



aWISH
ANIMAL WELFARE INDICATORS AT
THE SLAUGHTERHOUSE



Guidelines to validate sensor output for animal-based measurements of animal welfare

Jarissa Maselyne¹, Anna Stygar², Matti Pastell², Queralt Allueva Molina³, Angela Ramon Pérez³,
Pol Llonch³, Heng-Lun Ko^{3,4}

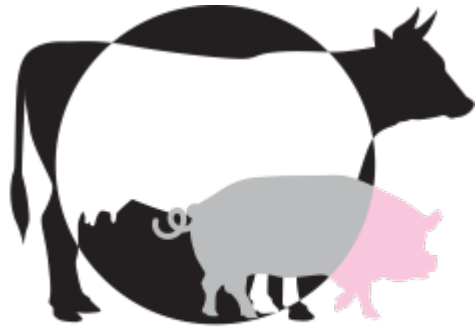
¹ ILVO, ² Luke, ³ Universitat Autònoma de Barcelona, ⁴ University of Murcia
& other consortium partners
awish@ilvo.vlaanderen.be



Funded by
the European Union



Joint initiative



ClearFarm

<https://www.clearfarm.eu/>
H2020, 2019-2024



aWISH
ANIMAL WELFARE INDICATORS AT
THE SLAUGHTERHOUSE

<https://www.awish-project.eu/>
Horizon Europe, 2022-2026

EU requests solutions for large-scale, objective, automated animal welfare monitoring

Sensor technology ?

→ Valid and reliable measurements are needed.

→ Measured data needs to be relevant for animal welfare.

=> Sensor validation is crucial for commercial and research settings



Funded by
the European Union



Guidelines for sensor validation

- Animal-based welfare indicators
- Livestock animals
- Give a general framework, different steps that can be re-iterated
- Use example cases
- Group of scientists from different backgrounds



→ to be used by the aWISH partners for sensor validation

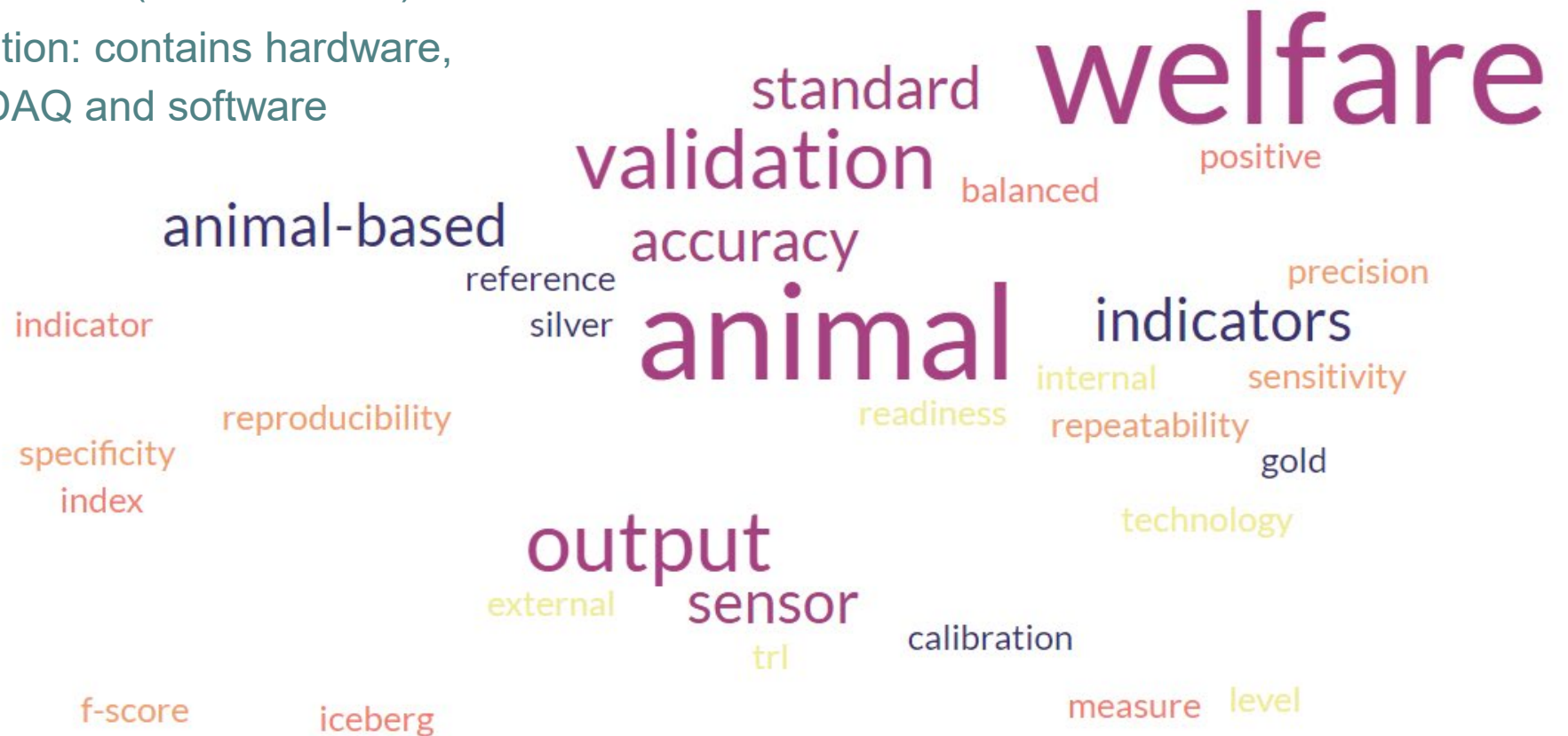
→ to be used by scientists and developers using PLF for animal welfare





Glossary

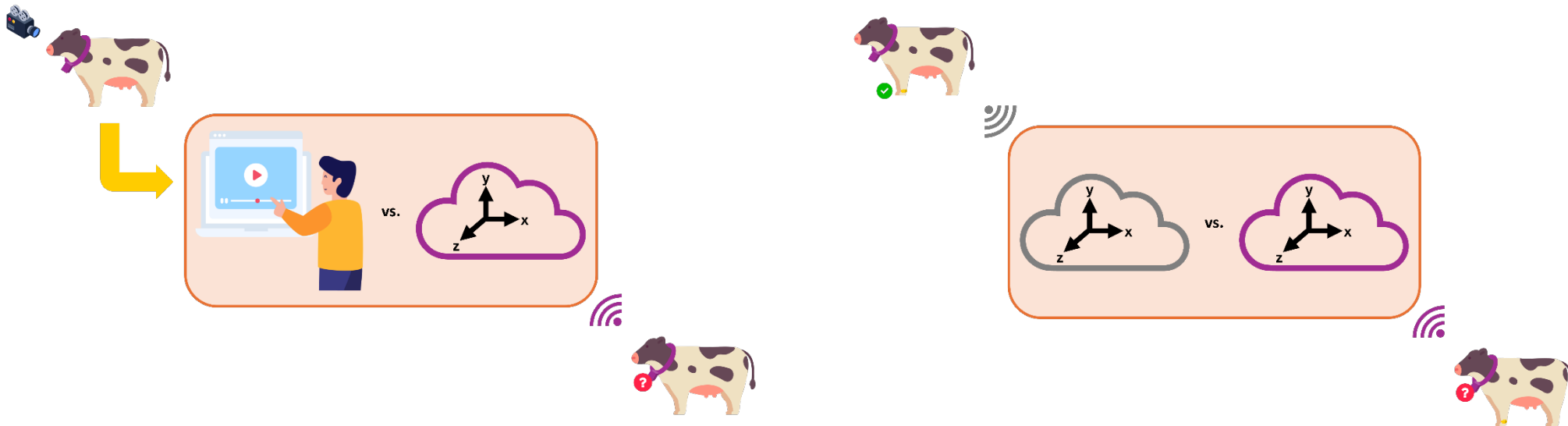
- **Sensor:** Device that produces an output signal for the purpose of detecting a physical phenomenon or stimulus (Fraden, 2010).
- Use broader definition: contains hardware, but also possibly DAQ and software





Glossary

- **Gold standard:** A methodology of measurement that is considered the current standard in the field, and that provides close to error-free data (Gold et al., 2010; Reitsma et al., 2009) to verify the accuracy of the sensor output.
- **Silver standard:** A reliable and sufficiently accurate estimation of the animal-based measurement that presents some limitations when it is used to verify the accuracy of the sensor output but can be used as an alternative to the gold standard.





Need standardized procedure

Guideline to validate sensor output of animal-based measurement and its relevance to animal welfare

Block 1: Output validation (i.e., if the sensor is measuring what it is designed to measure)

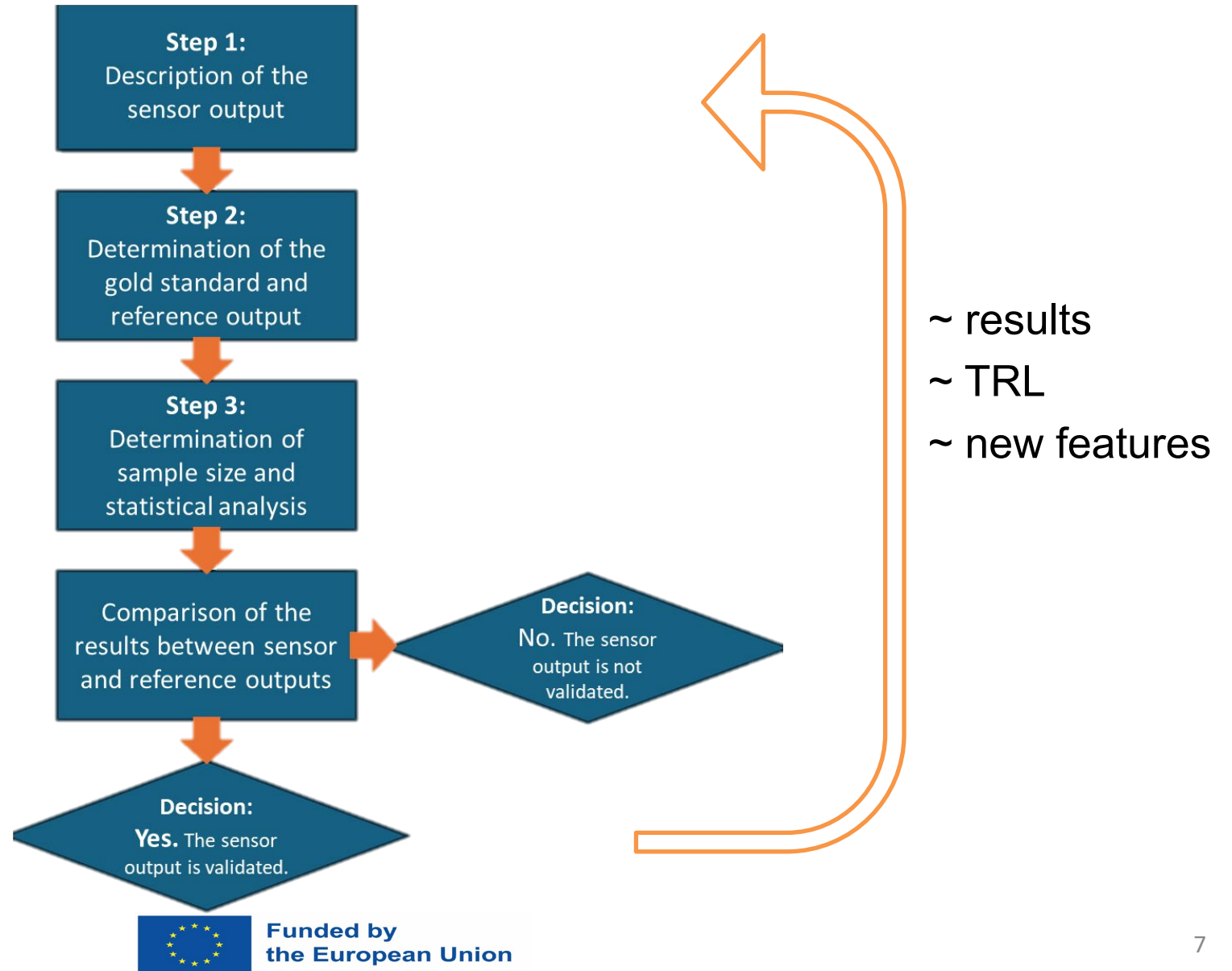
Block 2: Welfare relevance validation (i.e., if the sensor output can be used to measure animal welfare)





Block 1: Output validation

Is the sensor measuring what it is designed to measure?



+ **Step 4.** Limitations of the methodology



Step 1: Description of sensor output

The sensor's output as the **technical characteristics of the sensor** (i.e., type of sensor and set up, data storage, software used, etc.) and the **type of recorded data** (i.e., units, frequency of recording, at individual or group level, etc.) must be described in detail.

- model, brand, company, country
- hardware, firmware, software
- instrument precision, operating conditions, specs
- post-processing!!
- talk to the company





Step 2: Determination of gold standard and reference output

Gold standard or silver standard and how to select it?

Reference output of the biological information (i.e. the animal-based measure) that we want to obtain from the animals

- how accurate is the gold/silver standard?
- what are the limitations of this gold/silver standard?
- human observers: training, inter & intra observer reliability checks
- reference output frequency, level (individual, group?)
- reference output scale: categorical, tVAS, continuous?

You cannot be more accurate than the standard you compare with!?





Step 3: Determination of sample size and statistical analysis

Guidelines for sample size collection

Proper choice and description of statistical analysis

- External validation > internal validation!
- Related to distribution in reality
- Related to the reference output scale!
- Error propagation
- Effect of pre/post-processing, individual differences





Step 4: Limitations and challenges

Technical elements that can block the validation process:

- Confidential algorithms, changes throughout time
- Lack of comparable reference output
- Calibrations needed

Animal and environmental effects to consider:

- Effect of farm conditions and other factors
- Effect of individual animals on sensor output, processing
- Invasiveness? Habituation period – check effects for wearables



Conclusions

Guidelines for sensor validation

Step-by-step process that is linked to TRL levels and the actual application from the beginning

=> Important to know the sensors you are working with

=> Important to know limitations to the sensors and methodologies you are using





aWISH
ANIMAL WELFARE INDICATORS AT
THE SLAUGHTERHOUSE



awish@ilvo.vlaanderen.be

jarissa.maselyne@ilvo.vlaanderen.be



www.awish-project.eu



Example cases

	Computer vision system	Sound analyser	Accelerometer sensor	Sensor fusion: Automatic weighing station (Load cell + RFID)
Pig	1. Tear staining (individual level) 2. Frequency of locomotory play behaviour (group level)	-	-	-
Broiler chicken	3. Incidence of footpad dermatitis (individual level)	4. Number of stress vocalizations (group level)	-	-
Dairy cattle	-	-	5. Identification of different activities (individual level)	6. Body weight (individual level)





Glossary

- **Internal validation:** consists of validating the sensor using the same data set as was used for the development of the sensor (Altmann et al., 2008; Gomez et al., 2021). For example, accelerometer data is collected from one dairy cow herd, and the dataset is split in a training set and validation set for testing algorithms for posture identification. Different cows are present in the validation set, but as they are from the same herd, the sensor is validated only within the tested context.
- **External validation:** Validation of the sensor using a fully independent data set, from different herds not used for system development (Altmann et al., 2008; Stygar et al., 2021). This shows the accuracy of the sensor in contexts different from the development context. For example, accelerometer data is collected from different herds, and an independent test set is used to validate the algorithms, with data from herds that are independent from the training set and were previously 'unseen' by the algorithm.

