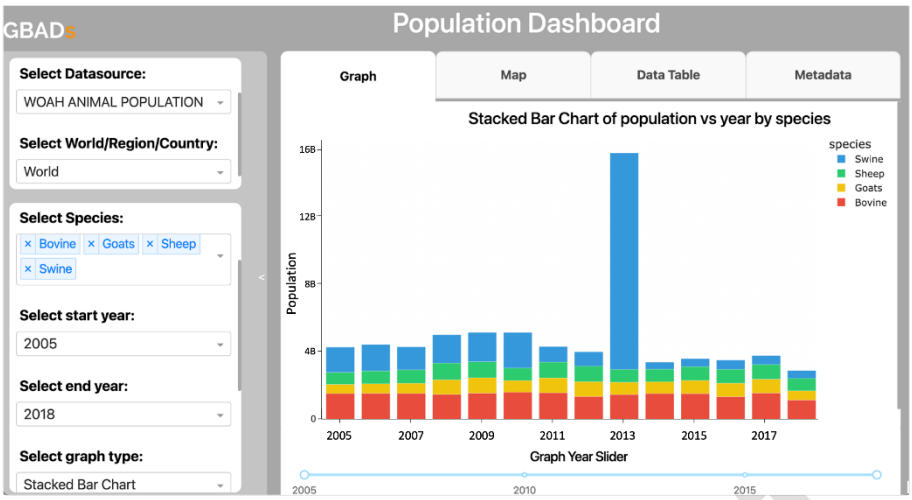
(a) WOAH (categories according to 2018 annual reporting guidelines) [7]
(b) Metadata from FAOSTAT – crops and livestock products (Live animals: QCL) [8]
(c) Metadata from FAOSTAT – climate change – enteric fermentation – Tier 1 [9]. The FAOSTAT emissions database is computed following Tier 1 2006 IPCC Guidelines for national greenhouse gas inventories. Vol. 4, chapters 10 and 11 (http://www.ipcc-nggip.iges.or.jp/public/2006gl/vol4.html)
(d) Eurostat (bovine population – annual data) [10]
All data accessed on 14 June 2022

Eurostat: statistical office of the European Union
FAO: Food and Agriculture Organization of the United Nations
FAOSTAT: Food and Agriculture Organization of the United Nations Corporate Statistical Database
IPCC: Intergovernmental Panel on Climate Change
WAHIS: World Animal Health Information System
WOAH: World Organisation for Animal Health

Table III

Changes in cattle and buffalo species categorisations over time as described in World Organisation for Animal Health annual reporting guidelines documents [7]

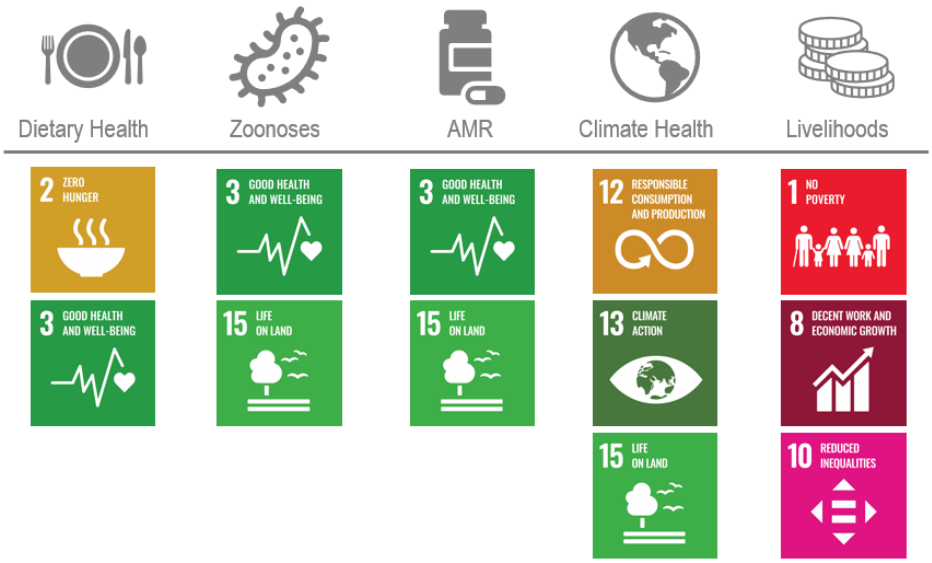
2009–2014	2015–2016	2017–2019
Cattle	Cattle	Cattle
Buffaloes (not *Syncerus caffer*)	Buffaloes	Adult beef cattle (2+ years)
		Adult dairy cattle (2+ years)
		Male and female cattle (1–2 years)
		Calves (<1 year)
		Buffaloes



GBADs: Global Burden of Animal Diseases
WOAH: World Organisation for Animal Health

Figure 1

Data visualisation (World Organisation for Animal Health global animal population from 2005 to 2018 [16]) **highlights likely error in swine data in 2013** (errors have since been rectified)



AMR: antimicrobial resistance

Figure 2

Global Burden of Animal Diseases (GBADs) and the sustainable development goals

Sustainable development goals [35] which GBADs outputs contribute towards in various sectors