The bottleneck is the PDF

Today, artificial intelligence can derive a linear conformation description from a photo, convert breeding values between the German and American systems, and justify a mating for each individual cow. What causes it to fail in practice is something else: the format in which breeding organizations release their data.

By Helmut Demmelhuber and Matthew Zonderop

There is one note in our work that reveals the most to us about the state of the art. It appears on our tools page, directly below each individual application, and reads, in essence: Please process six to eight animals at a time, then move on to the next batch.

Not because of a lack of computing power. Rather, because genomic data is primarily made available to farmers as individual PDF data sheets rather than as a structured bulk export. Analyzing an entire herd in a single pass does not fail because of artificial intelligence. It fails because of a file format.

That is the shortest, most honest summary of where we stand today—and at the same time, the reason why the most interesting question for the coming years is not a technical one.

What's Already Working Today

Over the past few months, we've built a series of tools and made them publicly available as free beta versions. Not as a product, but to find out what works in practice. Four areas of application have proven to be truly robust.

Developing breeding goals through discussion. A farm manager usually knows what's bothering him, but rarely how to prioritize it. A language model can conduct a structured interview that transforms vague priorities into a measurable, farm-specific breeding goal—for Holstein, Jersey, Brown Swiss, and Fleckvieh, each within the breed-specific trait logic. This does not require a single uploaded file. The value lies not in the calculation, but in the systematic questioning.

Individual mating based on genomic breeding values. A well-founded mating recommendation is generated from a cow's genomic data sheet and the bull profiles in the catalog—based on RZG or TPI for Holstein, JPI and NM$ for Jersey, and GZW for Brown Swiss. This is the essence of what a breeder used to do at the kitchen table with a catalog and a pencil—only faster and with more traits taken into account.

Conformation evaluation from photos. This is the application that surprises experts the most. Current image models can classify a Holstein cow using the linear description system and suggest a rating of EX, VG, or GP. For Fleckvieh, the same process works according to the Fleckscore system with a linear 1-to-9 description and the four composite scores for frame, muscling, legs and hooves, and udder on a scale of 68 to 93.

In our tests, the deviation in the photo-based linear evaluation was usually no more than one point from the actual classification—depending on image quality and perspective. This requires standardized photos taken from defined angles, including the hindquarters. In combination with genomic profiles, this can also be used to generate a preliminary evaluation based on the Triple-A system.

Two points must be made here. First: This is no substitute for a classer or a judge. A photo captures a single moment under uncontrolled conditions; assessing an animal at a standstill, in motion, and in comparison with the herd is something else entirely. Second: It is already useful today for an initial classification, for preselection of young animals, and for farms without regular access to a breeding consultant—partly because it puts subjective "off-day" performance into perspective. It is precisely in this level of differentiation that the true benefit lies.

Rotation strategies for embryo transfer. The most challenging case we've calculated so far. For a top-tier animal, a comprehensive strategy is developed for OPU and ET use over multiple cycles: initially inseminations with very high total breeding values, followed by targeted balancing bulls, and then phases focused on specific traits such as polled, red factor, or slick. This is a planning task involving multiple levels simultaneously—exactly the kind of problem that human working memory reliably fails to handle.

Conversion between breeding value systems. A converter between RZG and TPI may sound unspectacular, but it solves a real-world, everyday problem. Those who use American genetics are constantly comparing numbers that aren't directly comparable. A tool that performs this conversion with calibration and consistency checks prevents precisely the errors that would otherwise creep unnoticed into decision-making.

The Chain Behind It

Individual tools solve individual problems. What's more interesting is the sequence. Under the umbrella of **360° Breeding Planning**, three steps interlock, and they cannot be interchanged arbitrarily.

Step 1 – 360° Herd Analysis. Breeding values, genomic data, pedigrees, and herd management data are incorporated; the result is a strengths-and-weaknesses analysis of the herd as a whole: genetic levers, cow family patterns, inbreeding trends, and potential. An advanced variant identifies clusters within the herd—such as show animals, high-performance animals, and economically oriented animals—and derives different strategies for different groups based on this analysis. This is practically relevant because hardly any herd is uniform, yet most farms still mate their cows uniformly.

It often becomes apparent that certain traits occur more frequently—not by chance, but as a result of earlier breeding decisions. Leg conformation, udder health, or fertility may have been specifically improved over generations, but in some cases they may also have drifted unintentionally in an unfavorable direction. At the Carstens farm in Visselhövede, all genomic herd data was evaluated on this basis, and a breeding strategy was derived from it—down to first-, second-, and third-ranked bull recommendations for each individual animal, suitable for everyday barn operations.

Step 2 – Flexi-Herd Index. The breeding goal is converted into a weighting, and the weighting is used to create the farm's own ranking list for bulls, cows, and young stock. Instead of a general overall breeding value, an index is created that takes into account robot suitability, longevity, fertility, hoof health, milk composition, and conformation according to the farm's specific requirements.

Step 3 – Flexi-Mating. Based on this, individual animal decisions are made: which bull for which cow, taking into account genetic strengths and weaknesses, inbreeding, health traits, and performance data. The goal is explicitly not the highest individual breeding value, but balanced herd development.

The process can be summarized in four words: data, analysis, weighting, recommendation. Anyone who starts with the recommendation—and this is the norm in practice—gets an answer to a question they never asked.

The inconvenient aspect of farm-specific indices

When the farm manager sets the weights, the system optimizes based on the manager's own perception. Someone who considers hoof health to be their main problem will assign it a high weight—and may overlook precisely what they aren't even aware of in the first place. A general overall breeding value has an underestimated advantage over an individual index: it also includes what the individual does not currently consider important.

The crucial question for any farm-specific index is therefore not how finely it is weighted, but whether it is allowed to contradict the user's assessment—whether the chosen weighting is checked against the findings from Step 1 and a contradiction is flagged. A system that accepts self-assessments uncritically reinforces errors instead of correcting them.

Here's what it looks like when it works: In one of our ET evaluations, a bull that the author personally found likable was clearly excluded—with a breeding rationale that was entirely understandable. That is the litmus test. It's not that the machine is right, but that it can explain why. A system that merely reflects its own preferences is worthless. One that contradicts without providing a rationale is dangerous.

Same chain, different hemisphere

Perfect Cow Breeding Solutions follows the same process in New Zealand, but with a different set of traits. The system is based on NZAEL's RAS lists, the breeding values BW, PW, and TOP, genomic breeding values, and herd testing data; the economic evaluation relies on key metrics from DairyNZ. Cows are categorized into breeding tiers, and bull teams are assembled to complement the actual herd profile rather than target an average.

One notable point highlights the system's limitations: Foreign bulls are mapped to the New Zealand herd profile before they are even considered. Anyone who uses international genetics without this mapping is comparing figures that have nothing to do with one another.

The difference from the German approach is not technical but structural in nature. A pasture-based operation with block calving faces a strict, annually recurring selection threshold: whether or not the cow becomes pregnant within the designated time window. This shifts the emphasis heavily toward fertility and functional traits and shortens the timeframe over which a cow must prove her value. An index that works exceptionally well in Waikato would be ill-suited for a robotic operation in Baden-Württemberg—and vice versa.

Back to the bottleneck

And that brings us to the main topic. The technical capability to evaluate large herds using AI is available. What's missing is access to one's own data in a processable format.

Genomic results are primarily provided to farmers as individual PDF data sheets from the vit or Netrind system, with one animal per file. For home-bred breeding bulls, complete data exports are sometimes not available at all. The result is absurd: A farm with 300 genotyped animals has 300 documents, from which a table can be created only with considerable effort—while the same data is, of course, available in a structured format at the breeding organization.

A comparison with New Zealand illustrates the scale of the issue. There, farmers themselves export a standardized CSV file from MINDA, the central herd database of the Livestock Improvement Corporation, which, according to industry figures, covers approximately 93 percent of all dairy cattle in the country: phenotypic data, genomic profiles, production and health histories, live weights, and every lactation curve. A single file, available for download. On this basis, a New Zealand sharemilker can analyze a 900-cow herd across 26 traits before lunch. The difference does not lie in the technology. It lies in access.

Added to this are differing reference bases, base conversions, and varying trait definitions across organizations and countries. A model that is unaware of these differences compares numbers that are not comparable and presents the result in flawless sentences. A flawed Excel spreadsheet looks flawed. A flawed AI response looks like expert advice.

Our demand is therefore straightforward and concrete: **structured exports of a farm's own herd data in a machine-readable standard format, for every farm, at no additional cost.** Not as a concession, but as a matter of course—it is the data from their own animals. Whoever solves this issue will make more of a difference over the next five years than any new algorithm.

What can realistically be added in three to five years

More selectable traits. For German Holsteins, RZPersistence and RZFeedEfficiency were added in April 2025 with the transition to the single-step method; reliability increased by up to 14 percent depending on the trait, most notably for health and functional traits. Heat tolerance and methane reduction are currently under development. Each new trait expands the scope within which a farm-specific index can differentiate.

Traits derived from data that is collected anyway. From the MIR spectra of milk testing, it is possible to indirectly derive ketosis risk via blood BHB levels, energy balance, mastitis, ruminal acidosis, pregnancy, and individual animal methane emissions. This is already being used in Canada and by some German state milk testing associations. Measurements are taken monthly on every farm—but only a fraction are analyzed.

Image-based trait assessment in the barn. What currently functions as single-image analysis is evolving into continuous monitoring. Research on automated body condition scoring demonstrates reliable classification from camera images, with rear-view images yielding the best results. The benefit lies not in the camera itself, but in the fact that isolated assessments are transformed into trends.

Sensor data in mating decisions. Data on milkability, intervals between milkings, activity levels, and rumination patterns are available on many farms but are practically never factored into breeding decisions. This is less a technical problem than an interface issue—see above.

And what won't happen

No method can generate information that isn't already in the data. The heritability of a trait is a hard limit: where the genetic contribution is small, the prediction remains uncertain, no matter how advanced the calculations are.

Language models do not estimate breeding values and should not attempt to do so. And they make things up if left to their own devices. Research on AI advisory systems in agriculture identifies hallucination as a key risk—particularly contextual hallucination, where a piece of advice is technically correct in one region but harmful in another. Applied to breeding: a recommendation that makes sense in the New Zealand pasture system but costs money in a German robotic operation. That is precisely why the responsibility remains with the farm manager—and that is precisely why he must be able to understand the recommendation.

The bottleneck in the coming years will not be computing power. It will hinge on two questions: whether we can access our own data, and whether better data will actually lead to better decisions.

About the Authors

Helmut Demmelhuber is a management consultant based in Tübingen, Baden-Württemberg, and an expert in the use of artificial intelligence in cattle breeding. He grew up on a Bavarian dairy farm and has been actively involved in Holstein breeding again since 2025. He advises breeders and breeding organizations on the practical integration of AI into cattle breeding.
www.helmut-demmelhuber.com

Matthew Zonderop, as a 50:50 sharemiller, manages a herd of about 400 KiwiCross cows near Matamata in New Zealand's Waikato region and is the founder of Perfect Cow Breeding Solutions. He grew up on a New Zealand pasture-based farm and, drawing on his own on-farm experience, developed an AI-supported mating system that is now used on several New Zealand farms. He creates farm-specific mating plans based on 26 traits per animal and was a finalist in the Fieldays Innovation Awards in 2025.
www.perfectcow.co.nz

Boxes related to the article

Box 1 · What Works Today—and What Doesn't

What works today:

- Developing farm-specific breeding goals through a structured interview, without any files
- Individual mating based on genomic breeding values (RZG/TPI, JPI/NM$, GZW)
- Strengths-and-weaknesses analysis and clustering within a herd
- Linear conformation description and grading recommendation based on photos (Holstein: EX/VG/GP; Fleckvieh: Fleckscore)
- Conversion between breeding value systems with consistency checks
- Bull selection and semen management based on breeding goals and herd analysis

Not currently possible:

- Processing an entire herd in a single run—limited to six to eight data sheets per run
- Automatic data retrieval from herd management software or breeding organization
- Feedback based on actual progeny performance, because the first generation of daughters takes four to five years

The reason for the first limitation is not technical, but rather a formatting issue: genomic data is primarily available to farmers as individual PDF data sheets, not as a structured export.

Box 2 · Three Things That Are All Called "AI"

In cattle breeding, "AI" traditionally stands for **artificial insemination**. In this article, it refers to artificial intelligence. Within this context, it's worth making a further distinction because the three levels vary in reliability:

Statistical learning methods—the basis of breeding value estimation. Mature, proven, with established reliability.

Machine learning and neural networks—pattern recognition in images, sensor data, and spectra. Partially ready for practical use in trait assessment; in breeding value prediction, they have so far offered only incremental advantages over traditional methods.

Language models—process text, synthesize sources, and explain relationships. Useful for data processing and advisory services, unsuitable for breeding value estimation, and prone to plausible-sounding errors without validation.

Box 3 · The Question Determines the Answer

If you feed a language model with generic queries, you'll get generic answers. Four pieces of information dramatically improve the quality—you can remember them as a formula:

1. Who am I? In what capacity am I speaking? Farm manager, consultant, breeding supervisor—from which perspective are you working?

2. What exactly is the issue? Farm, herd size, housing, milking system, target group, context, problem.

3. What should be done? Analyze, compare, structure, explain—and in what format: list, table, or continuous text.

4. How should the system approach this? As a neutral analyst, an experienced consultant, or a critical sparring partner.

Example: “I am a dairy farmer and make management decisions for my farm. My farm has about 150 cows and uses milking robots. Suggest five specific measures to improve fertility management and briefly justify each one. Respond as an experienced, practical consultant for fertility and herd management—providing technically sound and understandable advice.”

Two additions that are particularly important in breeding: explicitly limit the data sources to be used (“use only the bulls from the uploaded files; no other lists or internet sources”). And always specify the breeding goal rather than assuming it.

Box 4 · Five Questions for Every Provider

- What data does the system use for its calculations—and how up-to-date is it?
- Are breeding values from different sources or reference databases compared, and how is this adjusted?
- Can the system explain why it arrives at a particular recommendation?
- What happens to my herd data, and who is allowed to reuse it?
- How would I know if the recommendation was wrong?

The last question is the most revealing. Anyone who cannot answer it has not built a decision-making system, but rather formed an opinion.