



Best practice guidance note on the use of digitalisation tools produced in the Digivet project

Orla Shortall¹, Carol Kyle¹, Lisa Boden², Carlijn Bogaardt³, Wonhee Cha⁴,
Matthew Denwood⁵, Jessica Enright³, Jenny Frössling⁴, Wiktor Gustafsson⁴,
Søren Nielsen⁵, Petter Hopp⁶, Alice Puk Skarbye⁵

¹ The James Hutton Institute, ² University of Edinburgh, ³ University of Glasgow, ⁴ National Veterinary Institute - SVA, ⁵ University of Copenhagen, ⁶ Norwegian Veterinary Institute

Contents

Executive summary.....	3
Introduction.....	3
Methods.....	3
Findings	4
Data ecosystems.....	4
Feedback on digitalisation tools.....	6
Summary and conclusion	8

Executive summary

Three workshops held in the UK, Denmark and Sweden explored challenges and opportunities in data ecosystems in relation to African Swine Fever (ASF), *Salmonella* Dublin (S. Dublin) and antimicrobial use (AMU) respectively. The workshops elicited feedback on four data digitalisation tools: movenet, sdubcontrol, goldeneye and AMView. Common challenges across countries and case studies included the quality of data entry, difficulties aggregating data, lack of active data curation, challenges with data sharing, under-use of data to provide return value and concerns about data privacy, particularly at farm level. Differences in perceptions of challenges and strengths between case studies and countries included different levels of data openness, with Denmark having the highest levels of data openness and sharing across the case studies. There were also different perceptions of the desirability of data openness and privacy between actors.

Feedback on data digitalisation tools considered the tools' users, utility, usability and wider social and ethical issues. Functions such as data anonymisation (movenet) and a data de-encryption "kill switch" to remove access to data at the end of a project (goldeneye) were seen to facilitate data sharing and security, though they were not a replacement for trust within data ecosystems. Rather they could make data management easier and more secure within a system where trust already exists. AMView, movenet and sdubcontrol could help provide information and support decision making. Challenges with using tools (sdubcontrol and AMView) for information provision and decision making include the accuracy of the data on which they are based, providing useful information at different scales while maintaining individuals' privacy, and difficulties with data interpretation. More specific guidance on the use of the tools produced within the Digivet project can be found in deliverable 2.4.

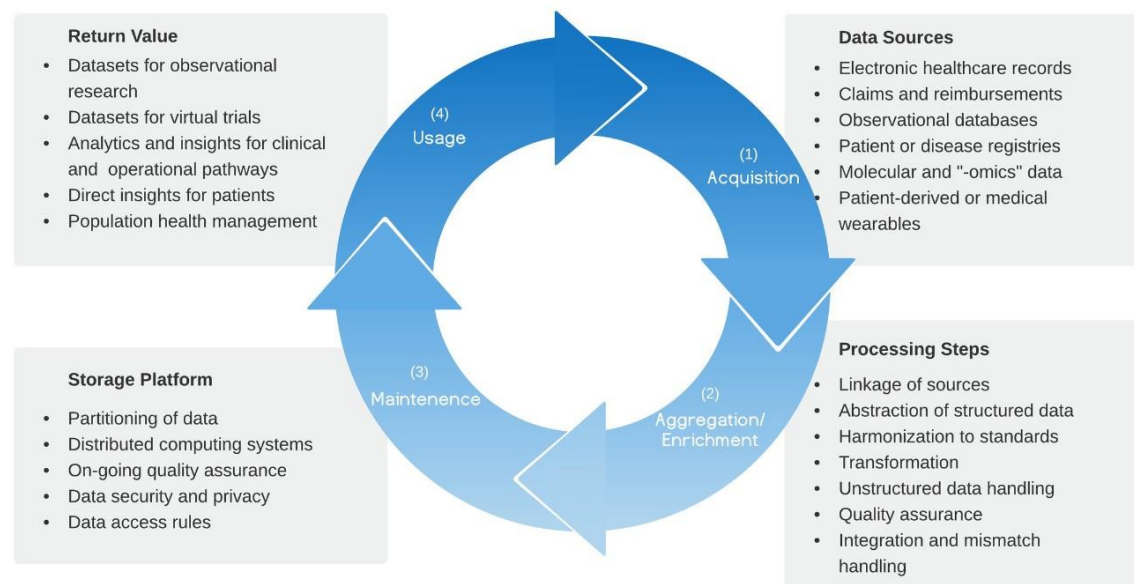
Introduction

This research note presents findings on the use of data digitalisation tools developed in the Digivet project based on the results of three workshops held in the UK, Denmark and Sweden in 2023. The workshops elicited feedback on data digitalisation tools developed as part of the Digivet project and explored strengths and challenges in data digitalisation across countries. The workshops focused on three case studies: African Swine Fever (ASF) (UK), *Salmonella* Dublin (S. Dublin) (Denmark), and antimicrobial use (AMU) (Sweden). Workshops were attended by stakeholders from academia, government and the agriculture industry from project partner countries: Denmark, the UK, Sweden, Norway and Estonia. More details of the workshops and results can be found in individual country reports produced as part of the Digivet project. Specific guidance on the use of each of the digitalisation tools is also published by this project as part of deliverable 2.4.

Methods

Facilitators took notes of discussions during the workshop. These notes were digitised and analysed using the data lifecycle theoretical framework, adapted from Zhang et al. (2022), which was produced in relation to data use in the human health field, as can be seen in Figure 1.

Figure 1 The data lifecycle taken from Zhang et al. (2022)



Findings

Data ecosystems

The workshop findings on challenges and strengths within the current data ecosystems were compared and contrasted across countries and case studies. The results presented in this research note show a comparison between the UK, Denmark and Sweden excluding Estonia and Norway because fewer stakeholders from Estonia and Norway attended the workshops, and there was not sufficient data on all case studies.

Table 1 Similarities in strengths and challenges within current data ecosystems across countries and case studies

Data Aquisition (connectivity, wearables, electronic identification, decision-support tools, diagnostics etc.)	Data Aggregation and Enrichment (data handling, interoperability, harmonisation etc.)	Data Management and Maintenance (partitioning, access, security etc.)	Data Usage and Return Value (Costs, relevance, user-friendliness, trust, transparency etc)
Data quality challenges: farmers, vets and stakeholders not inputting data, not updating data frequently enough, inaccurate or missing data.	Data entered in different formats that are difficult to aggregate (particularly AMU).	Data sharing is difficult because of lack of agreement and GDPR restrictions.	Data under-use – not used enough to support decision making (particularly at farm level in relation to AMU).
		Perception that data sharing is becoming more difficult.	Data access for use in academia can be difficult.
		Not enough active data quality control and curation.	There are concerns about data privacy at farm level.

Workshop participants reported similar challenges across case studies and countries in relation to data acquisition in terms of inconsistent data quality and frequency of data entry.

Farmers, vets or other stakeholders might not enter data correctly; may enter data too late to comply with legal requirements and for the data to be accurate, such as notifying pig movements for the ASF case study; and may enter data relating to AMU in different formats. Data aggregation could also be challenging if data was entered in different formats, particularly in relation to AMU.

Workshop participants all reported challenges in relation to data sharing, due to what they saw as legal complexity, procedural difficulties and at times a lack of agreement between stakeholders about the how data should be shared and used. There was a perception that data sharing had become more difficult in recent years, potentially because of 2018 changes in GDPR legislation. Data sharing was seen to be particularly challenging in the UK where the four devolved nations may collect and store data in different systems using different formats. There were also concerns about a lack of data curation and quality control issues in the management of data across countries and case studies, for instance in the removal of outdated information.

In relation to data use, workshop participants across the different countries stated that AMU data in particular could have better return value for stakeholders, in terms of being used by farmers and vets for benchmarking. Across the case studies, some participants stated that it can be difficult for academics to access data, which has an impact on the return value of the data. Participants expressed concerns about farmers' privacy as a reason for limiting data access. It was stated that reporting on sensitive farm level information could stigmatise farmers and create reputation risks for the industry.

Table 2 shows differences between the countries and case studies in relation to challenges and strengths in existing data ecosystems in the UK, Denmark and Sweden.

Table 2

Data Aquisition (connectivity, wearables, electronic identification, decision-support tools, diagnostics etc.)	Data Aggregation and Enrichment (data handling, interoperability, harmonisation etc.)	Data Management and Maintenance (partitioning, access, security etc.)	Data Usage and Return Value (Costs, relevance, user-friendliness, trust, transparency etc)
Different kinds of data collected e.g. Sweden and Denmark have a system of active surveillance for S. Dublin which the UK does not.		Data sharing seen to be easier in some contexts than others.	Perceptions of the desirability of data openness and privacy differed between countries and stakeholders.
Differences in AMU data collection across countries and sectors.		Differences in data openness between countries: Denmark has the highest level of data openness.	Perception farmers and industry bodies want more privacy in the use of data.

In terms of differences across case studies and countries, different kinds of data were collected across the countries. Data was collected at farm level on S. Dublin status in Denmark and Sweden as part of disease eradication schemes, in contrast to a system of passive surveillance for S. Dublin in the UK where there is no S. Dublin eradication scheme. Data collection for AMU also differed: in some countries and sectors data on antimicrobial use was entered by vets or pharmacies, or in the case of the UK beef and dairy sectors,

farmers or vets; whereas other sectors did not have this kind of system, but rather AMU data was aggregated from sales data. In the ASF case study, pig holding and movement data showed some differences between countries in relation to how holdings were identified and the time period in which farmers were required to register pig movements.

While participants described data sharing as challenging across all countries, it was also seen to be easier in some contexts rather than others. This was also true of levels of data openness, as the types and amount of data that were shared between stakeholders and made publicly available also differed. Denmark had the highest level of data openness and data that was accessible to the public. In Sweden less data was accessible to the public, and the UK had the most stringent data privacy of the three countries in terms of data that was accessible to the public. Participants discussed different perspectives on the desirability of forms of data openness and privacy. Stakeholders stated that despite concerns about farmers' privacy in Denmark there was a culture of trust in relation to data sharing and data openness, facilitated by existing practices. There was a perception that different actors such as farmers and industry bodies may have a preference for greater levels of privacy in order to protect the reputation of the industry.

Feedback on digitalisation tools

At each of the three workshops digitalisation tools developed in the DigiVet project were presented. In the UK workshop, the movenet R package was presented with an accompanying interactive app, which facilitates the effective use of livestock movement data in veterinary public health. At the Danish workshop two tools were presented: the encryption package goldeneye and the sdubcontrol tool to help stakeholders make decisions in relation to the S. Dublin eradication programme. At the Swedish workshop the AMView R Shiny app was presented which shows a visual representation of AMU through trends and maps. More details of the tools can be found in the individual country reports. The tools involved interventions at different points on the data lifecycle, as can be seen in Figure 2. Movenet involved interventions at the level of data aggregation, through the anonymisation function, and at the management and use stages of the data life cycle through facilitating social network analysis of livestock movement data. Goldeneye acted at the data

maintenance stage through the encryption function. Sdubcontrol and AMView had data use functionality through helping stakeholders understand and make decisions based on data.

Figure 2 Representation of Digivet digitalisation tools within the data lifecycle

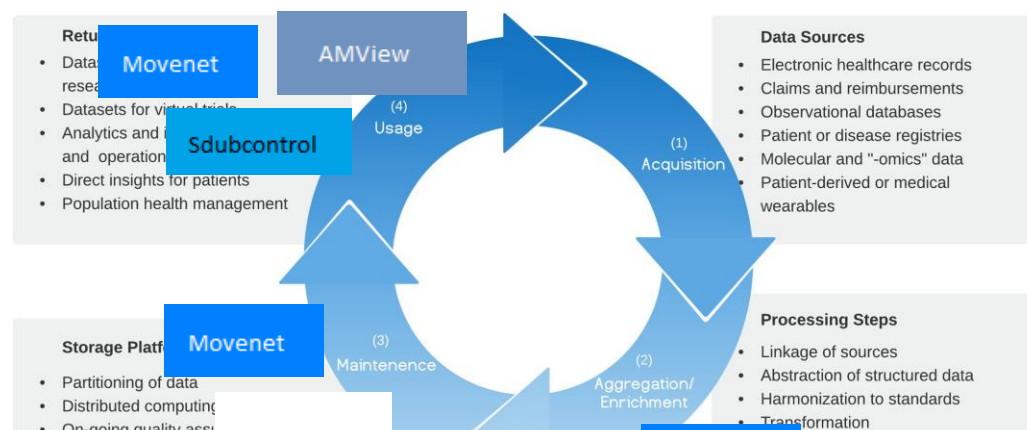


Table 3 shows stakeholder feedback on the tools provided during the workshops, organised thematically into comments on the target users, utility, usability and ethical and social issues within the data ecosystem.

Table 3 Stakeholder feedback on data digitalisation tools developed in the Digivet project.

Theme	Movenet	Sdubcontrol	Goldeneye	AMView
Target users	The tool is targeted at users with a certain amount of data management expertise.	The potential users could be broadened to vets, farmers and industry stakeholders as well as stakeholders involved in decision making on S. Dublin control programmes.	The tool is useful for researchers and academics to facilitate data sharing.	The target user could be better specified for the tool.
Utility	<p>The tool could make data sharing easier in countries where data needs to be anonymised.</p> <p>In countries such as Denmark where data anonymisation is less of an issue, the tool could be repackaged to emphasise its other uses.</p>	<p>The tool could be developed to support test sampling strategies.</p> <p>The tool could be further developed to help farmers eradicate S. Dublin.</p> <p>The tool could be used to show farmers the potential outcomes of herd level changes to disease risks.</p>	<p>The “kill switch” feature of revoking access to decrypted data after a certain date is helpful.</p> <p>The tool could provide an additional layer of security around data sharing.</p>	<p>Additional functionality could be added to make the output useful for understanding AMU at different scales.</p> <p>The data could be presented with a denominator to make the tool more informative.</p> <p>The utility of the tool depends on the accuracy of the data entered into it.</p>
Usability	The tool could be made more visually appealing to make it more accessible for users who are not data scientists.	<p>The tool could use visualisation instead of numbers to make it easier to use.</p> <p>The effects of changing inputs in the tool are complex so use by farmers could be mediated by vets and other advisors.</p>	Maintenance and updates of the tool would have to be undertaken to ensure the tool’s continuing usability.	The tool is intuitive and easy to use.
Ethical and social issues	Making data more accessible has knock-on effects in terms of resource needs to meet data sharing requests and the potential for the mis-use of data.	The tool could help motivate farmers to take a proactive role in S. Dublin control.	The tool does not replace the need for trust between partners in relation to data sharing.	The use of the tool for providing stakeholders with information, and potentially allowing farmers to benchmark their AMU would have to be carefully considered with respect to issues of privacy. Data access would need to be restricted at certain scales to certain actors to ensure privacy.

The data management tool goldeneye was seen to have the potential to enhance security in data sharing. Though stakeholders stated trust was still necessary between partners in the data ecosystem: there could be a way to overcome the tool’s “kill switch” data de-encryption feature. The tool could rather make data management easier and more secure within a system where trust already exists. Discussion about the movenet tool focused on how features of the tool could be better targeted and enhanced to meet the needs of diverse actors relating to differing stages of the data lifecycle. While facilitating data sharing through anonymising data was seen as beneficial, it was also recognised there were resource requirements to increased data sharing and this posed risks. The AMView and sdubcontrol tools both focused on the return value of data and discussions were had on how farmers and other stakeholders could best use the tools to support decision making around animal health.

Within the AMU case study data maps onto a particular behaviour: the use of antimicrobials on farm and/or antimicrobials prescribed by veterinary practices. For this reason, stakeholders stated that AMU data could be better used in benchmarking to help inform farmers' and vets' decisions about AMU and wider animal health management. However, participants also recognised that setting a desirable target for AMU benchmarking was complicated. While low AMU at farm level could be indicative of high biosecurity and standards of animal husbandry and health, setting a target of low antimicrobial use could also encourage farmers to reduce AMU to the detriment of animal health and welfare. Thus, data on AMU at any given scale is an imperfect measure of the desirability of outcomes. For this reason, the use of AMU data visualisation as a decision support tool needs to be carefully considered.

Summary and conclusion

Three workshops in the UK, Denmark and Sweden explored challenges and opportunities in data ecosystems in relation to ASF, S. Dublin and AMU respectively, and elicited feedback on four data digitalisation tools: Movenet, sdubcontrol, Goldeneye and AMView. Common challenges across countries and case studies included the quality of data entry, difficulties aggregating data, lack of active data curation, challenges with data sharing, under-use of data to provide return value and concerns about data privacy, particularly at farm level. Differences in perceptions of challenges and strengths between case studies and countries included different levels of data openness, with Denmark having the highest levels of data openness and sharing across the case studies. There were also different perceptions of the desirability of data openness and privacy between actors.

Feedback on data digitalisation tools consider the tools' users, utility, usability and wider social and ethical issues. Functions such as data anonymisation (movenet) and a data de-encryption "kill switch" to remove access to data at the end of a project (goldeneye) were seen to facilitate data sharing and security, though they were not a replacement for trust with the data ecosystems. Rather they could make data management easier and more secure within a system where trust already exists. Challenges with using tools (sdubcontrol and AMView) for information provision and decision making include the accuracy of the data on which they are based, providing useful information at different scales while maintaining individuals' privacy and difficulties with data interpretation. More specific guidance on the use of the tools produced within the Digivet project can be found in deliverable 2.4.

References

Zhang, J., Symons, J., Agapow, P., Teo, J.T., Paxton, C.A., Abdi, J., Mattie, H., Davie, C., Torres, A.Z., Folarin, A., Sood, H., Celi, L.A., Halamka, J., Eapen, S., & Budhdeo, S. (2022) Best practices in the real-world data life cycle. *PLOS Digital Health*, 1(1), e0000003.

<https://doi.org/10.1371/journal.pdig.0000003>