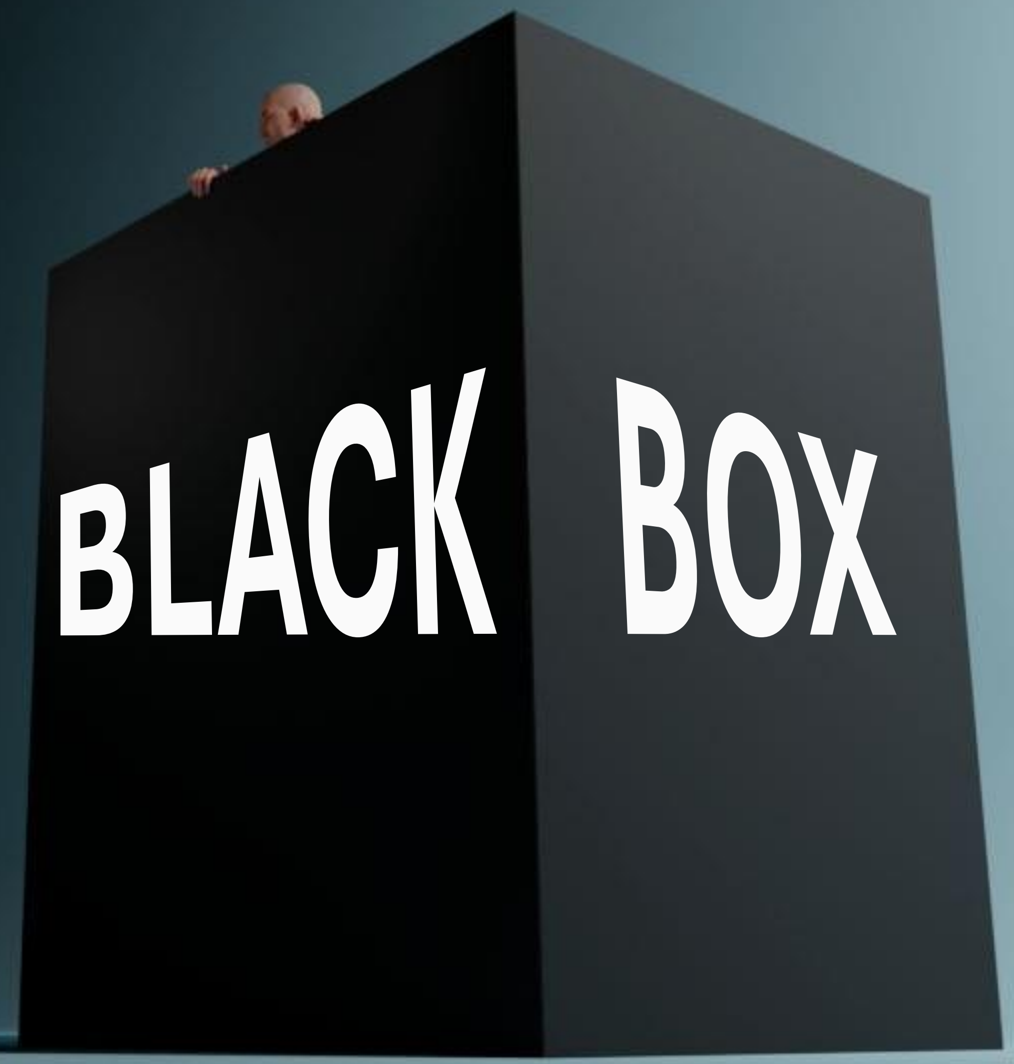
Thinking outside the BLACK BOX



Rethinking the use of Machine Learning for Automated Behaviour Analysis

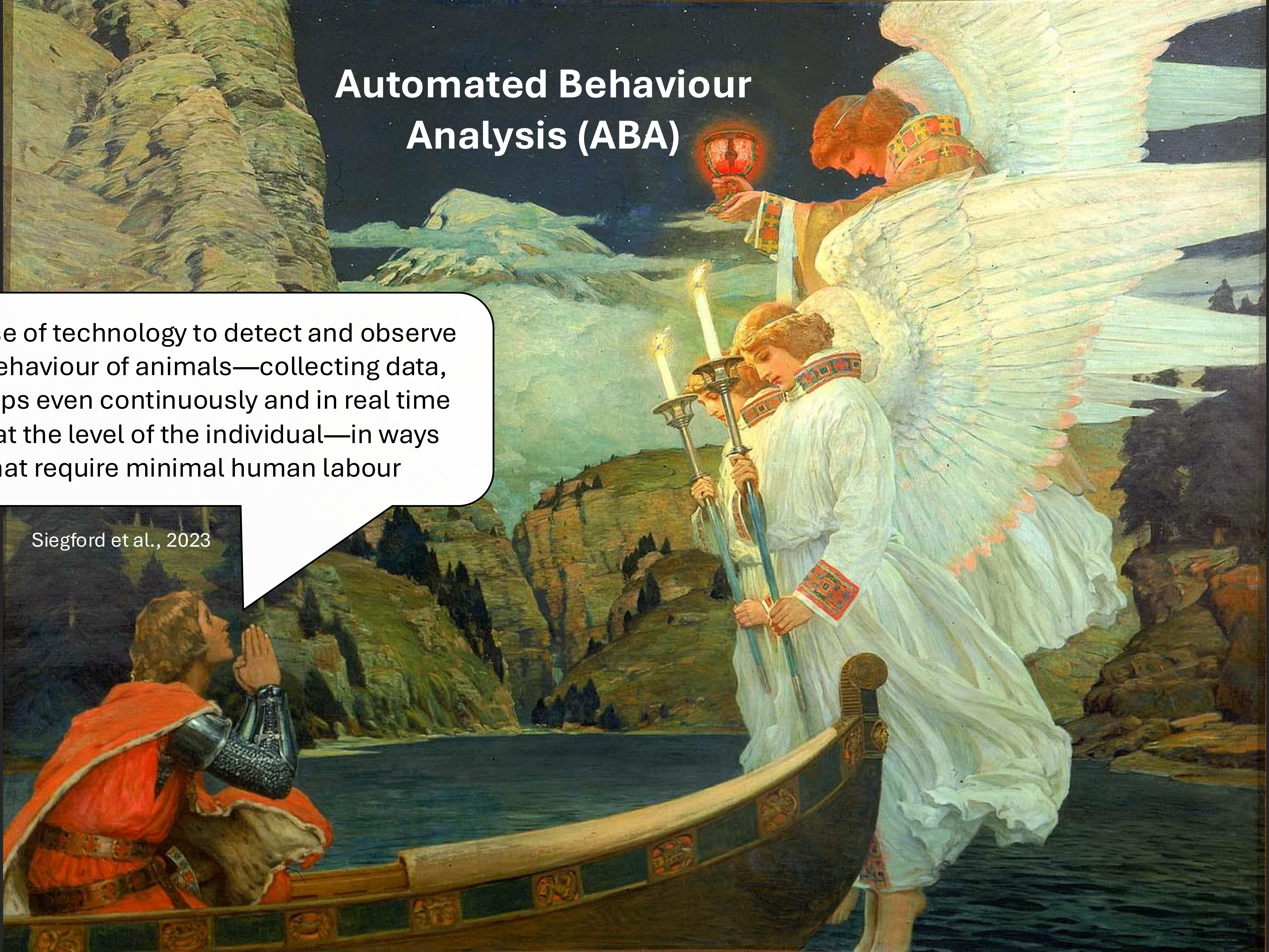
***Stijn Brouwers**, Niall McLaughlin*

June 29, 2026

Artificial Intelligence 4 Animal Science, Ghent

image: filtrabit

Automated Behaviour Analysis (ABA)



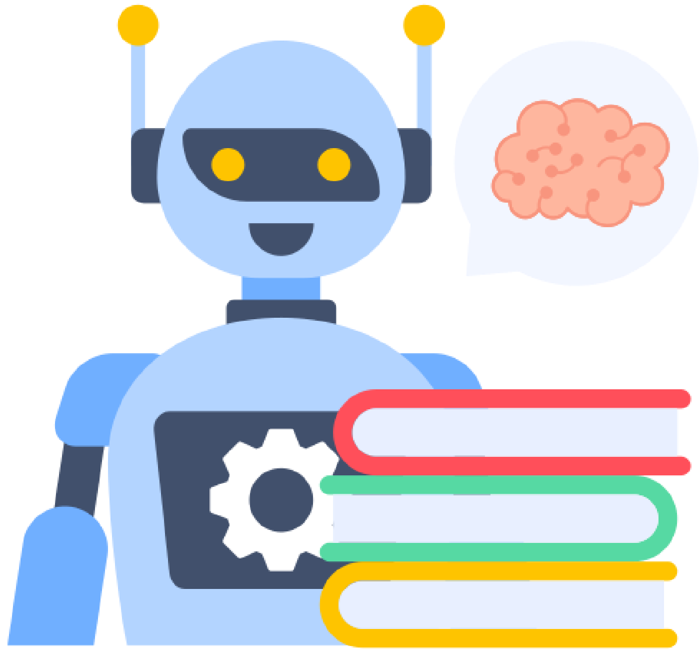
The use of technology to detect and observe the behaviour of animals—collecting data, perhaps even continuously and in real time and at the level of the individual—in ways that require minimal human labour

Siegford et al., 2023

Frederick J. Waugh, 1912

The Technological Revolution

Sensors

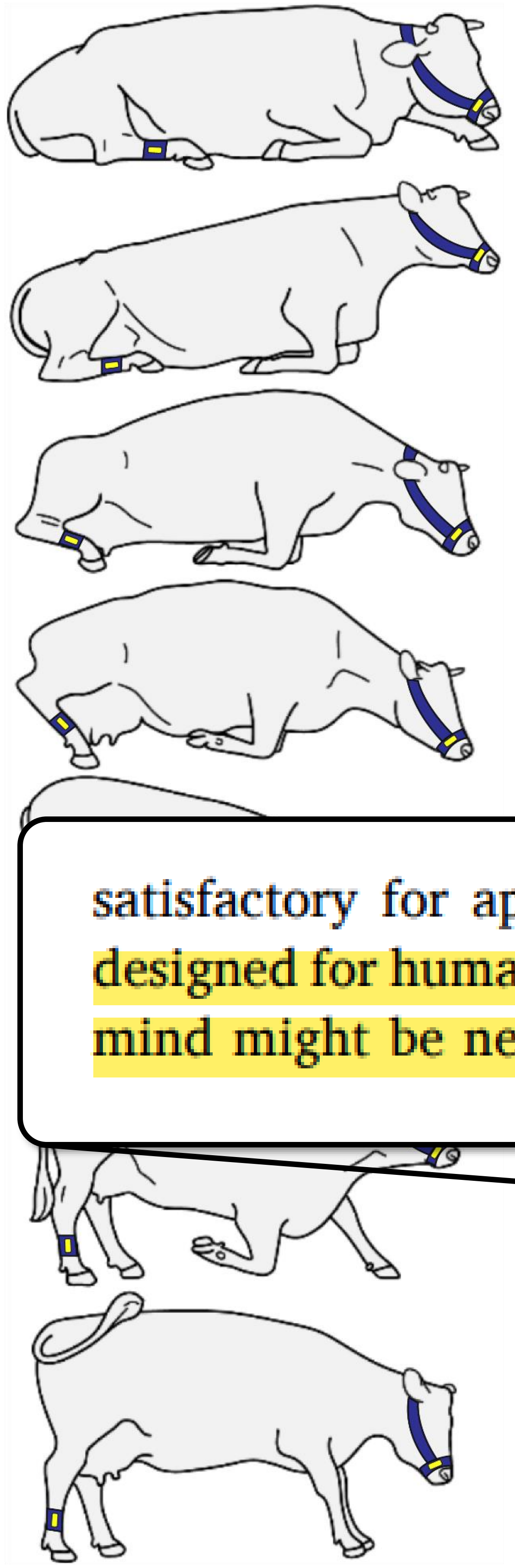


Machine Learning

PRACTICAL IMPACT
REMAINS LIMITED

So why does ABA underperform?

The Ethogram Problem



Towards a novel method for detecting atypical lying down and standing up behaviors in dairy cows using accelerometers and machine learning

Stijn P. Brouwers [a,b,*], Michael Simmler [c], Pascal Savary [a], Madeleine F. Scriba [a]

[a] Centre for Proper Housing of Ruminants and Pigs, Federal Food Safety and Veterinary Office (FSVO), Agroscope, Tänikon 1, 8356 Ettenhausen, Switzerland
[b] Animal Physiology, Institute of Agricultural Sciences, ETH Zurich, Universitätstrasse 2, Zurich 8092, Switzerland
[c] Digital Production, Agroscope, Tänikon 1, 8356 Ettenhausen, Switzerland

ARTICLE INFO

Edited by: Stephen Symons

Keywords:
Animal welfare assessment
Free-stall cubicle
Time series classification

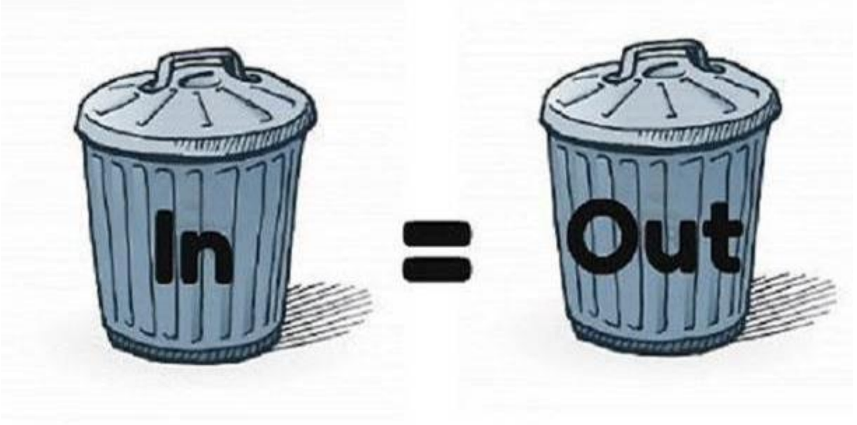
ABSTRACT

Free-stall cubicles are designed so that cows do not defecate in the bedding material, while still providing comfortable lying places. However, fixed cubicle elements, such as the partition and neck rail, restrict available movement space and may hinder cows from performing natural lying down and standing up movement patterns. Although there are various types of cubicle partitions that differ in shape, dimensions, or materials, there is no method other than visual observations to assess their effects on cow welfare. An automated detection system could improve the efficiency and promote objectivity of such assessments. Therefore, the aim of this research was

backwards, were identified as most promising candidates for accelerometer-based detection. These behaviors were detected with balanced accuracies of 0.67 and 0.74, respectively, and their learning curves indicated that more training data might further improve model performance. Overall, achieved performances were not yet satisfactory for application in the evaluation of new dairy cow housing installations. Potentially, ethograms designed for human observers are not optimal for machine learning and adjustments with machine learnability in mind might be necessary. The behaviors identified as promising are good candidates for further development into an efficient and objective method that could complement human observations in the assessment of dairy cow housing installations.

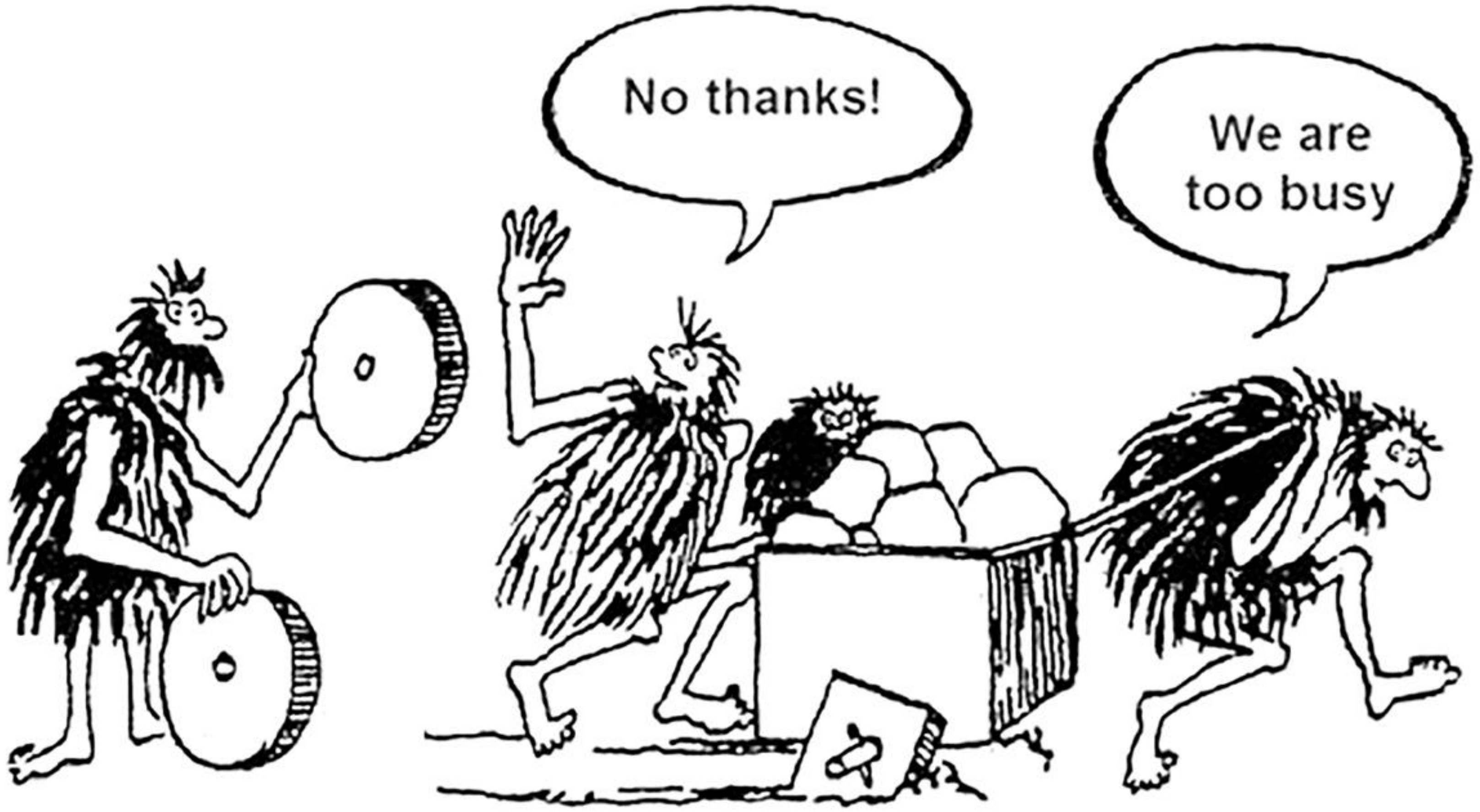
satisfactory for application in the evaluation of new dairy cow housing installations. Potentially, ethograms designed for human observers are not optimal for machine learning and adjustments with machine learnability in mind might be necessary. The behaviors identified as promising are good candidates for further development

head lunge
fluidity



See also: Nasser & Kasper, 2026

The Data Bottleneck



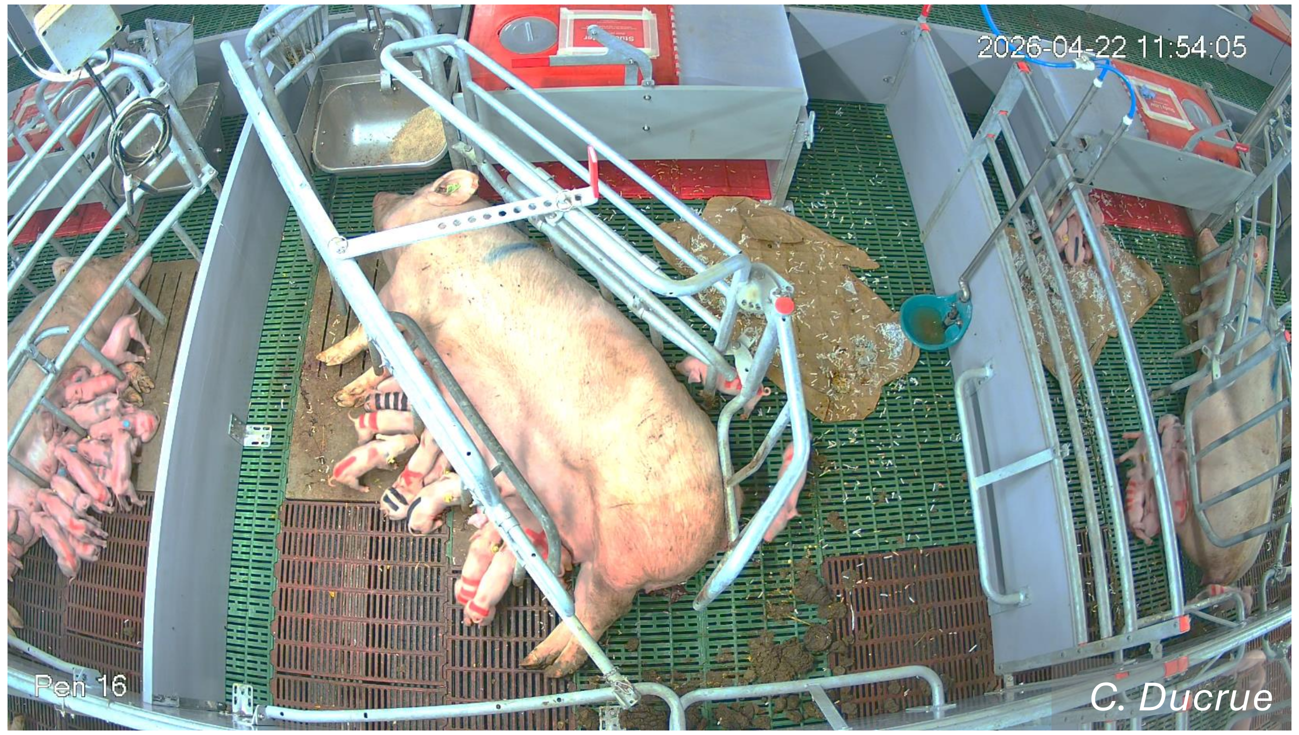
ABA intends to reduce human labour, yet its current development depends on vast amounts of manual annotation

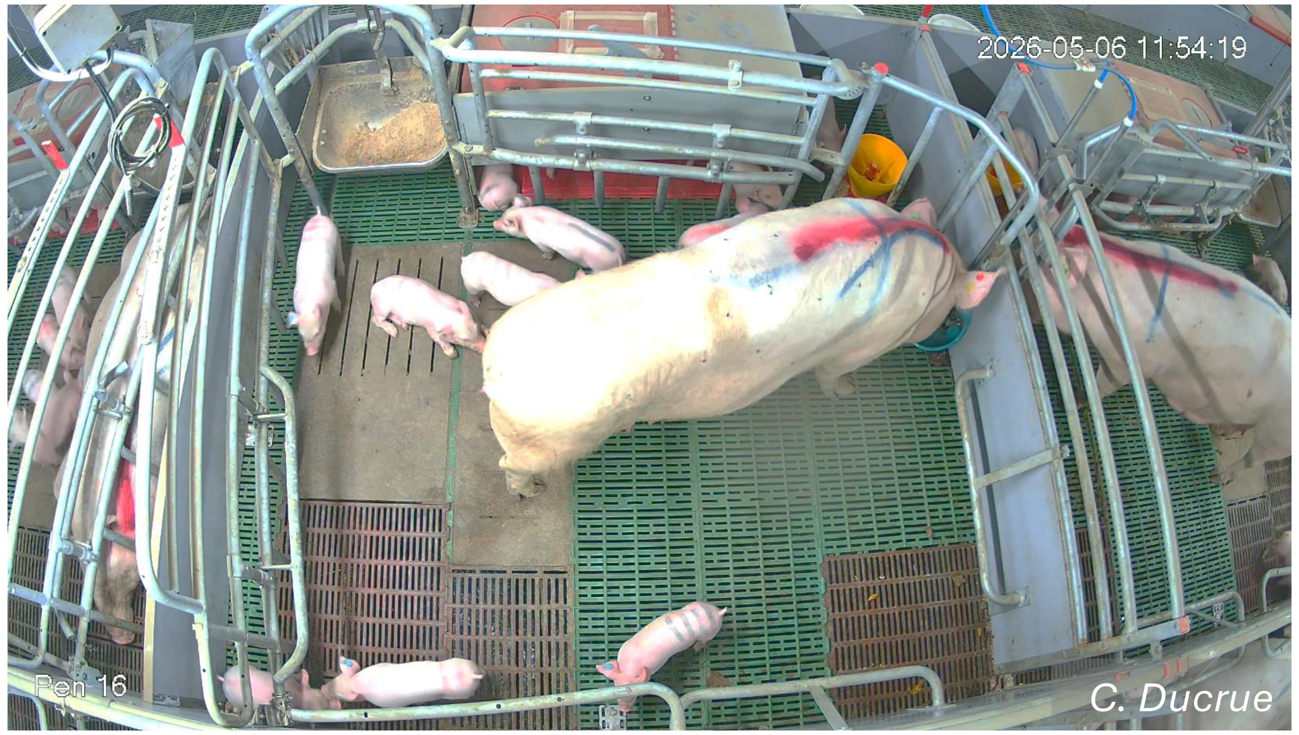
Teagasc
Agriculture and Food Development Authority
Queen's University Belfast


Clever Hans
Mary Evans Picture Library

See also: Pastell et al., 2025

Failure of Generalisation





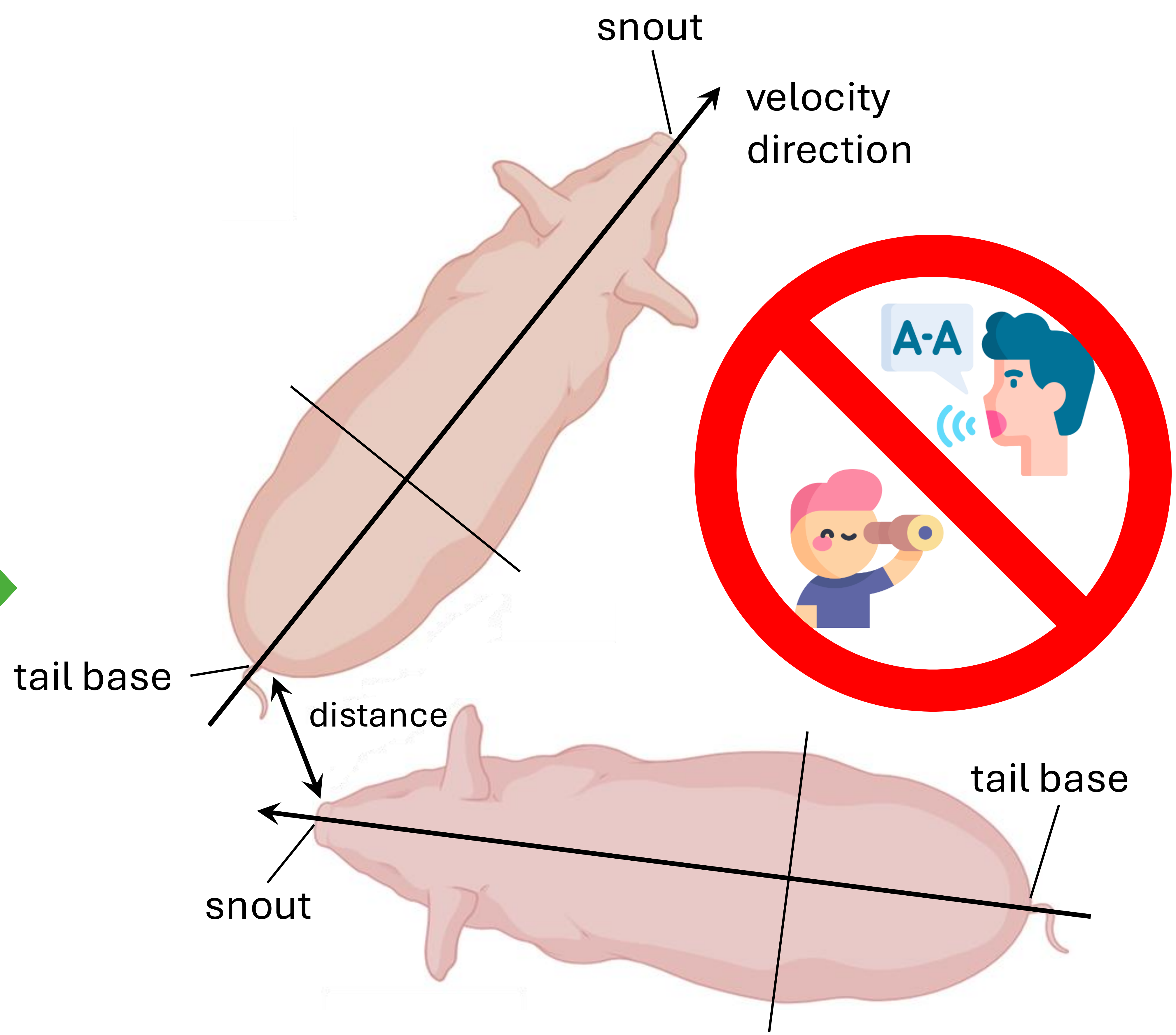
Teagasc Agriculture and Food Development Authority

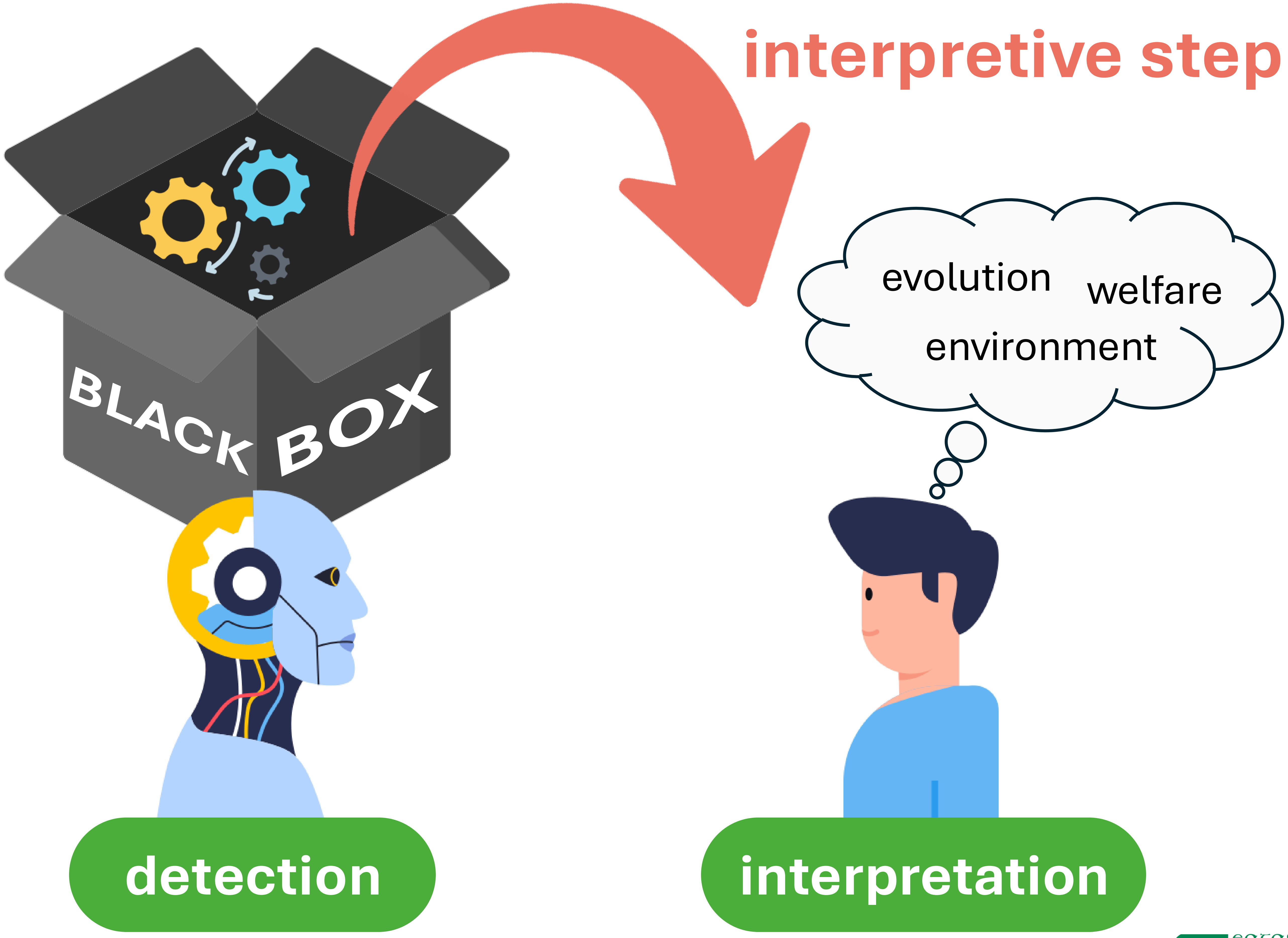
QUEEN'S UNIVERSITY BELFAST

So how can we move forward?

See also: Egnor & Branson, 2016

Ethogram
Playing
Fighting
Foraging
Resting
Running
Tail-biting
…




interpretive step
BLACK BOX
evolution
welfare
environment
detection
interpretation

Instead of this:

Table 4. Performance comparison of various models in different testing scenarios.

Method	Dataset	HOTA↑ (%)	MOTA↑ (%)	IDF1↑ (%)	MT↑ (%)	ML↓ (%)	FP↓	FN↓	IDS↓	FPS↑
Deep-OC-Sort (yolov8)	night-in scene	67.9	64.1	71.9	5	3	0	6436	25	15.7
	night-out scene	33.0	28.2	38.5	1	9	0	22,759	298	27.9
	daytime-out scene	83.9	90.0	78.2	13	0	575	527	244	36.0
FairMOT	night-in scene	62.5	68.8	56.8	6	1	760	4780	85	22.0
	night-out scene	39.0	33.5	45.5	7	4	6266	14,972	92	25.7
	daytime-out scene	80.0	90.5	70.7	12	0	535	585	161	40.5
CSTrack	night-in scene	56.2	64.7	48.8	5	3	1	6191	79	16.8
	night-out scene	48.4	52.8	44.4	10	4	1142	13,849	174	23.2
	daytime-out scene	79.5	91.1	69.4	11	0	333	679	194	45.6
SDGTrack	night-in scene	80.0	78.1	81.9	9	0	1558	3371	11	17.2
	night-out scene	75.8	69.6	82.6	17	2	1678	9342	10	41.3
	daytime-out scene	92.9	95.0	90.8	12	0	176	500	3	48.8

↑ indicates performance metrics where higher values are preferred; ↓ indicates metrics where lower values are preferred.

Liu et al. 2025

Does model output allow meaningful biological inference?

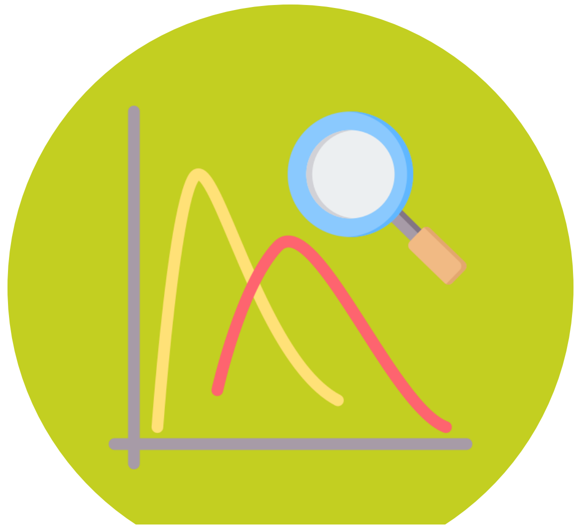
Biological Validation

Christensen et al. 2025

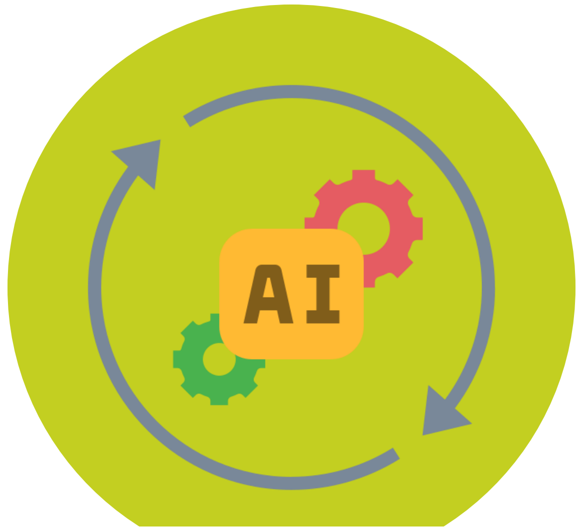
test simple hypotheses that we know a priori to be supported in our study system

models with ‘low’ standard metrics can still be powerful for hypothesis testing

Other paths



probabilistic methods

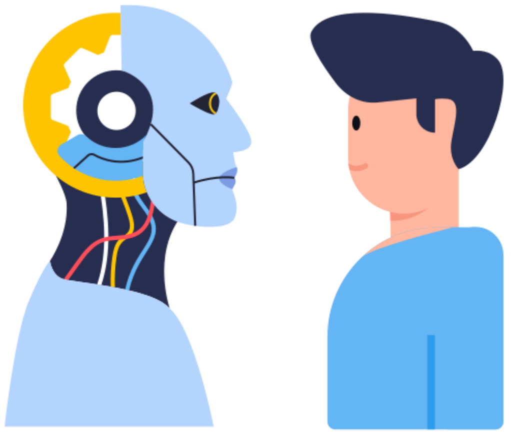


semi- and self-supervised learning

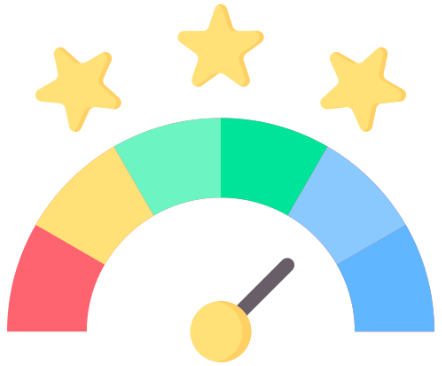
federated learning

Take-Home

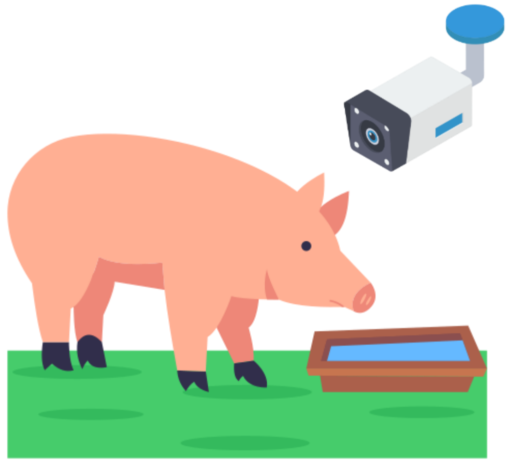
black box



human-in-the-loop



accuracy



biological meaningfulness

QUEEN'S UNIVERSITY BELFAST

teagasc
AGRICULTURE AND FOOD DEVELOPMENT AUTHORITY

Thank You!

stijn.brouwers@teagasc.ie

icons from flaticon.com

Bibliography

Brouwers SP, Simmler M, Savary P, Scriba MF. 2023. Towards a novel method for detecting atypical lying down and standing up behaviors in dairy cows using accelerometers and machine learning. *Smart Agricultural Technology*, *4*, 100199.

Christensen C, Ferreira AC, Cherono W, Maximiadi M, Nyaguthii B, Ogino M, Herrera DC, Farine DR. 2025. Moving towards more holistic machine learning-based approaches for classification problems in animal studies. *Animal Behaviour*, *230*, 123386.

Egnor SR, Branson K. 2016. Computational analysis of behavior. *Annual Review of Neuroscience*, *39*, 217-236.

Liu T, Jie D, Zhuang J, Zhang D, He J. 2025. Sdgtrack: A multi-target tracking method for pigs in multiple farming scenarios. *Animals*, *15*(11), 1543.

Nasser HR, Kasper C. 2026. Unlocking the potential of computer vision in precision pig farming: a call for a collaborative data and AI models platform. *Animal Frontiers*, vfag003.

Pastell M, Fred H, Norton T, Aernouts B, Taghipoor M, De Temmerman PJ, Maselyne J. 2025. Call for machine learning guidelines for precision livestock farming. 1st EAAP Conference on Artificial Intelligence 4 Animal Science. Zurich, Switzerland.

Siegford JM, Steibel JP, Han J, Benjamin M, Brown-Brandl T, Dorea JR, Morris D, Norton T, Psota E, Rosa GJ. 2023. The quest to develop automated systems for monitoring animal behavior. *Applied Animal Behaviour Science*, *265*, 106000.