H2020-SFS-2018-2020

DECIDE

Data-driven control and prioritisation of non-EU-regulated contagious animal diseases

Deliverable 3.1

Requirement analysis

WP3 – Integration of data tools in disease control programs

Authors Jenny Frössling (SVA), Gema Vidal (SVA), Fernanda Dórea (SVA), Wiktor Gustafsson (SVA), Luis Pedro Carmo (NVI), Victor Henrique Silva de Oliveira (NVI), Britt Bang Jensen (NVI), Jade Bokma (UGent), Yara Slegers (UU), Camille Delavenne (AUSVET), Lucía Dieste Perez (GD), Anouk Veldhuis (GD), Sébastien Picault (INRAE), Carla Gomes (AHI), Annette Boerlage (SRUC), Maria Costa (SRUC), Beatriz García Morante (IRTA), Carolina Mendes Galante Merca (UCPH), Leonardo Víctor de Knegt (UCPH)

Lead participant SVA

Delivery date V1: 05 July 2023
V2: 16 April 2024
V3: 02 January 2025

Dissemination level Public

Type Report

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No. 101000494.

Revision History

Author Name (Partner short name)	Description	Date
Jenny Frössling, Gema Vidal (SVA), Luis Pedro Carmo (NVI)	Draft deliverable v0.0	18.04.2023
Jenny Frössling, Gema Vidal, Fernanda Dórea, Wiktor Gustafsson (SVA), Luis Pedro Carmo, Victor Henrique Silva de Oliveira, Britt Bang Jensen (NVI), Jade Bokma (UGent), Yara Slegers (UU), Camille Delavenne (AUSVET), Lucía Dieste Perez, Anouk Veldhuis (GD), Sébastien Picault (INRAE), Carla Gomes (AHI), Annette Boerlage, Maria Costa (SRUC), Beatriz García Morante (IRTA), Carolina Mendes Galante Merca, Leonardo Víctor de Knegt (UCPH)	Draft deliverable v0.1	23.05.2023
Jenny Frössling, Gema Vidal, Fernanda Dórea, Wiktor Gustafsson (SVA), Luis Pedro Carmo, Victor Henrique Silva de Oliveira, Britt Bang Jensen (NVI), Jade Bokma (UGent), Yara Slegers (UU), Camille Delavenne (AUSVET), Lucía Dieste Perez, Anouk Veldhuis (GD), Sébastien Picault (INRAE), Carla Gomes (AHI), Annette Boerlage, Maria Costa (SRUC), Beatriz García Morante (IRTA), Carolina Mendes Galante Merca, Leonardo Víctor de Knegt (UCPH)	Draft deliverable v0.2	01.06.2023
Angela Bearth (ETHZ), Jasmeet Kaler (UoN), Miel Hostens (UGent), Mirjam Nielen (UU)	Revision v0.3	05.06.2023
Gerdien van Schaik (UU)	Review of v0.3	28.06.2023
Jenny Frössling, Gema Vidal, Luis Pedro Carmo	Final version (v1.0)	30.06.2023
Johannes Ripperger (accelCH)	Final checks and formatting of v1.0	05.07.2023
Jenny Frössling, Gema Vidal, Luis Pedro Carmo	Final version (v2.0)	16.04.2024
Jenny Frössling, Gema Vidal, Luis Pedro Carmo	Final version (v3.0) with Annex added	19.12.2025
Johannes Ripperger (accelCH)	Final checks and formatting of v3.0	02.01.2026

Contents

Abbreviations

Abbreviation	Description
AMU	Antimicrobial usage
API	Application programming interface
AWS	Amazon Web Services
BRD	Bovine respiratory disease
DAFM	Department of Agriculture Food and Marine *(in Ireland)*
EU	European Union
ICBF	Irish Cattle Breeding Federation
NSCP	National Salmonella Control Programme *(in Ireland)*
PCR	Polymerase chain reaction
PLF	Precision livestock farming
TRL	Technology readiness level
VDL	Veterinary diagnostic laboratory
WP	Work package

Partner short names

Short name	Organisation
UU	Universiteit Utrecht
UCPH	Københavns Universitet
UGent	Universiteit Gent
ETHZ	Eidgenössische Technische Hochschule Zürich
UoN	The University of Nottingham
SVA	Statens Veterinärmedicinska Anstalt
INRAE	Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environemment
NVI	Veterinærinstituttet – Norwegian Veterinary Institute
IRTA	Institut de Recerca i Tecnologia Agroalimentàries
SRUC	Scotland's Rural College
AHI	Animal Health Ireland Initiative
GD	Gezondheidsdienst voor Dieren B.V.
AUSVET	Ausvet Europe
SLW	SLW Biolab Weterynaryjne Laboratorium Diagnostyczne Spolka Cywilna
LELY	Lely Industries NV
accelCH	accelopment Schweiz AG

Executive Summary

This Deliverable is a summary of content and technical requirements of the tools being developed within the scope of the DECIDE project. It also includes separate descriptions of the context and requirements of each prototype tool.

Objectives of the Deliverable

With the help of this deliverable the breadth of content and technical requirements of the tools being developed in DECIDE was identified and organised in a structured manner. The objective is to guide further development and to promote collaboration between tool developers when seeking the most appropriate solutions. This can also be used by other WPs to identify how to link their work to and support the development done in WP3.

Activities

Project participants were engaged in various activities that contributed to the outcomes summarised in this deliverable:

- Monthly meetings between WP3 lead (SVA) and co-lead (NVI) team.
- Monthly meetings with partners involved in the development of tools.
- Meetings with smaller groups (i.e., species) for in-depth discussions to better understand each tool being developed.
- Regular communications with different partners to foster collaboration across species on how to best plan and advance tool development.
- Meetings with other WPs to coordinate efforts and look for synergies.
- Organised workshop (May 2023) to provide overview of different software and approaches to tool development, foster discussion and collaboration between partners, and team building.

Outcomes

- Overview of content and technical requirements of each tool being developed.
- Brief summary of most relevant communalities and dissimilarities regarding content and technical requirements of the tools.

Next steps

- Bridge the identified gaps, in particular regarding IT infrastructure, security and software development.
- Design the evaluation of the usability and usefulness of the different prototypes to be piloted.
- Link to ongoing work in WP1, WP2, WP4, and WP5.

1 Analysis and synthesis

1.1 Input sources for assessment of content and technical requirements

This deliverable is based on several sources and activities. The leadership of WP3 has conducted a series of meetings with the members of the species case studies to gather information on the content and technical aspects of each tool. WP5 also attended these meetings to ensure a close collaboration and alignment between WP3 and WP5. By participating in the meetings, WP5 will be able to use the information to better inform their project activities. A structured approach was followed to collect input using a pre-defined form developed in Excel addressing the following: member state(s), species, user group, number of users in pilot, potential number of users in the future, application/case leader, data owners/data location, type of data, epidemiological unit, temporal unit, frequency of data updates (in tool), data pipeline description, development platform, how users access and use the tool, technology readiness level at different time points (month 25, 37, and end of project), output and functions, and intended action of the user.

Another important source of information was the work conducted in collaboration with and lead by ETHZ and UoN in WP5, particularly the execution of focus groups and surveys aimed at better understanding user needs, which provided relevant input for the app developers. Aligning tool features with the user requirements will be critical for the successful implementation of the apps. For some tools, project members have also had continuous and direct dialogue with potential users, to enable their involvement and make the development of applications iterative.

We have also had monthly meetings with partners involved in the development of tools, and meetings between WP3 lead (SVA) and co-lead (NVI) team, to share experience from both previous and ongoing work on tool development. The expertise of project partners and previous experience has also been discussed and shared through the species data groups.

In May 2023, WP3 organised a workshop to share experience from using different software and approaches to tool development, foster discussion and collaboration between partners, and team building. The workshop also added insights into our understanding of content and technical requirements.

1.2 Overview of case studies and their progress

In total there are currently sixteen tools under development. The majority of those are being developed for cattle and pig species (five and seven, respectively) while two tools are being developed for poultry and salmon species each. Table 1 shows a list of all the tools being developed as well as the other names by which they can be found in other deliverables, if applicable. Tools are at different stages of the development cycle, and those in earlier stages are expected to achieve a lower TRL by the end of the DECIDE project compared to those in more advanced development stage. It should be noted that to meet the objectives of the project, some functionalities may be added to some tools, while the development of others may be discontinued. As per agreement, tools will be made open source when finalised.

Table 1. List of tools under development in the project.

Tool Name	Other names	Species	Current TRL**	Expected TRL**
Barometer		Cattle	4 – 5	6 – 7
Emulsion BRD		Cattle	3	4 – 5
Calf Health Support tool	Ley Calf Management Signals	Cattle	3	4 – 5
Purchase Risk Meter	Purchase Assistant	Cattle	4 – 5	6
FriskKo app	Purchase Assistant	Cattle	4 – 5	6
Poultry Netherlands	Dashboard with e.g., diagnostics footpad lesions, AMU	Poultry	3 – 4	5
Poultry Poland	Dashboard of production and diagnostic data	Poultry	3 – 4	5
Pig Peaks Farm	Dashboard with GD	Pigs	4	6
Pig Peaks Population	Dashboard with GD	Pigs	4	6
Abattoir dashboard	Abattoir SRUC	Pigs	2	4
Early Disease Warning	Early Warning Sensors	Pigs	1	2
Pig Disease Surveillance	DiseaseSurveillance IRTA Lab Diagnostic data map Pig Barometer	Pigs	1	2
Dutch Pig Pilot	Integrated slaughter – diagnostic data tool	Pigs	1	1 – 2
Pig HealthCheck*	PigHealthCheck AHI PigHealthCheck Dashboard	Pigs	2 – 3	5
Salmon Norway	Salmon NVI	Salmon	5	7
Salmon Scotland	Salmon SRUC	Salmon	4	6

* New functions added to existing tool

** Current TRL based on level by June 2023. Expected TRL based on level by the end of the DECIDE project. Technology readiness level (TRL): 1: basic principles of how a tool could aid decision-making processes among intended users; 2: technology concept for a tool is formulated; 3: experimental proof of concept for a tool is developed; 4: technology for a tool is validated in a lab setting; 5: technology validated in relevant environment (e.g., farms, with health associations); 6: technology demonstrated in relevant environment (e.g., farms, with health associations); 7: system prototype demonstration in operational environment; 8: system complete and qualified; 9: actual system proven in operational environment.

1.3 Analysis and summary of content requirements

The content requirements of the tools under development within the scope of DECIDE were assessed from various perspectives. Below we provide a summary from various viewpoints.

1.3.1 Interactivity and purpose of the tool

We explored the use of interactive interfaces, as well as the purpose of the tools being developed. All tools have some level of interactivity and multipage (dashboards with tabbed layouts) with visualisations that relate to different components of the decision-making or to help provide necessary context.

Based on their purposes, the tools can be grouped into decision-making support and exploratory or learning (including benchmarking) (Figure 1). These categories differed in the cognitive load they demand from the users. Cognitive load is the mental effort required by a tool to store and process information in the short-term memory (working memory), which is used for reasoning and decision-making. Cognitive load depends on the number and complexity of the elements that interact with the working memory. High cognitive load can hinder thinking and learning processes.

- Decision-making support tools require minimal user input and return a fixed output. Examples of such tools are the Purchase Risk Meter, FriskKo app, and the Calf Health Support tool. These tools demand low cognitive load from the user (e.g., cognition, time, mindfulness given to the tool) and can be used for task-oriented activities such as whether a farmer should purchase animals from a given farm, or which calves should be sampled given certain clinical signs. Emulsion BRD is also a decision-making support tool, designed to help choose and evaluate potential interventions. It requires higher cognitive load and higher levels of user input because the type of input for the model must be selected by the user, and changing the parameters of the model requires changing the code in the software.
- Exploratory or learning tools do not require user input but allow users to use slicers and filters, drilling up and down the levels of data hierarchy (i.e., pathogens, age groups, time windows, etc.). Examples of such tools are Dutch pig pilot, Poultry Netherlands, Barometer, Pig Disease Surveillance (IRTA), Poultry Poland, and Pig Health Check (AHI). These tools are expected to require higher cognitive load by the user and have exploratory and communication purposes. These tools foresee the inclusion of early warning algorithms and/or benchmarking to support decision-making. However, not all those tools will be able to implement such functionality during the DECIDE project. Those that will not be able to implement early warning algorithms will still be able to create the foundation for their implementation in the future.
- Mix purpose tools with an exploratory - learning and a decision-support components. These tools include benchmarking or early warning functionalities. Examples of such tools are Salmon Norway (NVI), Salmon Scotland (SRUC), Pig Peaks Farm (SVA), Pig Peaks Population (SVA), and Abattoir (SRUC).

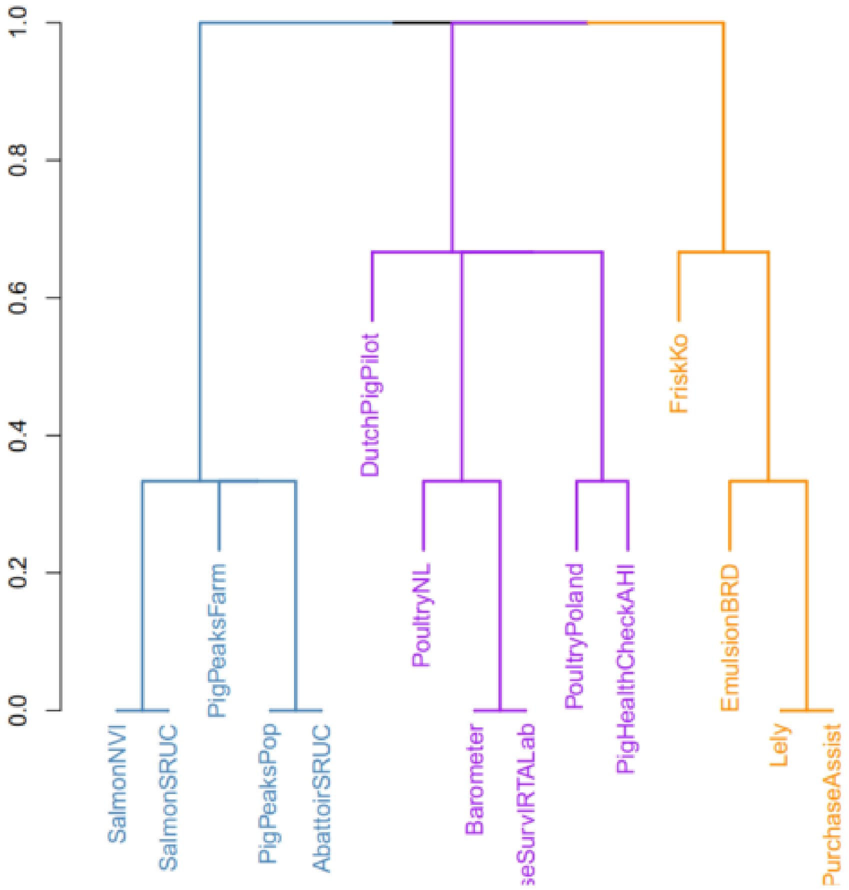

Figure 1. Hierarchical cluster dendrogram of tools being developed in the DECIDE project. Clustering has been made based on purpose of the tool (decision-making, exploratory, or mixed), intended audience (public, organizations, individuals), and data semantics (generation of alerts, benchmarking, or mixed). Blue cluster are tools with mixed purpose (exploratory and decision-support); purple cluster are tools that are exploratory; orange cluster are tools to support decision-making.

A hierarchical cluster analysis was performed to find similarities across the different tools. Results revealed that tools under development for different animal species may be more similar than those within the same species group (Figure 1). This suggests that developers, who have been working in species-specific groups, could benefit from reaching out to those working on similar tools for other species and/or countries.

1.3.2 Target audience/user

We also investigated the target audience of the tools:

- Open to the general public: Barometer, Pigs Disease Surveillance (IRTA), Salmon Norway, and Poultry Netherlands.
- For organisations: FriskKo app for a dairy farmers' association, Purchase Risk Meter for an association of cattle farmers, PigPeaks Population for an association of pig producers.
- For individual farmers: PigPeaks Farm for individual pig farms, Emulsion BRD, Calf Health Support tool, Poultry Netherlands, Poultry Poland, containing information that is not shared except with trusted individuals such as veterinary advisors.

Most of the tools being developed will have a combination of target audiences, requiring control over the type of information shared depending on the type of user, involving user authentication. For example, some tools (i.e., Barometer, Pig Disease Surveillance IRTA) are expected to provide higher level aggregated data open to the public while upon user authentication, farm-level data could be explored by other users such as diagnostic laboratories. Other tools will provide information across multiple farms to an organisation (i.e., Pig Health Check – AHI, Salmon Norway – NVI) while individual farmers would have a restrictive access to information from their own farm.

1.3.3 Recommendations for the next steps of tool development

- The DECIDE project involves different tools with varying levels of complexity and development. The degree of involvement of intended users in the co-creating process is also diverse among the tools, affecting the planning and piloting of the tools. Simpler tools that are co-created with users tend to have clear and realistic development plans and can be tested earlier in the project.
- Tools with more complex visualisations (multiple tabs, multiple plots or tables per tab) at the initial developmental phase would need to foster fluid communication with intended users to avoid the development of overly complicated dashboards with limited usefulness that also require a high level of cognitive load from the users.
- Tools with more complex visualisations already being used (i.e., Pig Health Check AHI, Salmon Norway) need to get user feedback to meet user's requirements in terms of the contents to move forward to the next phases of development.
- In further development of the tools, connection to other WPs should be investigated so their different perspectives can be included in the tool development (early disease prediction models, tool interoperability, disease burden, and user needs).
- Tools that are being developed without having the needs of the target user in mind may not be used by the intended user beyond the DECIDE project (e.g., due to low (perceived) benefit, high cognitive load).
- In the context of the European Union, tools should be made generalisable between countries or species to show that the data structure and methodology of the tool (e.g., the ontologies, models, economics) are applicable in similar farming systems.

1.4 Analysis and summary of technical requirements

In the whole data workflow, there are different steps with different technical requirements:

1.4.1 Technical requirement for data sharing or data transfer.

- Data is being shared through various methods depending on the size, format, and sensitivity of the data. Current methods are emails with attachments (.xlsx, .csv files), queries to databases (SQL database, Oracle database) using existing in-house applications (e.g., Rapportportalen at SVA) or APIs that generate files of different formats (.csv, .RMD), manual data entry, and text files (.doc, .pdf). For large databases, distributed queries may be required (to be confirmed). The data-sharing methods should comply with the relevant data protection and privacy regulations.

1.4.2 Technical requirements for data cleaning and data wrangling to prepare data for visualisation.

- R software and Python software are being used as stand-alone software or in combination. Frameworks used by different partners are PySpark, Databricks, Azure, and AWS.
- Cloud services for data storage such as Azure and AWS are being used.
- Version control (git) is used in some cases, integrated with platforms such as GitHub and Azure.

1.4.3 Technical requirements for data visualisation and tool deployment

- For data visualisation the R packages Shiny and FlexDashboard, Tableau, or a combination of different software (JavaScript, Python) are being used.
- Multiple programming languages are used (R, Python, JavaScript, HTML) with optional .css files for styling.
- Cloud service may not always be required during development but will be required later to distribute the tool to the end users.

1.4.4 Technical requirements for tool accessibility

- Most tools will be accessible as webpage, and only a few ones are expected to be accessible in local machines only.
- Most tools will require different levels of access for different types of users, requiring user authentication.
- Based on our insights from cattle farmer focus groups studies in WP5, farms need to have appropriate infrastructure with reference to Wi-Fi and connectivity as this can create inequality in access and has potential to widen and/or create digital divide on farms.

2 Prototype descriptions - cattle

2.1 Barometer

2.1.1 Context

Bovine respiratory disease (BRD) is one of the most important diseases in cattle, resulting in considerable economic losses, hampered animal welfare, and high antimicrobial use. Awareness of circulating BRD pathogens in a given region combined with early warning systems could support pathogen-oriented decision making, such as preventive measures (*e.g.*, vaccination), enhanced biosecurity, and targeted treatment. Therefore, the idea of the European Veterinary Barometer for BRD was raised by UGent. This tool offers the user a visualisation of pathogen-specific laboratory results of respiratory tract samples obtained from all cattle production systems over Europe together with geolocation. Next to the descriptives of respiratory pathogens in Europe, another aim is to create awareness about the importance of taking samples.

2.1.2 User category/categories

The European barometer will be open access; therefore, users can be of all kinds (veterinarians, laboratories, farmers, technical advisors, etc.). When the barometer is used by specific laboratories with their own data, they could decide to have different levels of access, depending on the sensitivity of the data.

2.1.3 Data source and data pipeline

The data owners are the collaborating laboratories from Belgium (DGZ, ARSIA, PathoSense), the Netherlands (Royal GD), and Ireland (Department of Agriculture, Food and the Marine). Their data consists of laboratory results (pathogen identification of respiratory tract samples), geographical data, production type, and if available, herd characteristics at the discretion of the data provider. These data have been saved in .xlsx and .csv files and sent to UGent. At UGent the data from the different laboratories are being processed (cleaned and formatted) using R and Python software to create RDM files which will be combined into one file. The RDM file is then implemented in Tableau for visualisation. Ontologies are being created in collaboration with WP1 to facilitate data interoperability. The dashboard is shown on a secured Azure website which will become open access in summer 2023 and will also be connected to the DECIDE website.

2.1.4 User interface

Currently, Tableau is being used for the interface to gather feedback during tool development while it is not yet available to the intended users. Once the website is launched, the European barometer will be openly accessible to the users. If laboratories decide to use the barometer for their own website, it is possible that login requirements are added. The frequency of updates is yet to be determined and will be in consultation with the laboratories. The ultimate goal is to have a continuous updating system.

2.1.5 Content and functions

Figure 2 shows a screenshot of the dashboard. The default settings show an overview of all the identified pathogens associated with BRD from all laboratories (A). Per individual pathogens, a specific-pathogen distribution over time (month/year) can be visualised (B). This is also possible per geolocation (C), per used sample type (swabs, bronchoalveolar lavages, necropsy, etc.) (D), per used diagnostic test (PCR, culture or nanopore sequencing) (E), and the production type from which the sample was obtained (beef, dairy, veal, mixed, etc.) (F). It is possible to change between count and proportion of positive samples for all measurements (G) or per country (H). In addition, one can filter based on interest (e.g., province, diagnostic sample, production type). When hoovering over the information button (I) more information on the specific pathogen will be shown (e.g., clinical relevance, incubation period, recommended diagnostic sample/test, infection route). The logos of the collaborating companies are shown (J), and when clicking on the logo one will be directed to the company website.

Veterinarians and farmers can use the information for on-farm decision-making to control BRD (vaccination, purchase support, etc.).

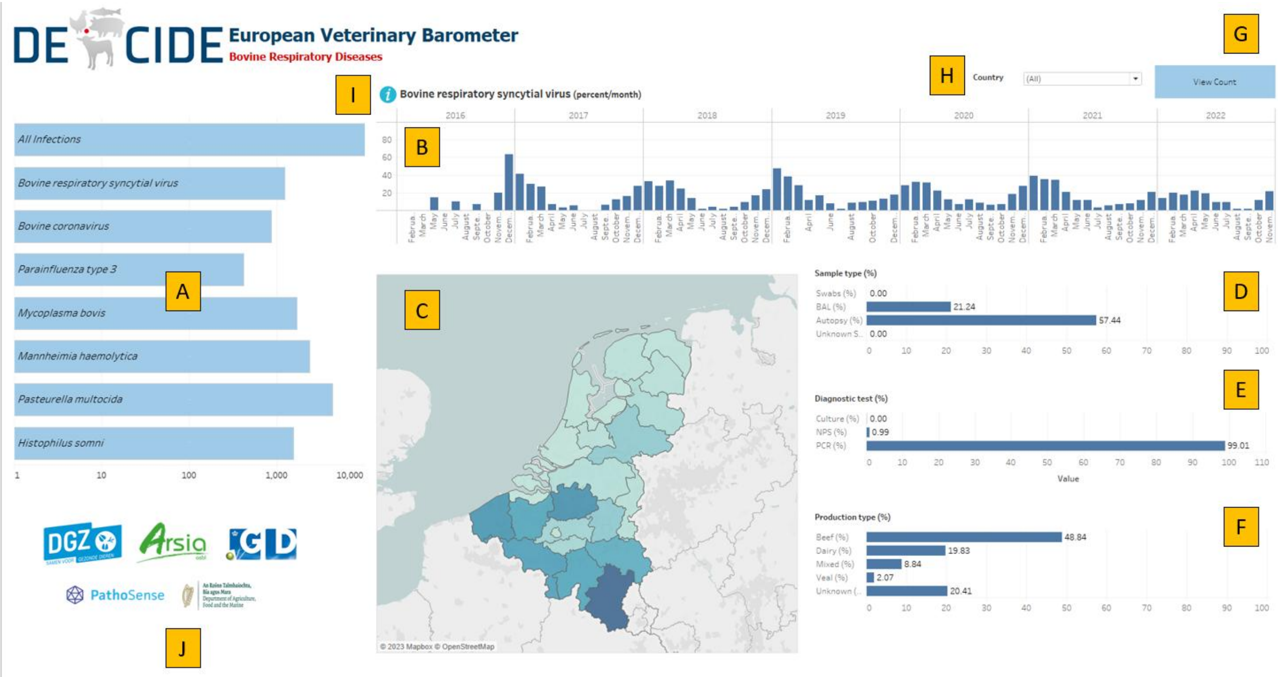

Figure 2. Screenshot of the European Veterinary Barometer, developed by UGent. A) The different pathogens involved, B) distribution of specific pathogens over time (month and year), C) per province, D) per sample type, E) per diagnostic test, and F) the production type. Results can be shown as either count or proportion of positive samples (G), and per country (H). The information button (I) provides more information on the specific pathogen. Logos of collaborating companies are shown (J).

2.1.6 Legal requirements

Given that the data owners are the laboratories, a data transfer agreement should be signed between the lab and UGent. An exception so far is the GD, due to their involvement in the DECIDE project. All data is distributed anonymously.

2.2 Emulsion BRD

2.2.1 Context

Bovine Respiratory Disease is a major health issue in young beef cattle, especially in fattening farms where they affect a large part of weaned calves in the first weeks after arrival. The diversity of pathogens involved (bacteria and viruses), leading to various clinical signs (mild or severe), makes BRD difficult to prevent, detect and treat, often leading to a high antimicrobial usage. A mechanistic BRD model is currently developed at INRAE to account for three different pathogens, risk levels and between-pens/between-building transmission. We intend to provide a web tool prototype based on this model to help farmers or veterinarians to better anticipate the risk of BRD outbreak in their farms in contrasted scenarios, and to compare and rank intervention strategies.

2.2.2 User category/categories

Veterinarians, cattle farmers, farm advisors.

2.2.3 Data source and data pipeline

The tool uses input from the users (composition of batches in the farm, risk levels, major pathogens, etc.) to describe the specificities of their situation of interest. These inputs are combined with other information such as parameter values involved in the model or predefined detection/intervention strategies, coming from observation data (temperature, clinical data, laboratory data), non-observable values estimated from observation data, existing knowledge, literature, or expert advice/assumptions.

2.2.4 User interface

The model is developed with EMULSION (open-source framework), and the tool is generated through the ATOM software (open-source) from the model and structured text specifications expressing the users' needs — see Figure 3. At this stage we do not anticipate login requirements as users do not have to store any personal information in the tool. The tool can be updated periodically for making features evolve according to user feedback, or to reflect evolutions in the underlying model (regenerating the tool from the model is an automatic process). If users require more specific adaptations of the tool or of the underlying model, they will need to be trained to use ATOM or EMULSION.

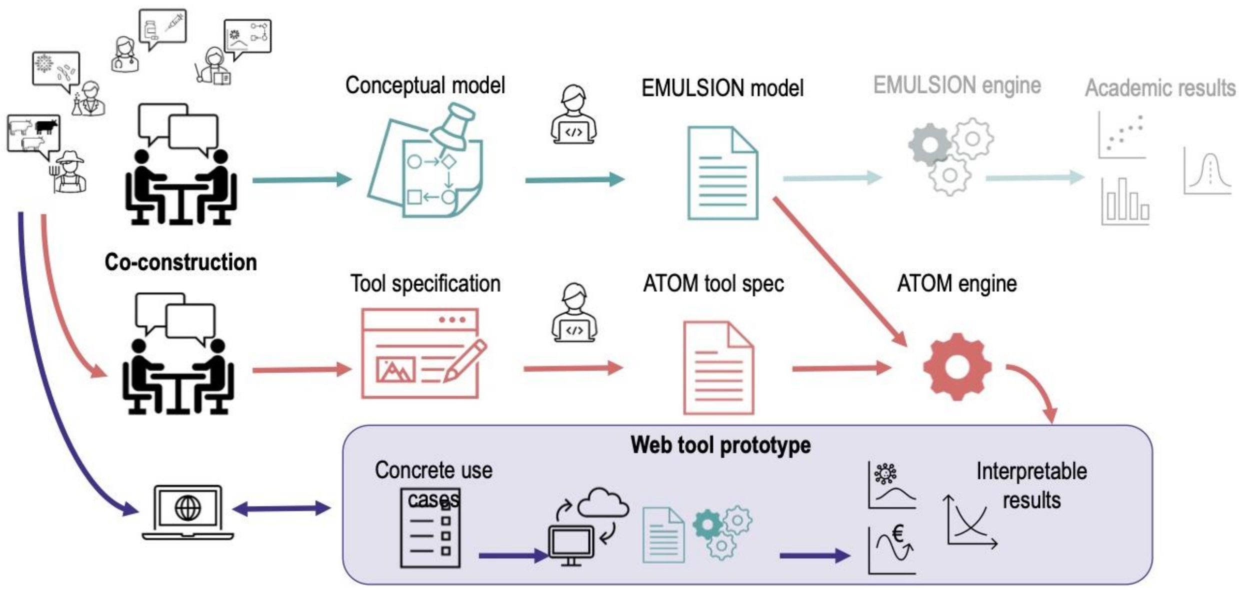

Figure 3. Diagram of the different components involved in the tool developed for the visualization of BRD models.

The tool prototype will be a web application based on Python/Django with EMULSION, Celery, Redis and Rshiny installed on the server. During development it can be tested on an internal server or on a local machine. It will require a dedicated server to make it usable from outside.

2.2.5 Content and functions

The model underlying the tool involves individual animals as epidemiological units and provides representations at different group levels (batch, building, region), with a time scale in half-days. The tool can be updated in real time or periodically (e.g., every 2-3 months) depending on the needs.

The tool will ask the user to define parameter values for adapting to his/her specific concerns, then the intervention strategies to compare. The tool will then run several stochastic repetitions for each scenario, build a synthetic graphical report to inform the user on predicted trends, and provide a ranking of envisaged interventions.

2.2.6 Legal requirements

Given the open-source nature of the source code of the frameworks (EMULSION and ATOM) and of the BRD model and tool specifications, and as we do not foresee the need to store user-specific information at the moment, no special legal requirements are needed.

2.3 Calf Health Support tool

2.3.1 Context

At this moment Lely has four test herds in which they test newly developed sensing techniques. Techniques that show promising results are daily camera visualisation of calves for growth monitoring, audio recording of coughing to monitor respiratory health, and data resulting from the automatic calf feeder to monitor feeding behaviour. The purpose of these sensing techniques is to detect deviations in feeding behaviour and to detect (pre-)clinical signs as a result of respiratory infections such as BRD, *Mycoplasma bovis* or *Mannheimia*. The value of early detection of such deviations increases when the farmer has the tools to identify the causative agent and subsequently, identify relevant measures to prevent further and future infections in the herd. In these farms, data is also available on calves' activity (via neck collar) and pen-level climate.

This study focuses on the combination of feeding and activity data, and climate/air quality data to detect respiratory infections in pre-weaned calves (>21 days old) followed by subsequent advice to farmers to support decision-making in the herd-level treatment of such infections.

2.3.2 User category/categories

Cattle farmers.

2.3.3 Data source and data pipeline

In the tool it is foreseen that an algorithm gives automated signals to the farmer when sensors indicate an aberration in individual feeding behaviour and activity combined with pen-level coughing frequency. Based on the signal, the farmer takes a non-invasive sample from a sample of calves in the pen including the one showing aberrations in the individual indicator and submits it to a veterinary laboratory. The lab runs a predefined set of diagnostics on most prevalent respiratory pathogens and returns the test results to the tool. The tool automatically informs the farmer about the pathogen involved and the most effective intervention measures to be taken to prevent further spread of the pathogen on the short term and recursion of the pathogen on the long term. The added value of conducting a diagnostic evaluation after a sensor alert is generated will be assessed based on the number of times a pathogen is found, the type of pathogen that is found, whether we see recurrence of the same pathogen in different calves from the same herd at the same and at different time points.

Future enhancement of the tool would include a combination of further sensor systems such as camera visualisation of body condition, and pen-level audio recordings (coughing).

2.3.4 User interface

To be decided depending on the results of the pilot.

2.3.5 Content and functions

To be decided depending on the results of the pilot.

2.3.6 Legal requirements

The farmers that are involved in the development of the prototype signed a mandate that they agree to make the data that result from this study available for the DECIDE project.

2.4 Purchase Risk Meter

2.4.1 Context

In Sweden, there is a National Surveillance program for *Salmonella* in place since 1995, when the country joined the European Union. If a herd is found to be *Salmonella* infected, the farmer can be economically compensated. However, to receive such compensation, the herd must meet some requirements, one of which is what is known as the 150-5 rule. This rule states that farmers trading live cattle with other herds should not buy more than 150 animals from 5 different herds for any given year. In 2022, the National Veterinary Institute (SVA) and the Swedish Board of Agriculture (JV) received the mandate from the Swedish government to review the National Surveillance program for *Salmonella*. As part of this review, SVA proposed to substitute the 150-5 rule with an output-based tool known as the "Purchase Risk Meter" to be used among cattle farmers. In this tool, county-level within- and between-herd prevalences estimated from the active surveillance testing are used to estimate the probability of introducing *Salmonella* in a herd when buying cattle, given the number of animals purchased and the origin of the selling herd(s).

Currently, this tool uses data from Sweden to be used on cattle herds. However, it can easily be adapted to other countries and livestock species with similar data on *Salmonella* or other pathogens.

2.4.2 User category/categories

Cattle farmers, farm advisors, trade brokers, the Swedish Board of Agriculture, veterinary advisors (SVA).

2.4.3 Data source and data pipeline

The tool uses input from the users (numbers of animals that they plan to buy, and the regional origin of animals/herds) and combines this with information at SVA on *Salmonella* seroprevalence (National Surveillance Program for *Salmonella* testing). Seroprevalence for each region is hard-coded in the application and updated every time new samples are analysed in the framework of the national surveillance program (every three years, approximately).

2.4.4 User interface

R Shiny application located in an R Shiny server hosted at and maintained by SVA. So far in the development, the server is set up for internal access only. For the application to be used outside SVA, we will need an externally accessible server.

2.4.5 Content and functions

Figure 4 shows a screenshot of the dashboard. The user enters the number of animals purchased from a given herd (**a**) and selects the county of the herd of origin from a drop-down menu containing all the counties in Sweden (**b**). Each purchase is saved as a row in a table that can visualised by the user (**c**), and multiple entries (multiple purchases) are allowed. Alternatively, an .xlsx or .csv file containing the same information (county and number of purchased animals) can be uploaded by the user (**d**), and the table with all the purchases can also be downloaded by the user. The user interface returns the probability (%) of introducing at least one positive animal from at least one of the selling herds (**e**). It also contains a gauge plot where indicates the overall probability as well as an assigned colour (red, yellow, green) corresponding to the acceptable level of risk. The probability and the colours of the gauge plot also match an icon of a smiley face that changes colour according to the colours of the gauge plot (**f**). In the future, other tabs will be added to the app such as maps, interpretation of results, or recommendations on how to reduce the risk of introduction of Salmonella (**g**).

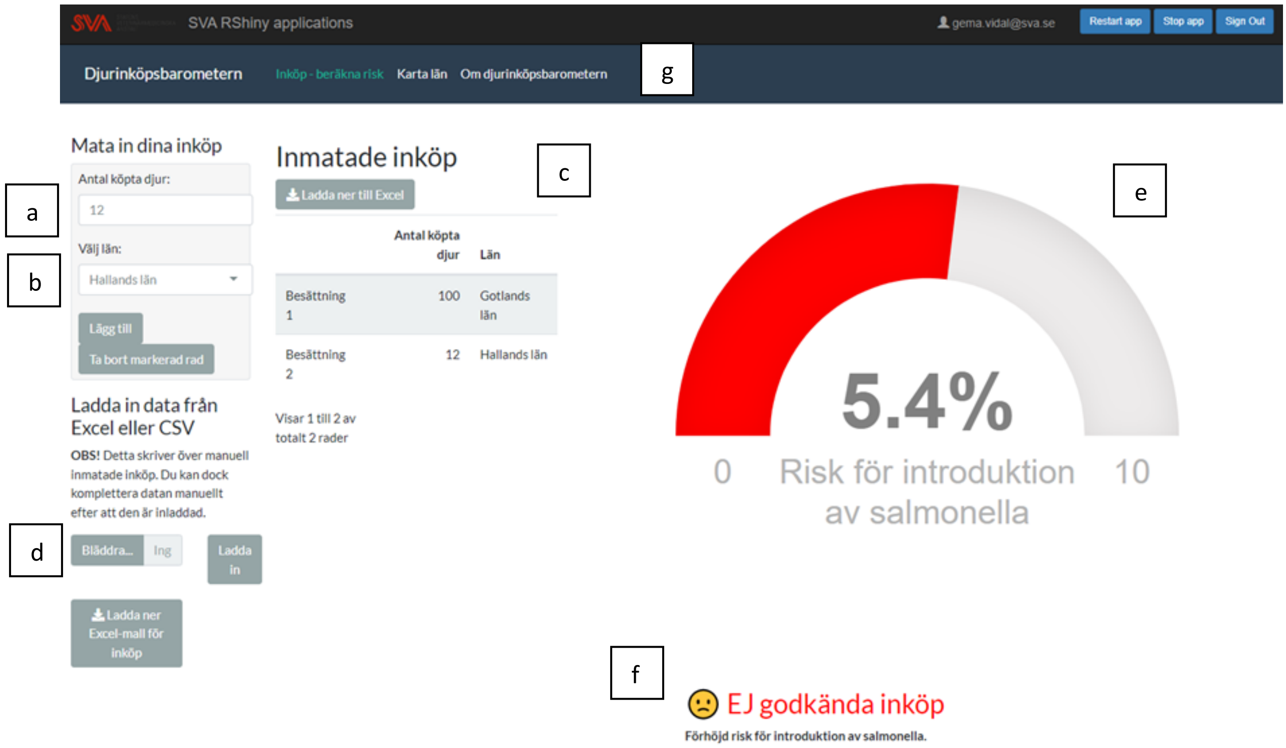

Figure 4. Screenshot of the Purchase Risk Meter tool, developed by SVA (SE). a) Number of animals to be purchased; b) County where the selling herd is located; c) Number of animals and county of origin for each one of the purchases entered by user; d) Upload/download of information entered by the user; e) and f) Visual representations of the output calculated by the tool; g) Other tabs with information to be developed in the future.

2.4.6 Legal requirements

Given the open-source nature of the data, no special legal requirements are needed at the moment.

2.5 FriskKo application

2.5.1 Context

The FriskKo application is a version of the Purchase Risk Meter tool described above. It is being developed to fit the needs of the main Swedish cattle association (Växa) and, specifically, to be developed in the context of their voluntary safe trade program (FriskKo). In such program, bulk tank milk samples from participant farmers are taken quarterly, and analysed at SVA for *Salmonella spp.*, *Mycoplasma bovis*, *Streptococcus agalactiae*, Bovine respiratory syncytial virus (BRSV), and Bovine coronavirus. The safe trade program incorporates participating herds into the Green List according to their last four laboratory results, i.e., those herds that have been tested negative to all pathogens in the last four samples are included in the Green List, whereas those that have had at least one test positive to at least one of the pathogens in the last four samples are not.

Regular meetings are being held between SVA and Växa advisors to develop the tool in an iterative manner. Discussed topics are user requirements, workflow, data pipelines, and hosting solutions for the final tool. Currently, Växa advisors require an interactive tool that can be adapted to include more input data other than bulk tank milk test results (i.e., herd size, trading networks, production system, etc.) and outputs beyond test results (i.e., matching list of farmers with similar health status), with different levels of access depending on the user.

2.5.2 User category/categories

Participant farmers of the FriskKo program, animal brokers, and farm advisors at Växa.

2.5.3 Data source and data pipeline

Växa is the data owner of the laboratory results for each of the participating farms. For tool development during the proof-of-concept phase, Växa shared with SVA a .csv file containing a series of laboratory results from bulk tank milk samples. The toy dataset is being used to develop the codes to clean and organise the data using the R software for statistical computing. The resulting datasets are being used by the tool that is being developed using the R Shiny package.

Long-term solutions for data pipeline are yet to be identified, and different software solutions and IT infrastructure will be identified depending on user's technical requirements.

2.5.4 User interface

During the development of the proof-of-concept, R Shiny is being used. Access to the tool will depend on the identified long-term hosting solution, although some level of user authentication is foreseen at this time.

2.5.5 Content and functions

In the first version of the tool (summer 2023), users will be able to enter the number of animals to be purchased and herd ID from where they wish to purchase the animals (seller farm). If the seller herd is in the list of FriskKo program participant herds, information on whether the seller herd is included in the Green List will be shown in the overview tab. Detailed results for individual pathogens can be seen in separate tabs (one tab per pathogen). Seller herds that either do not participate in the FriskKo program or do participate but are not included in the Green List due to positive tests to at least one of the pathogens, will be shown as "unknown status".

Next versions of the tool (summer 2024) will include the purchasing herd ID as user input and based on the last four tests, the output will be a list of herd IDs with the same health status. User interface will be developed based on regular discussions between SVA and Växa advisors.

2.5.6 Legal requirements

No legal requirements are needed at this stage. However, user authentication requirements are foreseen in the near future, and we are currently discussing the different legal implications for the different application hosting options.

3 Prototype descriptions – poultry

3.1 Prototype 1: The Netherlands

3.1.1 Context

In the Netherlands, poultry farmers are independent farmers and record their own data. Important sources of data are daily farm data, which can be different for each farmer, reports from the slaughterhouse, and national programmes (obligatory or voluntary). The Netherlands has multiple national monitoring programmes. These include registration of all antibiotic subscription, registration of transport of broiler flocks, scoring of footpad lesions at slaughter, and determining the vaccination status against Newcastle disease (of which the Netherlands is free) by hemagglutination inhibition assay. These substantial datasets are analysed and reported separately, but they have not yet been combined. Therefore, the purpose of our tool is to create a more complete view of broiler production in the Netherlands. This includes monitoring broiler health through time and elucidating associations between indicators of broiler health, performance, and antibiotic use. The first task of the dashboard is to visualise the broiler data on national level. Additionally, the dashboard could visualise farm-specific data, including benchmarking.

3.1.2 User category/categories

The tool is designed for poultry farmers and veterinarians, but also for animal health agencies.

3.1.3 Data source and data pipeline

Table 2. Datasets linked in the tool (Prototype 1: The Netherlands).

Data owner	Data controller	Dataset description	Type of dataset
AVINED	Royal GD	Antibiotic subscriptions	Oracle database, used as csv files
		Slaughter data	Csv file*
		Hygiene scan**	?
Royal GD		NDV titers	Excel files
		Other diagnostic lab results	Excel files
Government (public)		Poultry numbers of the Netherlands	Csv file

* An excerpt is sent to us as a csv file
** We currently have permission to use this dataset, but have not received it yet

In the future, this could be expanded with condemnation data from a large slaughterhouse.

Currently, the data pipeline, developed in collaboration with WP1, is as follows:

i. Datasets from AVINED and Royal GD are sent to UU by secured data transfer. The poultry numbers per region are extracted from the CBS (Central Bureau of Statistics) website. The data is stored at UU's Microsoft Azure Storage.
ii. The datasets are cleaned, filtered, and merged in Databricks using PySpark.
iii. The combined dataset is stored as a .csv file at a local computer.

3.1.4 User interface

The Dutch broiler data tool is in a very early stage of development. The first version of the dashboard is created in Tableau and could be published as a webpage. The webpage will either be public or accessible to relevant stakeholders. In the latter case, after logging in, they can see the data aggregated at different levels. If we want the possibility for farmers to see their own data, then an authentication step should be built in.
Currently, the tool is only updated when new data is sent from the data owners. It is not clear yet how often that will be. The ultimate goal is to automate this pipeline, such that the tool is updated regularly.

3.1.5 Content and functions

An early prototype version of the dashboard is shown in Figure 5. The data is on flock level but is aggregated at different levels in the tool. The output is an average or a fraction, e.g., the average footpad lesion scores for a certain year and broiler type. If individual farm level data is included, then this data will be visualised together with national averages for benchmarking. The user can choose different settings (filters) for the visualisations, to give the user more insight in the health status of broiler flocks in the Netherlands.

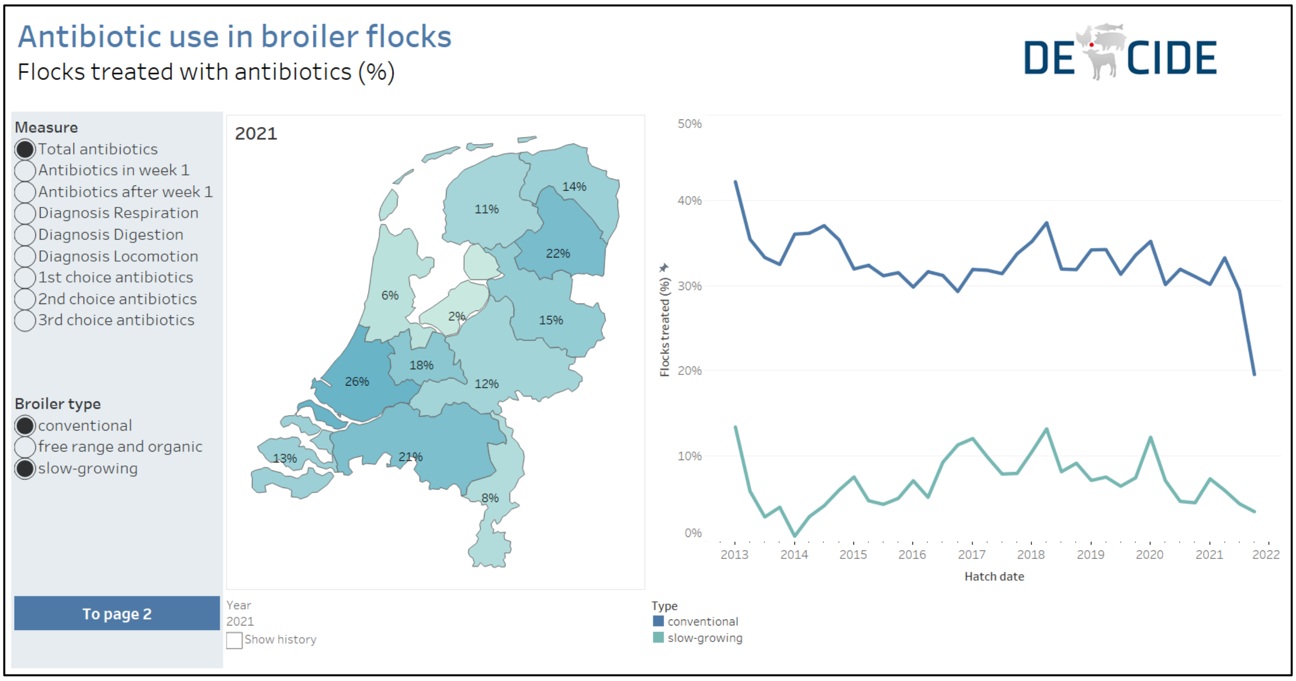

Figure 5. First page of a prototype version of the Dutch poultry dashboard. The user can select which measure is shown and which type of broiler chickens are included. The map can be viewed for each year separately.

3.1.6 Legal requirements

Data sharing contracts are in place between AVINED and UU for use of their data. Data of Royal GD can be used on the basis of an internship contract. This data is privacy sensitive.

The original datasets are linked on flock level. This is not possible with anonymous data, because the farm identification numbers in different datasets must be linked. However, in the data that is used within the tool, all privacy sensitive data is removed. This would be different if individual farm data were included in the tool, and possibilities to do so are still to be explored.

3.2 Prototype 2: Poland

3.2.1 Context

Poland is the biggest producer of broiler chickens in Europe with 80% of its production coming from highly integrated production systems. Despite this important level of integration, data tools allowing for the collection, visualisation and integration of production and health data to make informed decisions are usually not available. Indeed, most of the data are gathered in unstructured digitalised documents such as spreadsheets.

In this context, a first collaboration has begun between SLW Biolab and an integrator to monitor 250 flocks. This program randomly samples five birds at the end of each production cycle and performs necropsy, bacteriology, serology, parasitology, and PCR tests. The data collected through this first project has the potential to show the value of data integration (i.e., production data and laboratory data) and more automated data management solutions. To achieve this objective, the following work plan has been developed and is being implemented:

i. Precisely identify the integrator’s needs in terms of information and their current capabilities in terms of data and data management (tools available, internal capacities, etc.) (Spring 2023) in collaboration with WP5.

ii. Design a first dashboard using the integrated data from the monitoring scheme allowing for the first visualisation of production data and laboratory data together to improve knowledge (Summer-autumn 2023).

3.2.2 User category/categories

The main user will be the integrator and its employees such as veterinarians, producers, etc. The laboratory working in collaboration with the integrator is another potential user.

3.2.3 Data source and data pipeline
There currently isn't any automated pipeline as most of the data collection (production data and prophylaxis information) and management is done manually.

3.2.4 User interface
The user interface will be developed using an R Shiny dashboard. The application will be only available as proof-of-concept and will be only available for the data owner.

3.2.5 Content and functions
All data is recorded at the flock level. The content of the dashboard will allow the integrator to visualise not only the flock production results (number of birds placed, mortality, feed conversion rate, final average body weight, European production efficiency factor, death in transport, condemnation, bird sold, mean daily weight gain) as well as the footpad lesion scores, but also health indicators and compare performances of different flocks. The details of the content will be defined in more detail after the completion of a need and capabilities study planned over spring 2023.

3.2.6 Legal requirements
Data is owned by a Polish integrator and can be used in DECIDE according to an agreement signed between the integrator and SLW Biolab. This data falls under the joint controller agreement. The data does not contain personal data and there are very few identifiable elements concerning the farms (postal code and the number of flocks or birds per flock present on the farm).

4 Prototype descriptions – pigs

4.1 PigPeaks – Farm

4.1.1 Context
The majority of swine farms in Sweden manage their records using the production management software PigVision (Agrosoft®, http://agrosoft.eu/, 2014). These data allow large population coverage and the shortest time lag between a health event and its potential detection. Production records in these farms are generated continuously, not only on a disease event. This offers advantages such as timeliness, but also results in a lack of specificity, representing a challenge to the definition of the events to monitor and what deviations should be characterised as abnormal. These methodological issues had been previously explored in a project that aimed to investigate whether these regularly recorded production data could be used to construct health and performance indicators; and whether these indicators could be subjected to near-real time temporal monitoring (syndromic surveillance) to detect early signals of reproduction failure, which could be caused by undetected diseases or management errors. The system proved capable of providing early detection of unexpected trends when evaluated using simulated outbreaks of porcine respiratory and reproduction syndrome (PRRS). Communication of the results, however, proved difficult, and the work raised new challenges related to the conversion of the outputs of statistical algorithms into decision-supporting, actionable evidence. In the context of the DECIDE project, this prototype development aimed specifically at identifying how the data collected at farm level, plus the information generated after analyses of those data, could be displayed to farms in formats there were usable (easy to understand) and useful (had value in taking action to safeguard the health of animals).

4.1.2 User category/categories
Farmers (Swedish pig farmers) and their health advisors, including veterinarians.

4.1.3 Data source and data pipeline
Data generated on-farm and stored in the software PigVision (Agrosoft®, http://agrosoft.eu/, 2014 is analysed locally, at the same computer where data are recorded, and displayed in a local dashboard.

4.1.4 User interface
The dashboard is planned as a simple HTML page produced and available locally. The user opens the HTML page from a shortcut in their desktop.

4.1.5 Content and functions
The dashboard presents two main types of information, as shown in Figure 6. The first (panel A) is an overview of the herd, helping farmers with general management. In parallel to the technical work of analysing data and programming

the dashboard, focus groups and workshops are being conducted in collaboration with WP5 to understand what farmers and advisors would like to see in the dashboard. The assumption is that making the dashboard useful for farm management will encourage use, and result in better and more timely data recording. The second type of information (panel B) are the weekly time-series constructed for several different indicators of performance and health. These indicators are displayed in multiple tabs (top menu) reflecting the different stages of reproductive performance within a swine farm. Temporal aberration detection algorithms are applied to signal when unexpected increases (for instance an increase in mortality) or decreases (for instance a decrease in insemination success rates) are detected. All indicators are summarised at the individual (animal) level, per week.

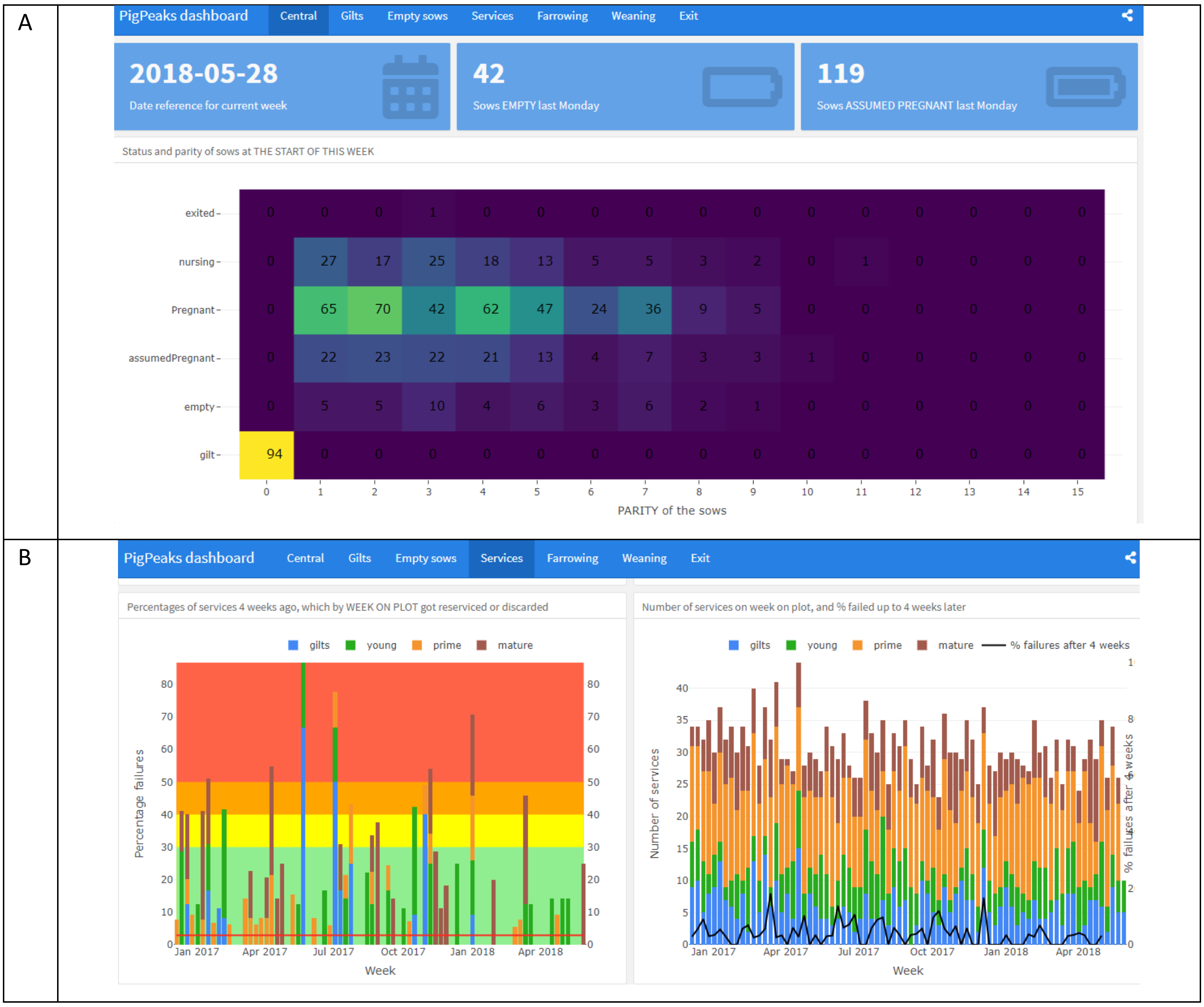

Figure 6. Prototype of a dashboard for monitoring production performance in swine herds in Sweden.

4.1.6 Legal requirements

Data is never removed from the farm. The dashboard is planned to be installed on the same computer used by the owner of the data, and they choose when to open the dashboard to see the results of analyses programmed to happen in the background. As no data sharing is required, no legal requirements apply.

4.2 PigPeaks – Population level

4.2.1 Context

This is a population level version of the prototype presented above (PigPeaks Farms). As explained then, the majority of swine farms in Sweden manage their records using the production management software PigVision (Agrosoft®, http://agrosoft.eu/, 2014). The software business model is based on yearly subscriptions (rather than sold licenses). In

Sweden, Agrovision is represented by a single vendor, Gård & Djurhälsan (G&D; https://www.gardochdjurhalsan.se/). G&D is a stakeholder in the project. As part of their service to farmers, G&D is responsible for the maintenance of backups for all those farmers who wish to send copies of their data to the secure server maintained by G&D. These farmers can choose to give G&D consent to use their data to publish yearly summaries that help farmers benchmark their own production against national performance indicators (https://www.gardochdjurhalsan.se/winpig/medeltal-och-topplistor/). This prototype aims to use the data available at the national (population) level to construct a dashboard of productive performance for sows and slaughter pigs that will fulfil three main objectives: 1) provide farmers with more regular, more detailed and more interactive benchmarking indicators; 2) engage farmers with a data-driven decision supporting system, starting from indicators they are already used to seeing, and allowing them to start tailoring these indicators to their specific decision needs; 3) support disease prevention and control at the population level for other stakeholders, supported by the guidance and needs of farmers who interact with the dashboard.

4.2.2 User category/categories

Farmers, their health advisors, veterinarians, and stakeholders involved in animal health at the population level (surveillance officials for example).

4.2.3 Data source and data pipeline

Swine production data recorded at the farm by farmers using the software PigVision is sent by farmers, voluntarily, to a central secure server at the company providing them with a PigVision subscription service (Gård & Djurhälsan), the dashboard is programmed to use data directly from this source and produce an interactive dashboard with population level indicators available at various temporal and geographical scales.

4.2.4 User interface

R Shiny app located in an SVA's R Shiny server. So far in the development, we have used an internal server. Later stages of development will investigate whether this should be placed in an external server to allow free access by external users or be kept in a restricted environment but providing specific registered access (login) to users.

4.2.5 Content and functions

The specific content will be decided in upcoming workshops coordinated by WP3 in collaboration with WP5.

4.2.6 Legal requirements

This dashboard is supported by G&D legal rights to use the data for statistical purposes as long as farmers can also benefit from the indicators generated. It also assumes that farmers will be the ones choosing to contribute their data to the system.

4.3 Abattoir Dashboards

4.3.1 Context

Scotland has a highly specialised, innovative, and concentrated pig farming sector. About 150 holdings keep 98% of the breeding sows. Farms are concentrated in the Northeast, where 60% of breeding sows and fattening pigs are kept. The pig sector generates around 3% of Scottish agricultural input and around 24,000 tonnes of pig meat are produced in Scotland every year. Processing of these pigs is also highly concentrated in a few abattoirs (the main one processes about 70% of all pigs in Scotland), which makes this step of the production chain ideal to carry out actions to improve animal health and welfare and the quality of the sector.

The purpose of this tool is to transform quarterly static (PDF) reports into interactive dashboards with the aim to provide a continuous monitoring and information on the pig health and welfare indicators available from Scottish pig health monitoring scheme (abattoir data), and to generate alarms when the number of lesions is higher than expected.

4.3.2 User category/categories

We envisage the creation of at least three categories of users: Scottish pig industry representatives, farmers, and veterinarians. Currently and in the near-medium term, only Scottish pig industry representatives will have access to the tool.

4.3.3 Data source and data pipeline

Data are owned by the Scottish pig industry (Scottish Pig Producers, Wholesome Pigs Scotland, and Quality meat Scotland) and collated, curated, managed, and stored by SRUC.

Data corresponds to individual pig checks at slaughter performed by independent assessors with the aim to investigate enzootic pneumonia-like lesions, pleuropneumonia, viral-like pneumonia, pleurisy and lung abscess; milk spot and hepatic scarring, papular dermatitis, pyaemia, tail damage, pericarditis and peritonitis.

In the data pipeline, assessment files (spreadsheets) are sent via email to SRUC after each slaughterhouse assessment day. Files are manually checked, and an R Shiny app is used to process the data and generate graphs for quarterly veterinary reports. Quarterly reports depicting the industry moving average of the last 10 years are plotted using R. Processed data is used to update R Shiny apps with mapping of slaughter lesions in farms all over Scotland. The new tool aims to display these graphs interactively on a dashboard with a time slice/window function.

4.3.4 User interface

The tool will be built using a R Shiny app and will be accessed by each user through a personal account requiring a password. The contents will be updated quarterly.

4.3.5 Content and functions

The objective is to display graphs representing the prevalence of slaughter lesions in finisher pigs processed in Scottish slaughterhouses. The main plot includes a visualisation of the proportion of pigs affected with each lesion out of all pigs assessed in each quarter (industry level), and out of all pigs assessed from each farm quarterly (farm level). Industry level representations will allow the continued monitoring of slaughter lesions prevalence in the Scottish pig industry over time, whereas farm level representations will allow the direct comparison between farm quarterly averages with industry averages (benchmarking). Graphs should be customisable to represent windows of time selected by the user.

4.3.6 Legal requirements

The only legal requirements are an up-to-date data sharing agreement between SRUC and the Data Owners (Scottish Pig Producers, Wholesome Pigs Scotland, and Quality meat Scotland). No data are to be shared with other institutions or partners. Initially, data visualisations will only be shared with the Scottish pig industry (Data owners). Slap marks and location may be shared with them, and access to the tool requires authentication. In a later stage, plots tailored to specific farms may be prepared, but access to these visualisations will require a more developed backend with authentication and user validation for each component of the tool.

4.4 Early Disease Warning

4.4.1 Context

Early detection and diagnosis of disease in pig farming is critical for timely intervention, which potentially improves treatment success and reduces disease impact. In this connection, precision livestock farming (PLF) is potentially one of the most powerful avenues for decision-making based on real-time livestock management. There is precedence for combining different sensor data streams collected in livestock production for detection of or forewarning about disease events. Considering that pig farming is one of the most technologically advanced and continuously growing production sectors, at IRTA we aim at combining and integrating information of different data streams (e.g., sensors, farm management software) to develop models capable of raising early warnings to occurrences of disease in growing pigs (nursery and fattening production phases).

4.4.2 User category/categories

Main users of early warnings are animal care takers (farmers), veterinarians and other health advisors. The use of the R coding for early warnings will be initially restricted to the data owners (providers).

4.4.3 Data source and data pipeline

One data owner is a pig integration company named Vallcompanys Group. This enterprise owns a fattening farm named Farm 5.0 (watch video in Spanish at https://youtu.be/W27tdNt2eSg) that has different sensors implemented, including environmental (temperature, humidity, ammonia, CO_2), water and feed consumption, and cough sensors (SoundTalks®). Through a web service, the IRTA team makes calls for data and data are then downloadable in CSV files. Before data flows into a data repository, it undergoes cleaning and is prepared in R software to ensure appropriate data integration for subsequent modelling.

Retrospective data from five growing pig batches from the Farm 5.0 is being currently used to develop a consistent framework for dynamic monitoring by means of state space models. The final aim is to use dynamic linear models as tools for making forecasts and giving warnings (i.e., distinguish between deviations from the normal patterns of data), so the users can early detect disease events based on data deviations.

A similar approach might be followed for abortion data provided by Gruppo Martini, a second data owner (provider) consisting of a pig integration company from Italy. The overall idea is to model abortions in breeding farms and set up alarms when risk(s) for higher mortality in later production phases (e.g., nursery and fattening) is reached. Thus far, data is in Excel files stored at IRTA repository and analyses are yet to be started.

4.4.4 User interface

An early warning platform development is not foreseen within the DECIDE project. The R coding will be provided to the data owner (Vallcompanys Group/Gruppo Martini), so the model can be further tested and validated in other farm systems.

4.4.5 Content and functions

IRTA will provide and implement R coding for data modelling directly on Vallcompanys/Gruppo Martini data repositories. Communication of warnings from the model application is yet to be discussed and designed.

4.4.6 Legal requirements

In this first phase, no legal requirements are foreseen. The Valorisation Office of IRTA may be involved in later stages to develop an overarching valorisation plan and a strategy to protect the intellectual property (final R coding/algorithms).

4.5 Spanish Pig Disease Surveillance Dashboard (Pig Barometer)

4.5.1 Context

In the last five years Spain bucked the European trend by continuing to expand its pig production. In fact, since 2015 Spain has the largest pig sector in Europe. In this context, the importance of health surveillance cannot be understated, as it safeguards the health and welfare of animals, the safety of food to consumers, and the quality assurance for trade in pigs and pig meat products. However, regular pig endemic disease surveillance is limited, and it is still needed to establish off-the-shelf guidelines for disease surveillance at the national level.

At IRTA, our objective is to develop a swine disease reporting system through a collaborative project among multiple veterinary diagnostic laboratories (VDLs) from Spain. The goal is to aggregate swine diagnostic data and report it in an intuitive format (i.e., web dashboards), describing dynamics of pathogen detection over time, specimen, age group, and geographical area. Sharing information on endemic and emerging diseases affecting the swine population in Spain will assist veterinarians and producers to make informed decisions on disease prevention, detection, and management.

4.5.2 User category/categories

Main users will be VDLs (data providers), veterinarians and other health advisors, farmers, and producers. The Spanish pig disease surveillance dashboard is meant to be public (open access). However, a pilot dashboard will be first co-developed with the participant VDLs, and probably main pig industry actors. Therefore, restricted access (login requirements) will be probably necessary until the final dashboard is launched.

4.5.3 Data source and data pipeline

The Spanish pig disease surveillance dashboard will be developed to share retrospective and prospective surveillance information derived from submissions to participant VDLs from Spain. To ensure anonymisation, data is planned to be provided aggregated by case (submission) on a quarter basis by participant VDLs. A case will be defined as “positive” whenever a sample of a given submission tests positive. If all samples from a submission are negative, the case will be defined as “negative”. Before data flows into a data repository and is visualised, it will undergo data processing (e.g., filtering and aggregations) to ensure appropriate data integration and standardisation. As no data is yet available, data pipeline needs to be further defined.

4.5.4 User interface

The dashboard is meant to be updated on a quarterly basis, adding all diagnoses (positive/negative) allocated since the previous quarter’s update. In the development, we will use an internal (IRTA) server. Dashboard software (e.g., R Shiny, Tableau, Power BI) is yet to be decided.

4.5.5 Content and functions

The dashboard will be used to visualise the surveillance information that is gathered from submissions to the VDL network which includes: unique ID per case, geolocation (province), sample collection date, pathogen, result, sample type, age group, and farm type. The screens will be designed so you have a panel of filters to view diagnoses within an age group, sample type, specific time period, etc. For example, users can select diagnoses of a specific pathogen in nasal swabs from sows in 2019. No prototype dashboard has been yet created; thus, the same functionality as the one described for the Barometer tool needs to be further defined. Examples of existing disease surveillance dashboards from other countries that can guide our pilot dashboard are the Great Britain disease surveillance dashboards (http://apha.defra.gov.uk/vet-gateway/surveillance/scanning/disease-dashboards.htm) and the United States Swine Disease Detection Dashboards (https://fieldepi.research.cvm.iastate.edu/domestic-swine-disease-monitoring-program/).

4.5.6 Legal requirements

A data sharing agreement to be signed by IRTA and the participant VDLs is under preparation. Due to the aggregated nature of the data, the geolocation (province) and farm anonymisation, no special legal requirements are foreseen at the moment.

4.6 Dutch pig pilot: Linking information sources to improve pig health by vets

4.6.1 Context

In the Netherlands there is a lack of adequate tools for swine veterinarians to collect information and to compile data from different sources to properly interpret this information. Therefore, Dutch swine veterinarians are facing a challenge when trying to use available data to take decisions and give advice to their farmers to improve swine health and performance. Veterinarians in the Netherlands get data from the Online Monitoring reports (https://www.gddiergezondheid.nl/Diergezondheid/Monitoring/Hoofdpunten-Monitoring-Varkens), from laboratory and necropsy results coming from Royal GD, from other diagnostic laboratories, and from the slaughterhouse, among others. However, the data is displayed in different formats and at different sites. This poses a challenge to veterinarians, who must invest too much time to put all the data together, if they do it at all. A tool where all the data could be visualised together would help veterinarians to efficiently analyse and use the data for making decisions.

4.6.2 User category/categories

The main users will be swine veterinarians. However, during the project, the possibility of developing different versions for farmers or other swine health advisors will be explored.

4.6.3 Data source and data pipeline

The tool uses input from the Royal GD lab database (laboratory determinations including PCR, serology, and biochemistry determinations). In addition, the addition to other data sources will be also explored. This includes slaughterhouse data from VION and others such as weather data, antibiotic use, or farm data. However, agreements are not in place yet to use those external data.

4.6.4 User interface

User interface still has to be determined. Probably we won't be able to deliver a ready-to-access and ready-to-use tool at the end of the project.

4.6.5 Content and functions

This will depend on the datasets finally used. A sketch in Figure 7 shows how we envision the dashboard with hypothetic Royal GD's data combined with slaughterhouse and weather data. Options to include data entered manually by the users such as lab results from other laboratories and text (e.g., farm visit reports or punctual interventions) are also included in the Figure.

4.6.6 Legal requirements

An agreement with VION is in process to combine a subset of Royal GD and VION datasets as a proof of principles. Confidentiality issues will be solved by anonymising the datasets and by removing all personal data from the datasets. Dataset-specific agreements are needed to also combine data from other external sources such as antibiotic use or farm data. To overcome this challenge, we will probably work with a small subset of farms as a proof of principles in the way that we can ask directly individual owners for approval of using their data for the purpose of the project. Weather data is public, so no special legal requirements are applicable.

Figure 7. Sketches of the potential layout of the dashboard made with toy data. In the first figure three graphs are shown, one for the farm clinical data (DDD, daily antibiotic dose), lab results from Royal GD (e.g., Salmonella, PRRS and Influenza) and data from the slaughterhouse (e.g., lung scores). The second figure also includes data that could be filled in by hand (e.g., farm visit report and lab results from other labs).

4.7 Pig HealthCheck System (Ireland)

4.7.1 Context

The Pig HealthCheck programme is an Animal Health Ireland (AHI)-led programme co-funded by pig producers and the Department of Agriculture Food and Marine (DAFM), with the aim of improving the profitability and sustainability of the Irish pig industry through improved animal health. The programme started in 2019 and currently comprises five key areas:

- Biosecurity: by conducting annual biosecurity assessments on farm.
- Animal welfare: by conducting annual assessments of risk factors for tail biting, based on a tool developed in collaboration with Teagasc, DAFM and AHI.
- Animal health: by capturing, analysing, and reporting of abattoir data from ante- and post-mortem (AM/PM) meat inspection, based on the system being developed by DAFM.
- Antimicrobial usage (AMU): by analysing AMU, using the database created by DAFM for recording AMU by pig farmers.
- Veterinary public health: by reviewing the National Salmonella Control Programme (NSCP), making recommendations to augment the existing programme by providing direction to increase engagement by both farmers and Veterinary Practitioners with the NSCP and to improve the outcomes achieved.

A key element of the programme is the development of an online system where farmers, their veterinarians and other farm advisors can have access to farm data, which are benchmarked with the national data and allow farmers to monitor their status for a range of measures and to compare this with the national profile. Currently this online system receives data from the five areas mentioned above and these data are displayed in several dashboards with three common functions: displaying of the latest results, comparison of those results with the national average, displaying the farm results over time.

Currently, this online tool has been accessed by the farmers of 144 pig units (there are around 350 commercial pig farms in Ireland). To improve the tool, it is important to understand how the tool is being used by farmers and their advisors (veterinarians and other farm advisors) in terms of their usability and usefulness. The lessons learned can then be applied to other countries and dashboards in terms of the functions they should contain. An evaluation study of this tool will be done during the next year.

4.7.2 User category/categories

Pig farmers, their veterinarians, other farm advisors and AHI. Potential in the future for government (DAFM) to have access to some components of the tool.

4.7.3 Data source and data pipeline

The tool uses input from several sources:

- Biosecurity and tail biting risk assessments: data input from users (veterinarians) into the tool. For biosecurity the BiocheckUGent scoring system is used and the tool as an Application Programming Interface (API) with the BiocheckUGent online system to retrieve the scores for each survey.
- Ante mortem data, Salmonella serology data, antimicrobial usage (AMU) data: data collected by DAFM and transferred to the database provider daily or monthly (frequency varies with frequency of update of the original data).

All these data are stored with the database provider - Irish Cattle Breeding federation (ICBF).

The types of data are:

- Scores from 0 to 100 for biosecurity.
- Scores from 0 to 4 per risk category, stocking density, number and type of enrichment items for the risk assessment for tail biting.
- Percentages for several lesions recorded at antemortem.
- Antimicrobial usage per mg/Kg of population in the farm, split by route of administration and type of animals receiving treatment for the AMU data.
- Seroprevalence score for the Salmonella serology data.

4.7.4 User interface

The tool was developed using an Oracle database and mixture of different front-end technologies for the application, including PHP, JavaScript, CSS and HTML.

Users access the tool via an online system: https://pigs.icbf.com/sign-in. They have a username (e.g., herd number) and a password set up by the user. For this a master list provided by DAFM was used with the email address for each farm. For veterinarians and other advisors, AHI curated the list with their email addresses so they can set up their password. Antemortem data are updated daily (end of the day), AMU and Salmonella serology data are updated monthly, while biosecurity and tail biting data are updated annually.

4.7.5 Content and functions

The epi unit is the farm and results are aggregated at farm level.

There are two input surveys in the tool: biosecurity and tail biting risk assessment, where the veterinarians input the data directly into the tool (Figure 8).

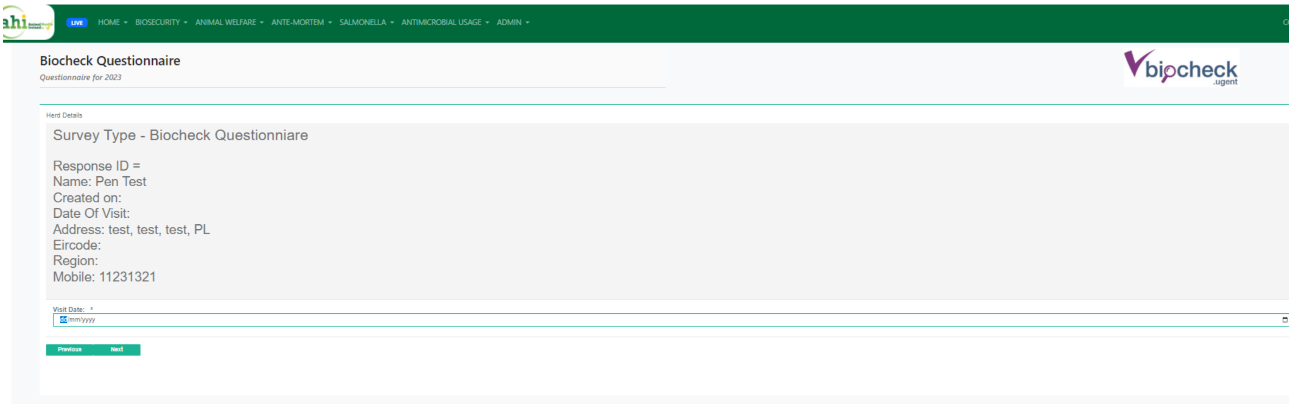

Figure 8. Screenshot of the first page of the biosecurity survey.

The three main functions in all five dashboards are:

- Displaying the latest results (Figure 9)
- Benchmarking the results with national results (Figure 10)
- Displaying results for the farm over time (Figure 11)

Scores | Year on Year | External Biosecurity | Internal Biosecurity | Recommendations

Herd	Description	Last Score	National Avg
	PURCHASE OF BREEDING PIGS, PIGLETS AND SEMEN	100	89
	TRANSPORT OF ANIMALS, REMOVAL OF CARCASSES AND MANURE	95	84
	FEED, WATER AND EQUIPMENT SUPPLY	53	50
	VISITORS AND FARMWORKERS	100	78
	VERMIN AND BIRD CONTROL	90	97
	LOCATION OF THE FARM	100	100
	DISEASE MANAGEMENT	100	73
	FARROWING AND SUCKLING PERIOD	36	57
	NURSERY UNIT	71	64
	FINISHING UNIT	100	100
	MEASURES BETWEEN COMPARTMENTS, WORKING LINES AND USE OF EQUIPMENT	50	67
	CLEANING AND DISINFECTION	65	83

Recommendations:
- Avoid use of CIA antibiotics where effective alternatives are availablea
- Minimise entry of visitors
- Change needles if blunt or heavily contaminated

Figure 9. Screenshot of the biosecurity dashboard showing the latest result for a farm.

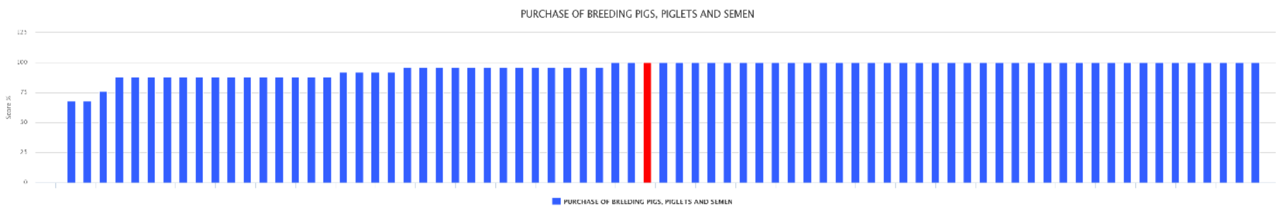

Figure 10. Screenshot of the biosecurity dashboard benchmarking the farm result for Purchase of breeding pigs, piglets, and semen with the other farm results. Red bar shows farm results, while the blue bars are other farms results.

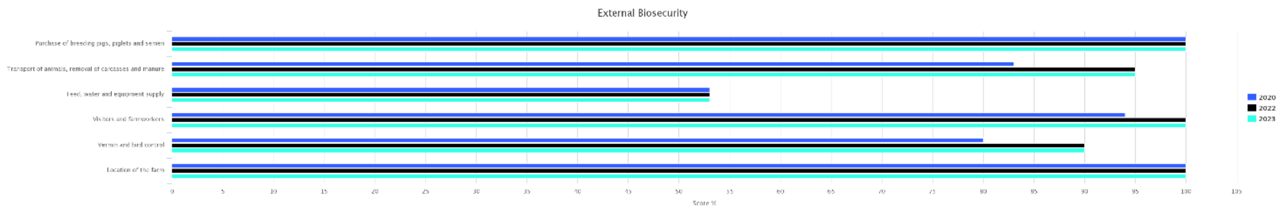

Figure 11. Screenshot of the biosecurity dashboard showing the results for external biosecurity for a given farm over time.

Users can use the dashboards to identify problems at farm level (high AMU, biosecurity areas that are requiring improvement, welfare risks, lesions with high prevalence, etc.) and to follow management changes at farm level (e.g., to verify if the introduction of a vaccine has reduced AMU).

4.7.6 Legal requirements

Currently anyone accessing the tool for the first time needs to sign the Terms and Conditions of usage before being able to see any data. The API for retrieving the biosecurity scores is subject to a license agreement with BiocheckUGent. There are data processing agreements for the data received from DAFM.

5 Prototype descriptions - salmon

5.1 Salmon Norway (SalmonSyS)

5.1.1 Context

Mortality is an important indicator of fish health and welfare in salmon aquaculture. In countries that are major producers of captive salmon, high mortality rates are a significant challenge. In 2022, more than 58 million fish died after being transferred for production in marine farms in Norway, which amounts to more than 3 out of 20 fish dying during the year. A range of factors, including disease outbreaks, environmental events such as algae blooms, and treatments aimed at controlling sea lice infestations, can contribute to the increased mortality of salmonids.

There is a need for improvements in the way salmon mortality information is provided for benchmarking. The presence of multiple data sources and storage formats presents a challenge to the process of data integration, collation, and analysis, which are crucial for distributing information about salmon mortality and its determinants. By continuously monitoring mortality data, valuable insights can be obtained for more effective farm management and disease surveillance.

5.1.2 User category/categories

The decision regarding which group of users to prioritise in the next stage of tool development will be based on discussions among DECIDE partners and findings from the qualitative study conducted in WP5. The current tool prototype has been designed to supply to a diverse range of potential users, including salmon farmers, the private sector (veterinarians, fish biologists, vaccine companies, etc.), and the public sector (the Norwegian food safety authority, the Norwegian directories of fisheries, the Norwegian veterinary institute, other research institutions, etc.).

5.1.3 Data source and data pipeline

The data used in the tool comes from various sources, including farm management systems at aquaculture sites, databases maintained by laboratories and competent authorities. Farmers usually use web platforms for reporting production data to competent authorities. The Norwegian Veterinary Institute (NVI) can query data from sources it does not maintain, or retrieve data in different formats when it is available (open data), e.g., from the Norwegian Directories of Fisheries website: https://www.fiskeridir.no/Akvakultur.

The Norwegian Directories of Fisheries are the data owners of farm registries, which include information such as geolocation and farm license type, stocking month, fish weight, number of stocked fish, and number of dead fish. The Norwegian Food Safety Authority maintains environmental data such as sea surface temperature, treatments data, and clinical records. Disease outbreak data and laboratory data are kept in NVI's sample journal system database.

5.1.4 User interface

Our tool prototype was developed using the flexdashboard package in R, which enables the creation of interactive data visualisations using Rmarkdown language. The prototype is currently available as an HTML document, with interactive features already incorporated. However, if necessary to enhance the tool's reactivity and make the visualisations more dynamic, we could further develop the tool using the R Shiny package, which was suggested by the developers of flexdashboard.

For publishing the tool, NVI has traditionally used an external server and has published various R Shiny apps (see: http://apps.vetinst.no/). However, we are planning to change our publishing solution and acquire products from https://posit.co/, which includes an R Shiny server, to improve our overall hosting capabilities.

5.1.5 Content and functions

Figures 12, 13, and 14 show screenshots of the SalmonSyS dashboard prototype. The dashboard consists of four tabs, each serving a specific purpose.

In the first tab (Figure 12), users are presented with a snapshot of data for a particular month. It includes the number of active farms during that month, as well as farms categorised by their mortality levels: those above 2% (non-baseline levels) and those below 2% (baseline levels). Additionally, a map displays the geographical distribution of farms based on their mortality levels across the Norwegian coastal waters.

The second tab, which is not depicted due to sensitive data, offers further details regarding farms with mortality above 2%. Users can observe the duration of increased mortality levels through a gradient colour palette that represents alert levels. Moreover, an interactive map allows users to click on specific farm markers to investigate potential factors contributing to the high mortality, such as disease outbreaks and the number of sea lice treatments.

The third tab of the dashboard replicates similar functionalities to those in the second tab. There is another map that displays factors associated with mortality, specifically for farms with registered mortality at baseline levels. Additionally, this tab includes a time series graph illustrating the current and historical mortality distribution at both national and regional levels (Figure 13).

The fourth tab provides a time series view, focusing on the expected versus observed mortality for individual farms throughout the production cycle. When the observed mortality surpasses the upper limit of the expected mortality, as determined by a mortality model, it serves as an early warning for potential adverse events in a farm (Figure 14). Currently, our team is engaged in validating this early warning system's performance in detecting outbreaks of pancreas disease, an infectious disease associated with increased mortality and other indirect losses, such as reduced growth in salmon raised in aquaculture sites.

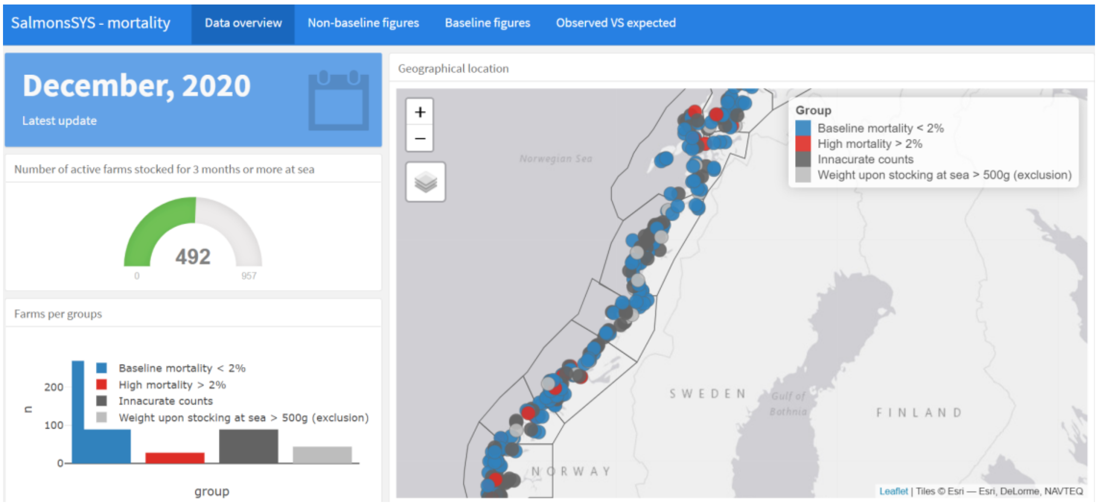

Figure 12. First tab of the SalmonSyS dashboard prototype.

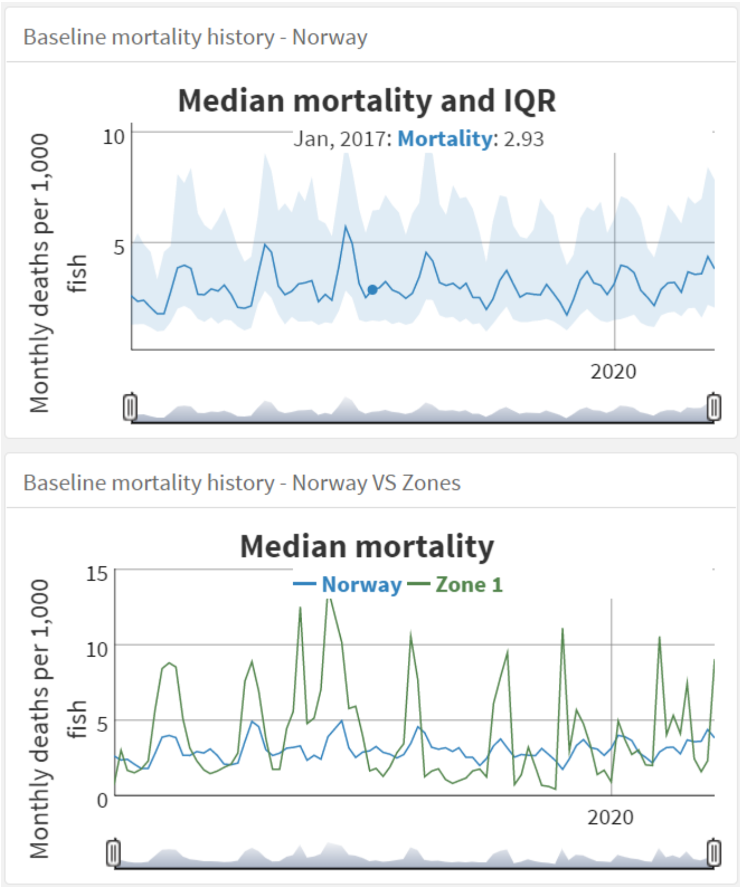

Figure 13. Time series plot of mortality in the third tab of SalmonSyS prototype.

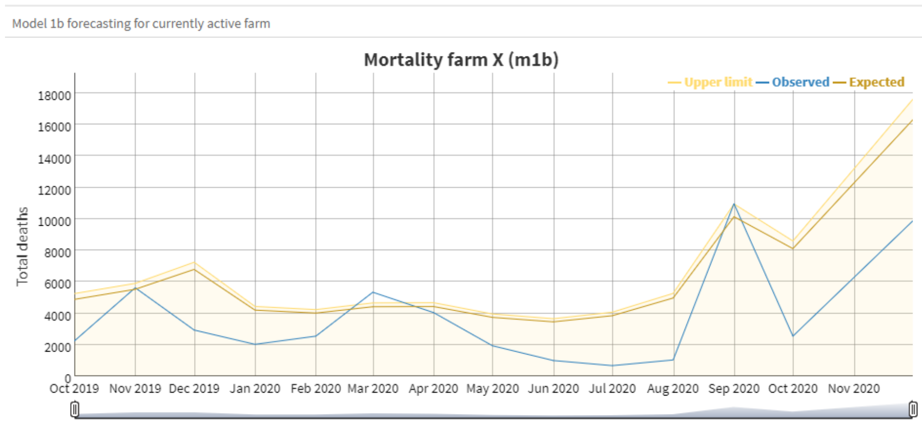

Figure 14. Time series plot for early warning system in the fourth tab of the SalmonSyS prototype.

5.1.6 Legal requirements

Some of the data in the dashboard, particularly production and mortality data, is sensitive and cannot be made publicly available. To accommodate different user requirements and ensure data security, it may be necessary to share data at varying levels of granularity based on the end user's specific needs and permissions.

5.2 Salmon Scotland

5.2.1 Context

Scotland is the third largest farmed salmon producer in the world. Mortality poses a significant constraint on salmon production, and reducing mortality can improve sustainability through improved health and welfare of salmon. Primary causes of mortality can be infectious, such as viruses and parasites, and non-infectious, such as harmful algae blooms and husbandry actions like damage from treatments. Salmon spend the last 1.5 years of their production cycle being constantly exposed to ever-changing environmental conditions in open net pens in the ocean. Living in tolerable environmental conditions is crucial for salmon survival, and extreme environmental conditions are a key concern because of the complex interactions between the salmon host, its environment and infectious or non-infectious agents.

Purpose of the salmon tool will be to monitor monthly salmon mortality and provide alarms for when mortality is higher than expected. We are currently evaluating the possibility of using environmental variables, such as temperature and dissolved oxygen, to improve the predictions of mortality. If these show to improve the model, we plan to add an early warning system for when monthly mortality is observed to be higher than expected according to the environmental factors at that time.

5.2.2 User category/categories

Fish producers, health teams of fish farmers, veterinarians.

5.2.3 Data source and data pipeline

The data tool uses monthly site level production information (mortality, food, total production) collected and reported by Marine Scotland and remote sensed satellite data from Copernicus and NASA sources. This data will be processed using dynamic linear models that will be recalibrated periodically.

5.2.4 User interface

A development plan for the user interface has not been initiated. We will co-develop this tool with the Norwegian tool. Login requirements are anticipated; updates are planned to be monthly.

5.2.5 Content and functions

The plan is that users will be able to zoom into a map of Scotland to their sites and regions, similar to Norway. The tool may also show a table with environmental observations from satellite sources for comparison. Development of this has not started yet.

5.2.6 Legal requirements

Given the open-source nature of the data, no special legal requirements are needed at the moment.

Annex

Table A1. List of requirements for the various tools in D3.1. Greyed out cells indicate that the specific requirement is not yet known for the tool and needs to be decided at later development stages.

			Cattle Barometer (Barometer)	BRD tools (Emulsion BRD)	Calf Health support tool	Purchase risk meter	Purchase Assistant (FriskKo)	Poultry Barometer (Poultry Netherlands)	Broiler virus circulation (Poultry Poland)	PigPeaks Farm	PigPeaks Population	Abattoir Inspect (Abattoir Dashboard)	Early Disease Warning Pigs	Pig Disease Surveillance (Pig Barometer)	Dutch Pig Pilot (Pig Herd Health Dash-board)	Pig Health Check System	Salmon Norway (SalmonSys)	Salmon Scotland (Salmon Mortality Monitor)
1	Data sharing agreement (legal contract)		x		x		x	x	x		x	x		x	x	x	x	
2	Data transfer solution and input																	
	2.1	Electronic files	x				x	x	x		x	x					x	x
	2.2	In-house database		x		x				x		x			x		x	
	2.3	Connection to external database (API)											x			x	x	
	2.4	Direct input into the tool by the user		x		x	x									x		x
3	Data preparation and wrangling																	
	3.1	Software for data preparation	x	x		x	x	x	x	x	x	x	x	x		x	x	x
	3.2	Cloud services for data storage		x				x					x			x		x
	3.3	Version control		x		x	x		x	x	x	x	x	x		x	x	x

Table A1. Continued

			Cattle Barometer (Barometer)	BRD tools (Emulsion BRD)	Calf Health Support tool	Purchase risk meter	Purchase Assistant (FriskKo)	Poultry Barometer (Poultry Netherlands)	Broiler virus circulation (Poultry Poland)	PigPeaks Farm	PigPeaks Population	Abattoir Inspect (Abattoir Dashboard)	Early Disease Warning Pigs	Pig Disease Surveillance (Pig Barometer)	Dutch Pig Pilot (Pig Herd Health Dash-board)	Pig Health Check System	Salmon Norway (SalmonSys)	Salmon Scotland (Salmon Mortality Monitor)
4	Data visualisation and tool development																	
	4.1	Data visualisation software	x	x		x	x	x	x	x	x	x	x	x		x	x	x
	4.2	Software for tool development	x	x		x	x	x	x	x	x	x	x	x		x	x	x
	4.3	Tool server	x	x		x	x				x	x				x	x	
5	Tool accessibility																	
	5.1	Internet connection	x	x		x	x	x			x	x				x	x	x
	5.2	Device (smartphone, laptop, desktop)	x	x		x	x	x	x	x	x	x	x			x	x	x
	5.3	Browser	x	x		x	x	x	x	x	x	x				x	x	x
	5.4	User authentication					x				x	x				x	x	x