

**FABRE TP session at Porto's EAAP 2022:
"Animal farming 4.0: the role of Big Data in animal breeding"**

This year at EAAP in Porto, FABRE TP organised a challenge session, "Animal farming 4.0: the role of Big Data in animal breeding", on the 7th of September. The session aimed to discuss with research, academia, and the private sector the growing role of **Big Data in livestock production and the challenges for breeding companies and research.**

Our chairman, *Johann Sölkner (BOKU)*, opened the challenge session by discussing the relationship of our last year's session on new phenotypes to big data and how Europe is slowly integrating technological sensors that are important, as these **raw data can be powerful** when managed and used collaboratively. The keynote speech of the session was given by *Sjaak Wolfert (WUR)*. Sjaak told in his speech how digitalisation had taken ample space in the whole supply chain, from scanning your food in the supermarket for product origin and ecological impact to sensing equipment on farms. That there is significant potential in the digital transformation of smart agriculture. He further addressed that so much data is generated, which is beneficial in support of sustainable food system development in diverse ways. The key message from this speech was that **"we will see a lot of technological developments and innovations going on. Ultimately, the organisational changes that come along are even more challenging, and it is essential also to consider that."**

Daniel Azevedo (COPA COGECA) could unfortunately not attend in person and therefore prepared a video for the session. He emphasised the importance of data sharing in agriculture. He discussed the digitalisation of the agri-food system and its **opportunity to create value and business opportunities for farmers**. He also explained how it provides farmers with a larger room for manoeuvre to make good farming decisions. Furthermore, he said it is essential to create transparency, define responsibilities and generate trust. The tool to ensure transparency and trust is the [EU code of conduct on agricultural data sharing](#) by contractual arrangement, developed by the agriculture private sector. The key message from Daniel was that **"sharing data is key"**, and the definition of the framework to share and access data; the strategy is the first step to handle before talking about the technology needed.

We were very grateful to hear about the legal aspects of precision livestock data management with *Patrick Majcen (Austria Chamber of Agriculture)*. Patrick closed the first part of the session and talked about how data should be used and the strategies to be adapted based on the data type. He also emphasised, like Daniel, the importance of **trustworthy pooling and sharing of agricultural data between private stakeholders**. Patrick explained the role of the data act in the farming system and explained its position during the paradigm shift from data ownership to data access rights. Patrick also drafted an overview of the current developments at the EU level (Data Governance Act and Data Act)

After a short coffee break, *René van der Linde (ICAR)* gave insight into ICAR and the survey results among their members based on data challenges in animal breeding. René explained the importance of data recording for animal performance evaluation and that data comes from multi-partners in the Agri-chain. She explained the challenges faced with robotic or sensor data and how ICAR helps through the ICAR ADE Standards. They provide a common standard and framework to support their members in easily integrating the data from those different technologies.

This presentation was followed by **Tomás Norton's (KUL)** presentation on computer vision as a powerful phenotyping tool. He showed how it is possible to monitor weight and individual behaviour like eating, drinking, and restlessness. Tomás also explained the difference between traditional machine learning and deep learning computer vision and how to deal with the challenges regarding computer vision technologies. Further, he explained that there is a need for a tailored approach for each species. Still, the common goal remains that a more efficient and affordable process should be a priority in animal monitoring. The key message of this speech is again a message that most of our speakers shared **“Collaboration and openness will enable the community to move faster!”**.

Helene Soyeurt (Université de Liege) could unfortunately not attend in person, she presented online. She gave us an insight into how **machine learning in dairy production bridges the gap between research and the field**. We got a crash course on machine learning and data mining, then dived into applied examples in the dairy production sector. Helene also gave us insight into solutions and future development of machine learning in the animal breeding sector. She finalised her presentation by saying how **better communication** with specialists in data visualisation, scientists and farmers is essential in bridging the gap.

Panel discussion industry and research institute

Moderated by Johann Sölkner (BOKU) & Ana Granados (FABRE TP)

Panelists: Sjaak Wolfert (WUR), Patrick Majcen (Austria chamber of agriculture), René Van de Linde (ICAR), Tomás Norton's (KUL) and Helene Soyeurt (Université de Liege)

The panel discussion was initiated with a question regarding the selling of data based on the sequencing of the sensor data bought from breeding companies which is quite an important topic for the breeding and recording companies. While a photo is a creative act, a sensor photo doesn't fall into that, thus not being covered by intellectual property rights. Therefore, these situations should be evaluated based on the contracts agreed upon by the parties. Even though the germinal products are not the property of the breeding companies or associations, data obtained from these could well be. This needs a final clarification on handling sensor data and **ownership of raw and processed data for recorded animals and their offspring**. Suppose it's not covered by trade secrets or intellectual property rights. In that case, there will be Open Access tool in the process of generating these data. The data act is just a proposal right now and what raw data derived data is, is a huge question.

Another critical topic is the growing concern about the **environmental impact of digital technologies**. We develop a lot of different sensors and analyses and it would be imperative to have a study on what we need exactly because many things can be redundant. For farmers, it's also important to have the best price possible with the best output. A better production must compensate the balance between the price for the analysis and technology in terms of financial output. The idea is that when we have the analysis we want to do more by combining the already existing data, in this case there's no new environmental impact. However, increasing the number of technologies used, especially with ruminants, could have a more significant impact. In particular when the transfer of samples from the farm to the lab is considered. The way we process the data is also something that we should pay attention to. The kind of developments in terms of blockchain and the use of CPU's which process large volumes of data showed us that a huge amount of energy is required. We should also look to the traditional approaches and data already collected and see

what value we can extract from these without additional footprint. **Transparency** is also another concern in this discussion. There should be a limit for the sensitive personal data taking into account privacy.

Data collection standardisation is still challenging, making the exchange of data difficult. Organisations like ICAR are trying to standardise data collection; however it still depends mainly on the local organisations and their relations with manufacturers that produce these sensors to collect data. Another challenge that adds up to this is the different versions of software and devices used, which create large amounts of data. Therefore the aim is to have a format in which manufacturers of the sensor technologies agree to provide the data in this format. This is, in particular, relevant for data collected by milking robots to which there is no access and not possible to standardise. Data definition should also be worked on with the users and manufacturers to move forward. We should have a **common data infrastructure** in the future, and we are already working on that, but that is a huge challenge also in terms of who will invest in this. For data relevant to final applications, there are private investors interested however, we have a lot of more fundamental research where we want to maintain this common data infrastructure. Who's going to invest in that also in the longer term? This is comparable to the system set up for research for publication mechanism. We can share our knowledge with Open Access publications. Still, we should move to a similar structure for data as well. Yet the question remains on who is going to invest in that. Is it the government or the private sector? The European Union is working on this issue by establishing ten common data spaces. At the same time, the European Standard Institute is researching developed common interface standards. A new project will be starting soon in the digital Europe, which will work on the agri-data space. There will be a number of partners working on that with input from the sector.

Local projects are also being implemented, such as the **OneSmart** project from Belgium. This project aims to develop a numerical platform to collect data from different stakeholders. The breeding association manages the platform because it's a farmer cooperative and the idea is that this structure is independent. The idea is that the farmers can just log in to this platform and have everything at the same place, avoiding the need to use several platforms for each data type. This could be the future where the farmers can be a part of such spaces that belong to independent structures.

The overarching message in this discussion is collaboration, but what kind of collaboration are we talking about? We need collaboration from many different levels. Farmers, for example, are generally in a weak position, and they cannot invest themselves in comparison to big companies in a data management platform. You also see some joint data initiatives where they try to collaborate on that level. The manufacturers of these machines also have to collaborate. This also depends on the level of innovation, and if you're working on a market-trending product, then you have to work together. There is also a need for good animal scientists who are interested in bringing insight into that data. The complementarity in terms of the different skill is very important and how we will use the information.

EFFAB and FABRE TP, with their different and complementary expertise areas, could also contribute to this collaboration. In particular, with its role as a knowledge provider in research projects and its position as a sectoral representative for policy making, there are certainly contribution and collaboration areas in achieving data standardisation as well as responsible and sustainable data sharing.

One of the most challenging points is sharing the data with an open mind to collaborate. It can only be successful with the participation of all actors. Many different companies are developing different aspects with very interesting features, but how can we use that in the field? We need good communication, and EFFAB can support that by showing the importance of being together with scientists, private companies and farmers trying to create the link between them. EFFAB and FABRE TP comprise the private sector and knowledge institutes, making it a strong network.

To conclude, it is crucial to include private companies in the research projects and share the benefit through transparency, trust and data interoperability. EFFAB and FABRE TP have a vital role in achieving these goals.