Athleet.Ai

PERFORMANCE INTELLIGENCE
WHITEPAPER SERIES

PAPER FOUR OF SEVEN · GRASSROOTS AND ACADEMY FOOTBALL

Casting the Net Wider

AI as a force multiplier for talent discovery, development and retention in grassroots and academy football

Paul Forrest
Founder and Chief AI Officer, Athleet.AI

October 2025

About this series

This is the fourth of seven papers examining how artificial intelligence can be used responsibly in performance sport. Each paper stands alone. Together they set out a single argument, which is that the durable advantage in sport comes from better decisions rather than from more data, and that better decisions require systems designed around the person who is accountable for them.

This paper concerns children. That changes what may be built, what may be recorded and who decides, and the difference runs through every chapter rather than sitting in a compliance appendix at the back.

Editorial balance. Roughly seven parts in ten of this document is evidence and sector analysis, two parts in ten is original framework, and one part in ten concerns the Athleet.AI platform itself.

HOW TO READ THE EVIDENCE TAGS

Every substantive claim carries a tag stating what kind of evidence supports it.

PEER-REVIEWED A primary study in a peer-reviewed journal.

REVIEW A systematic review or meta-analysis.

CONSENSUS A consensus statement, governing-body position or regulator guidance.

SURVEY A national survey, official statistic or methodological commentary.

PRACTICE Practitioner opinion, including the author's own.

ATHLEET.AI The author's own framework, proposed rather than proven.

Where a figure could not be traced to a primary source it has been left out rather than estimated.

Legal, regulatory and safeguarding notice. Chapter 10 describes data protection law, regulator guidance and safeguarding standards. Nothing in this document constitutes legal advice. The provision of legal advice sits outside the terms of any engagement with the author or with Athleet.AI, and the material is presented to support discussion and further review by qualified advisers.

Nothing here replaces a club's obligations under its governing body's safeguarding policy, or the judgement of a designated safeguarding officer. Where this paper draws a practical implication that a source does not itself state, the text says so.

ILLUSTRATIVE MATERIAL

Every club, player, scout and scenario in this paper is invented. Cases are marked **illustrative scenario** wherever they appear. Published research and official statistics are cited and numbered; invented material is never presented as evidence.

CITATION

Forrest, P. (2025). *Casting the Net Wider: AI as a Force Multiplier for Talent Discovery, Development and Retention in Grassroots and Academy Football.* Athleet.AI Performance Intelligence Whitepaper Series, Paper Four. Version 1.0, October 2025.

ALSO IN THIS SERIES

- **Paper one.** *The Responsible Performance System.* Governance and trust in sports AI.
- **Paper two.** *The Intelligent Track.* AI and event-specific expertise in elite athletics.
- **Paper three.** *Closing the Coaching Gap.* Responsible AI in grassroots athletics.
- **Paper five.** *The Connected Endurance Athlete.* Integrated preparation in triathlon.
- **Paper six.** *The Augmented Touchline.* Responsible AI in elite football.
- **Paper seven.** *The Augmented Golfer.* AI across technique, practice and strategy.

Contents

FRONT MATTER

A note from the author 4
Executive summary 5

PART ONE · A NARROWER NET THAN FOOTBALL BELIEVES

1 Why the existing net is narrower than football assumes 8
2 The pathway's hidden filters 9
3 AI as a force multiplier 11
4 From snapshot to trajectory 12

PART TWO · DEVELOPING THE PLAYERS YOU ALREADY HAVE

5 Smarter training for developing players 15
6 The coach capacity problem 17
7 The connected player record 18
8 Late developers and deselected players 19

PART THREE · WIDENING ACCESS WITHOUT SCALING BIAS

9 Widening access without scaling bias 20
10 Safeguarding children and families 21
11 Transitions and career durability 23
12 Value for clubs, sponsors and communities 23
13 A federated talent ecosystem 24
14 Implementation roadmap 26

END MATTER

Pathway readiness checklist 28
Glossary 29
Method and evidence grading 29
About the author 31
References 32

A NOTE FROM THE AUTHOR

The players nobody watched twice

Football believes it has the widest talent net in world sport. It watches more children, in more countries, more often, than any other game. The evidence assembled in this paper suggests something less comfortable, which is that the net is wide and it is also full of holes cut in a very particular shape.

Birth month cuts one. Biological maturity cuts a much bigger one. Where a child lives cuts another, and so does what a family can afford in petrol and Saturday mornings. None of these holes has anything to do with football ability, and all of them operate before anybody has watched a player for a second time.

I have spent my working life on how organisations turn what they know into what they do, and I coach in a different sport at a very different level, which is where the parallel becomes uncomfortable. Every coach I know can name a child they stopped seeing at 13 who they suspect would have been worth watching at 16. Nobody can check, because nothing was written down.

That is the gap this paper addresses. It argues that artificial intelligence can widen the net by helping people observe more, remember better and look again, and it argues at some length that anything which scores, ranks or predicts a child should be refused outright. The second argument is not a caveat attached to the first. It is half the paper.

TWO PAGES THAT CIRCULATE ON THEIR OWN

Executive summary

Football's talent problem is not finding the best player in Saturday's match. It is that the system for recognising potential is narrow, biased in ways it can name, and almost entirely without memory. Players are seen once, judged against children who are a year older or a year further into puberty, and then either signed or forgotten, with no record surviving either decision.

This paper argues that the honest opportunity is capacity and memory rather than prediction. A system can help people observe more players, record what they saw in a usable form, hold the context that makes a comparison fair, and prompt somebody to look again. It argues equally firmly that no model should rank a child, score their potential, or be trained on a club's own past selections, because such a model learns the club's history rather than the game.

1

Early success predicts remarkably little. Across 189 study samples, 89.2 per cent of internationally ranked under-17 and under-18 athletes never reach international level as seniors. Junior performance explains around 2.2 per cent of the variance in senior performance.

2

Maturity bias is larger than the relative age effect and grows with age. Across 12 academies and 1,011 players, maturity selection bias reached a very large effect by under-18. In one national pathway, no late-maturing player appeared in the under-15 or under-16 international squads at all.

3

Birth month remains a powerful filter that washes out later. Players born in the first quarter made up 43 to 46 per cent of European national youth squads against 9 to 18 per cent born in the last, and yet early-born professionals are no better valued in the transfer market.

4

Scouts watching the same children do not agree. When 83 scouts rated the same 24 under-11 players, inter-rater reliability coefficients ran between 0.02 and 0.09. Scouts also report believing they can judge adult potential from around the age of 13 and a half.

5

Selection often rewards physical maturity rather than football. In a four-year study of Dutch players aged eight to 12, sprint speed over 30 metres was the strongest single predictor of academy selection.

6

The AI research in this area is not ready to be sold. No systematic review of machine learning applied to talent identification in football had been published when this paper was written. The nearest critical review, covering AI across athlete identification, selection and development, concludes that the scope and the usefulness of these methods remain relatively unknown.

What follows for a club, an academy or an association

The recommendation is cheaper and slower than the technology allows. Record observations properly, with the date, the observer and the context. Record birth month and maturity band alongside every judgement so that a comparison can be made fair. Keep the record when a player is not taken, and build a routine that looks at it again. None of this requires a model and all of it removes bias that is currently invisible.

What should be refused is narrower and firmer. No score attached to a child. No ranking. No prediction of adult potential. No system that communicates with a child without an adult in the loop. Data protection law asks for an impact assessment before deploying anything that profiles children, for human decision-making rather than a human rubber stamp, and for a contract that establishes whether a supplier trains its own models on club or player data.

None of that is expensive, which is the part that tends to disappoint. The changes that would widen the net most are administrative rather than technical, and a club can make the first of them on a Tuesday evening with a phone and a shared record. What a system adds is consistency, reach and recall, and it adds those only once somebody has decided what is worth writing down.

Chapter 14 proposes a regional pilot across an academy and a network of community clubs, with an independent evaluation partner and measures on players observed, follow-up rate, diversity of the observed pool, and re-entry. The evidence base for AI in talent identification is currently too thin to support anything else, and a pilot of that shape is the only way I know of to make the case rather than assert it.

A reader who wants the argument without the product can stop at Chapter 13. Chapters 1 to 5 set out what the pathway actually selects for, chapters 6 to 9 what a system can honestly do about it, and chapters 10 to 13 what it must be prevented from doing.

PART ONE

A narrower net than football believes

Four chapters on who actually gets seen, what the judgement is worth, and why the honest opportunity is memory rather than prediction.

CHAPTER ONE

Why the existing net is narrower than football assumes

Football watches an enormous number of children and remembers almost none of them. Those two facts together define the problem.

The scale of English football alone is remarkable. Figures attributed to the Football Association and reported in a parliamentary library briefing put participation at 15.7 million people, supported by more than a million volunteers across roughly 18,000 accredited clubs.[30] SURVEY The underlying report could not be located during the preparation of this paper, so those numbers are cited at one remove and should be read as an order of magnitude. Even discounted, they describe a recruitment pool no other sport in the country can approach.

Against that pool sits a filtering process that is short, local and unrecorded. A scout watches a fixture. An impression forms. A player is invited or is not, and in the second case nothing at all persists. The child returns to community football carrying a judgement that nobody wrote down and that nobody will revisit, and the club retains no record that it ever saw them.

THE CORE CLAIM

The opportunity in talent discovery is capacity and memory. It is not prediction, and the published evidence on prediction is unusually clear about why.

The consequence is that football is unable to learn from its own decisions. A club cannot ask which of the players it declined went on to succeed elsewhere, because it does not know who it declined. It cannot ask whether its scouts are systematically missing a particular kind of player, because it holds no data on the players they passed over. Every season the process starts again from nothing.

1.1 What early success is actually worth

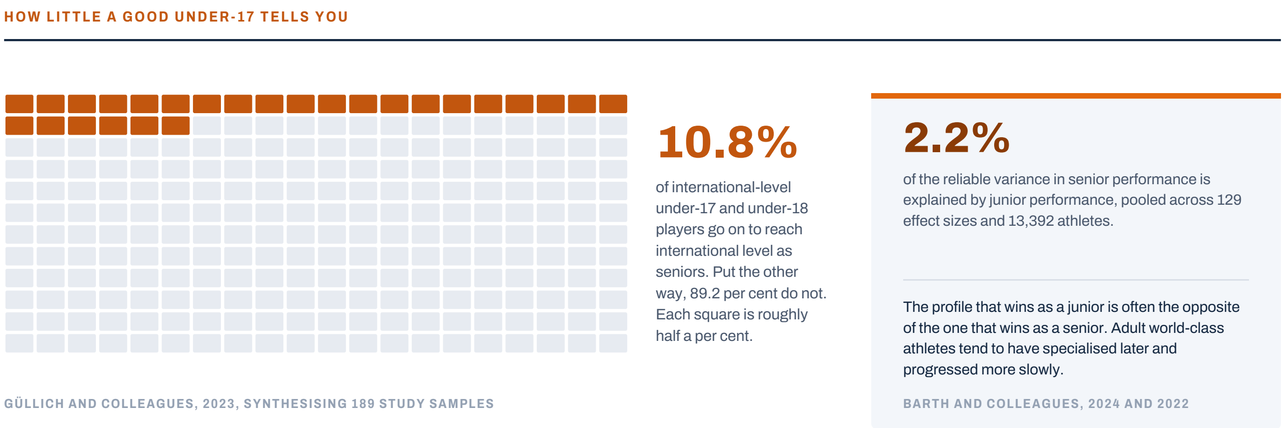


FIGURE 1 The junior-to-senior transition, as the evidence describes it. Each square is roughly half a per cent.

Güllich and colleagues, 2023; Barth and colleagues, 2024 and 2022. REVIEW

A synthesis of 189 study samples found that around 89.2 per cent of international-level under-17 and under-18 athletes never reach international level as seniors, meaning successful juniors and successful seniors are largely disjoint populations.[16] REVIEW A meta-analysis of 129 effect sizes across 13,392 athletes found junior performance explained about 2.2 per cent of the reliable variance in senior performance.[15] REVIEW A third synthesis of 71 studies covering 9,241 athletes found that adult world-class athletes typically specialised later and progressed more slowly than those who peaked as juniors.[17] REVIEW

Those three findings are drawn across Olympic sports rather than football alone, and a football-specific effect could differ. They are nonetheless the most methodologically robust evidence available, and they point the same way. The profile that wins at 16 is frequently the opposite of the profile that succeeds at 26.

THE ARGUMENT IN ONE PARAGRAPH

Football should stop trying to identify who will be good and start making sure it can still find them later. That means observing more players, recording the context that makes a comparison fair, keeping the record of the ones not taken, and building a routine that looks again. None of it requires a prediction, and all of it is within reach of clubs that own nothing more sophisticated than a shared spreadsheet today.

CHAPTER TWO

The pathway's hidden filters

Five filters operate before anybody has watched a player twice, and not one of them measures football.

THE PATHWAY'S HIDDEN FILTERS

Birth month HELSEN AND COLLEAGUES, 2005	Q1-born players made up 43 to 46 per cent of European national youth squads, against 9 to 18 per cent born in Q4.
Biological maturity SWEENEY AND COLLEAGUES, 2023	In one national pathway, no late-maturing player appeared in the under-15 or under-16 international squads at all.
Maturity, again CURNYN AND COLLEAGUES, 2025	Across 12 academies and 1,011 players, maturity selection bias grew with age and reached a very large effect by under-18.
What is measured FORTIN-GUICHARD AND COLLEAGUES, 2022	Sprint speed over 30 metres was the strongest single predictor of academy selection between the ages of nine and 12.
Who is doing the looking LÜDIN AND COLLEAGUES, 2023	Eighty-three scouts rating the same 24 under-11 players agreed with one another almost not at all.

None of these filters is about football ability, and all of them operate before anybody has watched a player twice.

FIGURE 2 Five documented filters in the talent pathway, none of which is a measure of football ability.

References 1, 5, 6, 18 and 31. PEER-REVIEWED

2.1 Birth month, and why it fades

The relative age effect is the best-known filter and the least serious of the two age-related ones. Analysis across European youth football found players born in the first quarter of the selection year making up 43 to 46 per cent of national youth squads against 9 to 18 per cent born in the last quarter.[1] PEER-REVIEWED A meta-analysis of 38 studies across 14 sports and 16 countries confirmed the effect as consistent and generally small to moderate, strongest among adolescent males at representative level.[2] REVIEW

Two later findings complicate the picture usefully. An analysis of English under-18, under-21 and senior squads found significant bias toward relatively older players persisting through the youth ages and no significant skew at senior level.[4] PEER-REVIEWED A large study of European professionals found early-born players over-represented and no better valued in the transfer market.[28] PEER-REVIEWED The effect is therefore a selection bias rather than a performance advantage, and it washes out by adulthood, which is to say after most of the damage is done.

There is a corollary that deserves more attention than it gets. An 11-season study at one English academy found that relatively younger entrants who survived selection progressed to professional contracts at a markedly higher rate than the relatively older ones.[3] PEER-REVIEWED That is a single-club sample and should not be generalised. If it holds more widely, the players a club works hardest to identify are precisely the ones it currently filters out.

THE UNDERDOG FINDING

Across 11 seasons at one English academy, relatively younger entrants who survived selection were markedly more likely to reach a professional contract. Single-club evidence, and worth knowing.

2.2 Maturity, which is the bigger problem

Biological maturity is the filter clubs discuss least and should discuss most. Across 12 academies and 1,011 players, one study found a small-to-moderate relative age effect alongside a much larger maturity selection bias that grew with age, reaching a very large effect by under-18.[5] PEER-REVIEWED In a national talent pathway of 159 players, the maturity bias was moderate to very large and increased with age, with no late-maturing player appearing in the under-15 or under-16 international squads.[6] PEER-REVIEWED

Both studies are cross-sectional and neither follows players into adulthood, so they document who gets selected rather than who succeeds. What makes the finding serious is that the bias grows through exactly the years in which academies commit resources, which means the sport is progressively concentrating investment on the children who happened to grow first.

Maturity-matched competition has been trialled as a response. In a maturity-banded tournament involving 115 academy players, earlier maturers reported greater physical and technical challenge, and later maturers reported more opportunity to show technical and tactical ability.[10] PEER-REVIEWED That is perception data from one tournament, so it establishes that the format changes the experience rather than that it changes outcomes.

2.3 Where a child lives, and what a family can afford

A systematic review of the birthplace effect in football confirms that players from smaller cities are over-represented among professionals, with population density and proximity to talent centres rather than facility quality appearing to be the operative mechanism.[8] REVIEW The original cross-sport work established the pattern two decades ago.[7] PEER-REVIEWED Neither study identifies coach or scout judgement as the mechanism, and I have not found work that isolates it, so the practical inference below is mine rather than theirs.

Socioeconomic position appears in the football literature too, though the available study is small and its measure imperfect. In 58 academy players at one English club, those rated by coaches as having higher potential came from significantly higher socioeconomic backgrounds using a postcode-derived classification.[9] PEER-REVIEWED The authors' own framing matters here, since coach-rated potential is itself a subjective judgement and the finding may reflect scouting bias as much as anything about the players.

CHAPTER THREE

AI as a force multiplier

The honest claim is that a system helps people see more and remember better. Any stronger claim runs ahead of a literature its own authors describe as exploratory.

The temptation in this field is obvious. Scouting is subjective, inconsistent and expensive, so a model that ranks players looks like an answer. The published evidence is unusually direct about why it is not.

No systematic or scoping review of machine learning applied to talent identification in football had been published when this paper was written. That is a finding rather than an omission, and it is the first thing to put to a supplier selling in this space. What does exist is a critical review of how artificial intelligence is being used across athlete identification, selection and development, which concludes that the scope and the utility of these approaches, and the difficulties attached to them, remain relatively unknown, and that a balanced approach embracing both innovation and critical evaluation is needed if the tools are to help rather than disrupt.[19] REVIEW A companion paper is more specific about why. It names data availability, quality and ownership as the first obstacle, and the labelling problem in supervised learning as the second, since longitudinal work is complicated by missing data and by imbalance in the outcome itself. The development of a junior into a top-tier senior is rare, and a rare outcome produces biased models that generalise badly.[20] COMMENTARY

Underneath the methodological weakness sits a deeper problem that better data would not fix. Any model of talent needs a label, and the only labels readily available are past human decisions, meaning who was signed, who was retained, who reached a professional contract. Every one of those labels encodes the filters described in Chapter 2. A model trained on them will reproduce birth-month bias, maturity bias and geographic bias faithfully, and will present the result with a confidence interval.

THE LABEL PROBLEM

A model trained to predict who gets signed learns who the club has historically signed. That is not talent identification. It is an expensive mirror.

The clearest published account of this failure comes from healthcare, where an algorithm used cost as a proxy for need and systematically underestimated the needs of Black patients because less had historically been spent on their care.[43] PEER-REVIEWED The mechanism transfers directly. Selection history is football's version of cost.

3.1 What a system can honestly do

THE WIDER TALENT NET

The net today

- **Seen once**
  A single fixture, in a single week, in one weather.
- **Seen locally**
  Within an hour of a training ground, and mostly in affluent postcodes.
- **Seen early**
  Scouts report believing they can judge from around 13 and a half.
- **Seen and forgotten**
  The player not taken leaves no record anywhere.

The net that could exist

- **Seen repeatedly**
  Several observations across a season, held against each other.
- **Seen wider**
  Community clubs and schools contributing to the same record.
- **Seen in context**
  Birth month, maturity status and access recorded alongside the observation.
- **Seen again**
  The deselected player retrievable at 17, because the record survived.

Widening the net is a capacity problem before it is an analytical one. Nothing here requires a model to rank anybody.

Athleet.AI framework, drawing on references 1, 6, 15 and 19.

FIGURE 3 The net as it operates and the net that could exist. Nothing in the right-hand column requires a model to rank anybody.

Athleet.AI framework, drawing on references 1, 6, 14 and 18. ATHLEET.AI

Four jobs are both useful and safe. Capture, meaning a scout dictating an observation at the roadside rather than typing it on Sunday night, which is the difference between an observation that reaches the record and one that does not. Context, meaning birth month and maturity band attached automatically so that a judgement can later be read fairly. Retrieval, meaning the ability to find every observation of a player across four years in seconds. And prompting, meaning a system that remembers a coach said a player was worth another look and asks about it in March.

None of that ranks anybody. All of it addresses the actual constraint, which is that human attention is finite and human memory is worse than anybody admits.

CHAPTER FOUR

From snapshot to trajectory

One observation ranks a group. Several observations, held against each other, describe a development rate, and only the second is worth acting on.

The judgement football currently runs on is a snapshot. A scout forms an impression in 90 minutes and that impression, once recorded as a decision, is rarely revisited. The evidence on how reliable such impressions are is worth reading before building anything on top of them.

THE JUDGEMENT THE PATHWAY CURRENTLY RUNS ON

0.02 to 0.09

The range of inter-rater reliability coefficients when 83 scouts rated the same 24 under-11 players.

Two scouts watching the same child essentially disagree.

13.6 years

The age from which scouts, surveyed in one study, reported believing they could judge adult potential.

Scouts also reported integrating information intuitively, despite describing structured procedures.

38 biases

Identified in a recent framework of cognitive biases relevant to talent identification decisions.

First impression, outcome and availability biases are all named among them.

What follows, and what does not

Unreliable human judgement is an argument for structuring and recording observation, and for gathering more of it. It is not an argument for replacing the judgement with a model, since a model trained on these decisions inherits exactly this unreliability and presents it with a decimal point.

FIGURE 4 Three findings about scouting judgement. The conclusion is an argument for structuring and recording observation, and against replacing it with a model.

References 13, 14, 31 and 32. PEER-REVIEWED

When 83 scouts rated the same 24 under-11 players, inter-rater reliability coefficients ran between 0.02 and 0.09, which is to say that two scouts watching the same children essentially disagree.[31] PEER-REVIEWED The players were very young, and reliability may improve at ages where technical differences are clearer. A survey of 125 scouts found they reported believing they could predict adult potential from around 13.6 years of age, focused on easily observable technical attributes, and ultimately integrated information intuitively despite describing structured procedures.[14] PEER-REVIEWED A recent framework catalogues 38 distinct cognitive biases relevant to these decisions.[32] SURVEY

One finding cuts against the obvious conclusion and deserves to be stated. In a study of 96 coaches and scouts rating player videos, unstructured overall ratings performed marginally better than structured collection combined mechanically, when tested against later market value.[13] PEER-REVIEWED The differences were marginal and market value is itself contaminated by human judgement, so this should not be read as intuition beating structure. It is a caution against assuming that imposing a form on a scout automatically improves anything.

4.1 What trajectory adds

SNAPSHOT AGAINST DEVELOPMENT TRAJECTORY

What changes if you wait

- Development rate becomes visible, and it is a different signal from current level.
- Response to coaching can be observed rather than assumed.
- Maturity status can be estimated and the comparison adjusted.
- The player who was behind at 13 is still in the record at 16.

Illustrative curves, drawn to show a mechanism rather than measured data. The evidence that the two trajectories exist and cross is at references 3, 5, 6 and 18.

One observation ranks a squad. Several observations, held against each other, describe a trajectory. Only the second is worth acting on.

FIGURE 5 Two illustrative trajectories and the point at which a club typically selects. The curves show a mechanism rather than measured data; the evidence that both trajectories exist is at references 3, 5, 6 and 17.

ATHLEET.AI

Development rate is a different signal from current level and it is invisible in a single observation. A player improving quickly from a low base and a player static at a high one look identical on a Saturday and are entirely different propositions. Capturing the difference requires only that observations are dated, attributed and kept, which is a records problem rather than an analytical one.

What clubs currently select on tells against them here. In a four-year study of 110 Dutch players aged eight to 12, sprint speed over 30 metres was the strongest single predictor of academy selection.[18] PEER-REVIEWED The outcome in that study is selection rather than adult success, so it documents what scouts select for. Sprint speed at 11 is substantially a measure of how far through puberty a child is.

ILLUSTRATIVE SCENARIO

The invented January and the invented August

Two invented boys are both 11. One was born in September and is 14 months further into his growth. He is quicker, stronger in contact and wins more of the ball. The other was born in August, is technically the better player and loses most physical duels.

A scout watching once will take the first, and will be able to give good reasons. A record that noted both birth dates, both maturity bands and both coaches' comments would let somebody ask a different question in two years' time. The system is not required to know which boy is better. It is required to make sure the second one still exists in the record when the answer becomes visible.

CHAPTER FIVE

Smarter training for developing players

Finding a promising player achieves nothing if the environment then loads them like a small adult through the fastest growth of their life.

The management problem examined across this series applies with particular force to a growing child. What the stage of development requires, what the coaches plan, how the player responds, what is then decided, and whether the decisions are building something that lasts.

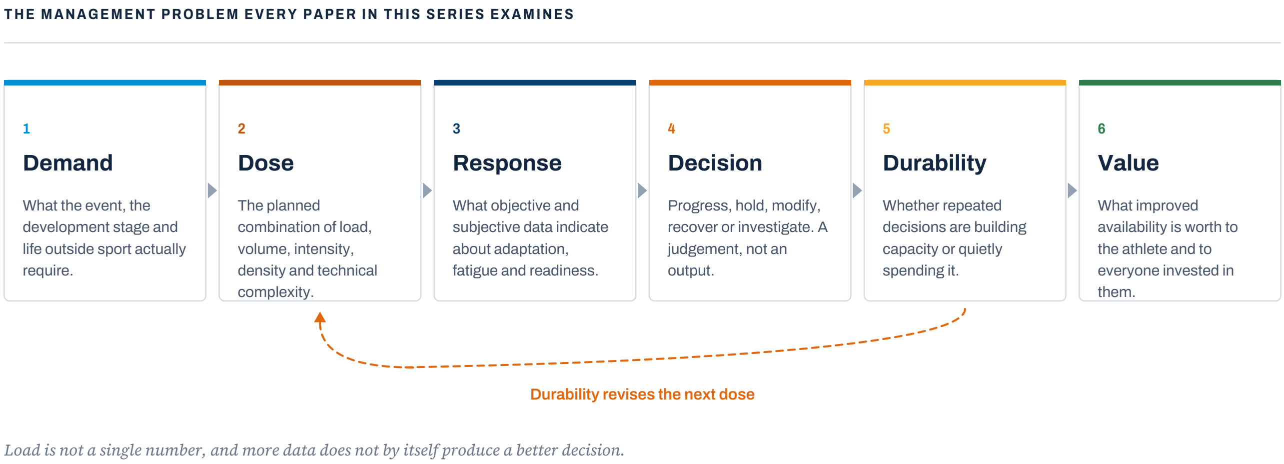
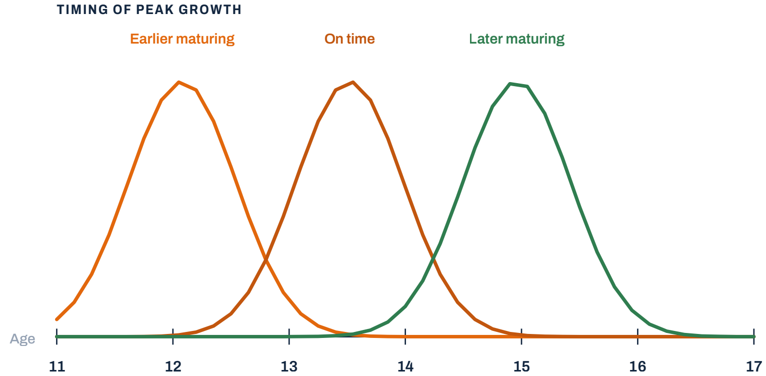


FIGURE 6 Demand, dose, response, decision, durability and value, read at pathway level.

Athleet.AI framework, used across the whole series. ATHLEET.AI

5.1 Growth changes the athlete underneath the plan

GROWTH, MATURATION AND TRAINING DEMAND



Three children of the same chronological age, growing at different times. The bands are illustrative. The consequence is not.

What a club can do about it

- Record maturity status as a band, never as a number attached to a child.
- Compare a player against others at a similar stage rather than the same birthday.
- Expect technical regression around fastest growth, and plan for it.
- Never let a maturity estimate reach a selection meeting as evidence of potential.

Maturity estimation carries substantial individual error, and the direction of that error is systematic. It belongs in a coach's thinking and nowhere near a decision record.

FIGURE 7 Three children of the same chronological age, growing at different times. The bands are illustrative; the consequence for selection and loading is not.

Athleet.AI framework. ATHLEET.AI

Around the period of fastest growth a young player's coordination, limb lengths and force production all change, and technical performance frequently regresses temporarily. A club that treats that regression as a decline in ability rather than a consequence of growth will release exactly the wrong children.

The practical response is a record rather than a model. Note maturity status as a band, compare players against others at a similar stage, expect and plan for technical regression, and keep the maturity estimate out of any selection meeting as evidence of potential. The estimation methods carry substantial individual error, as the companion athletics paper in this series sets out at greater length, and the direction of that error runs the wrong way for exactly the children a club should be protecting.

A HARD RULE

Record maturity as a band and never as a number attached to a child. The estimation error is large and its direction is systematic.

5.2 The load nobody is counting

A talented 14-year-old may play for an academy, a school, a district team and a Sunday side, and each organisation sees a reasonable amount of football. Nobody sees the total. This is the most tractable safety problem in youth football and the answer is a question rather than a sensor. Ask what else the player does in a week, record the answer, and revisit it each term.

My own view, which goes beyond what the evidence establishes, is that a shared record of total weekly football across organisations would do more for young players' durability than any wearable a club could buy. It is also, as Chapter 13 sets out, exactly the kind of thing a federated architecture is for.

CHAPTER SIX

The coach capacity problem

One coach, 20 players and a Tuesday evening is the constraint that actually shapes how much development any child receives.

Individual development is discussed as a philosophy and delivered as a resource. A coach responsible for 20 players across a season has, realistically, minutes per player per month to think about any individual one of them, and the plan defaults to the group because attention is finite.

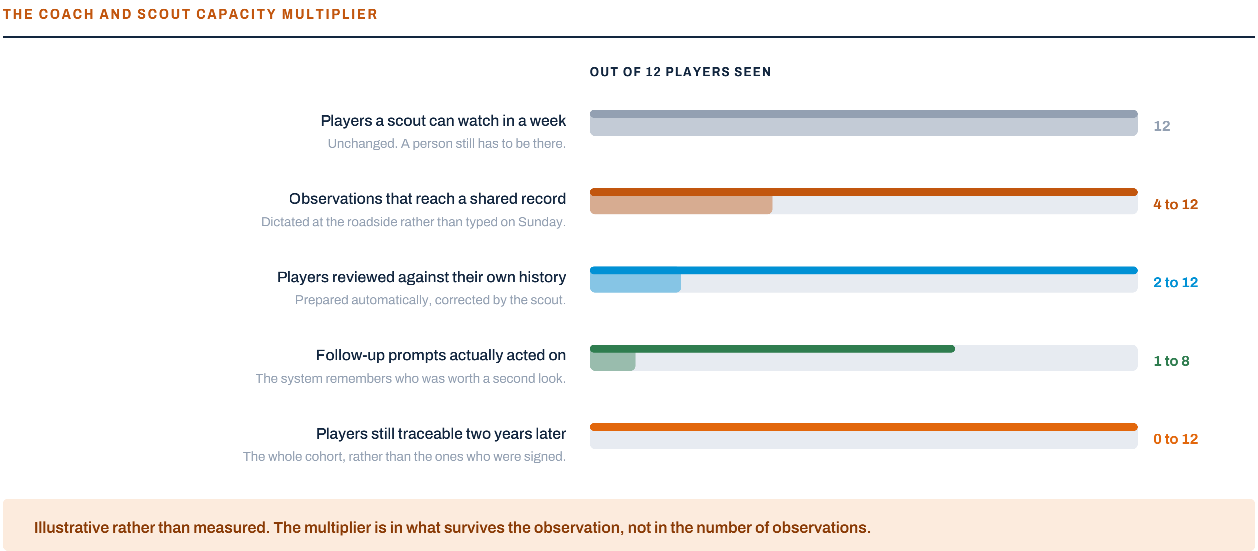


FIGURE 8 The capacity multiplier. Proportions are illustrative. The gain is in what survives the observation rather than in the number of observations. Athleet.AI framework. ATHLEET.AI

The multiplier does not work by having a person watch more football. It works by ensuring that what they saw reaches a durable record, that the record is compared against the player's own history, and that somebody is prompted to follow up. A scout who watches 12 players and files two reports has effectively watched two.

Research on coach education in grassroots football maps how policy is formulated and enacted, and describes a pathway shaped as much by administrative structure as by developmental need.[34] PEER-REVIEWED Work on English scouting practice is thinner still, with the available account principally establishing that rigorous evidence about how recruitment decisions are made is scarce.[33] SURVEY That scarcity is itself a finding, and it should temper any claim that a product is codifying established best practice.

SPECIFICATION

What a capacity tool must do, and must not

- Capture an observation in under a minute, at the side of a pitch, in the rain, on a phone.
- Attach date, observer, fixture, birth month and maturity band without anybody typing them.

- Retrieve every previous observation of the same player instantly, including ones by other people.
- Prompt a follow-up when a player has been flagged and not revisited.
- Produce no score, no ranking and no prediction about any child, at any point, for anyone.

CHAPTER SEVEN

The connected player record

The asset is the record, and its value comes as much from what it refuses to hold as from what it holds.

THE CONNECTED PLAYER RECORD

Observation	Context	Development	Decision
A scout's written note	Birth month within the year	Individual objectives	What was decided
A coach's session comment	Maturity status, as a band	Progression review notes	By whom, and on what date
Video, where consent allows	Club, school and travel	Attendance and continuity	What the family was told
Who watched, and when	Other sports played	Response to coaching	When it will be revisited

What this record must never contain

- A ranking, a score or a probability attached to a child.
- Any clinical interpretation of pain, injury or growth.
- Anything a parent could not be shown on request, in plain language.

FIGURE 9 The four columns of a connected player record, and the three things it must never contain.

Athleet.AI framework, informed by references 35 to 39. ATHLEET.AI

Four kinds of information make a record useful. What was observed, with the observer named and the date attached. The context that makes the observation comparable, meaning birth month, maturity band, and the practical circumstances of access. The development picture, meaning objectives, attendance, continuity and response to coaching. And the decision, meaning what was concluded, by whom, what the family was told, and when it will be revisited.

The exclusions matter as much. No score, ranking or probability attached to a child. No clinical interpretation of pain, injury or growth, which belongs to qualified people and not to a talent record. And nothing a parent could not be shown on request, in plain language, which is a simpler test than any policy and catches most of what should not be there.

That last test does real work. A club that would be uncomfortable showing a parent a note is a club that has written something it should not have, and the discomfort arrives faster and more reliably than a compliance review.

THE TEST

Could a parent be shown the whole record, in plain language, on request, without the club being embarrassed? If not, the record is wrong.

CHAPTER EIGHT

Late developers and deselected players

Release is the most consequential judgement football makes about a person, and the one it documents least well.

Every academy releases most of the children it signs. That is arithmetic rather than failure, since the places at the top are few. What is a failure is the manner of it, and the sport's habit of retaining almost nothing afterwards.

THE TALENT RE-ENTRY LOOP

Roughly nine in ten internationally ranked juniors do not become internationally ranked seniors. A pathway with no way back is discarding most of its own future.

FIGURE 10 The talent re-entry loop. Each step is an organisational commitment rather than a technical feature.

Athleet.AI framework. ATHLEET.AI

The qualitative evidence on release is short on sample size and long on clarity. An interpretative study of four released players described release as a multi-stage process experienced as traumatic, with delayed low mood and identity difficulty, worsened by inadequate aftercare.[23] PEER-REVIEWED Interviews with eight parents across five English clubs found that families wanted direct, honest communication during the retain and release cycle and saw clubs and families as jointly responsible for managing the harm.[24] PEER-REVIEWED Both are very small samples and neither claims prevalence. Both describe something clubs recognise immediately.

A study of 12 academy players aged nine to 15 reported performance anxiety, fear of deselection and active concealment of difficulty from coaches and parents.[25] PEER-REVIEWED Research on how academy coaches enact the English development plan found work intensification, reduced autonomy and pressure to present compliance for audit.[26] PEER-REVIEWED Neither is a comfortable read and both describe the environment into which any new system would be introduced.

8.1 Building a route back

Re-entry requires four things, none of them technical. A decision recorded with its reasoning. A conversation with the family that a person owns. A genuine route back to community football that the club actively supports. And a record that per-

TWO VERY SMALL SAMPLES

The release literature rests on studies of four players and eight parents. They establish mechanism rather than prevalence, and this paper treats them accordingly.

sists so that the player can be reconsidered at 15, 16 or 18 on the strength of evidence rather than a chance sighting.

Given that roughly nine in ten internationally ranked juniors do not become internationally ranked seniors, a pathway with no way back is discarding a large share of its own future. My own view, going beyond what any study establishes, is that re-entry rate is the measure that would most improve English football's talent system if clubs were asked to publish it.

8.2 What the progression numbers actually say

Conversion statistics circulate widely and survive scrutiny poorly. A verified ten-year follow-up of 198 academy players at two Spanish clubs found 6.1 per cent became professional, of whom seven reached the top division, meaning roughly 3.5 per cent of the original cohort.[21] PEER-REVIEWED An English study of 382 academy graduates across 33 clubs measured how much they played rather than how many turned professional, and found only a weak relationship between the number of graduates a club produced and the minutes they received.[22] PEER-REVIEWED

No verified England-wide academy conversion rate appears in the peer-reviewed literature. The widely quoted round numbers could not be traced to a primary source during the preparation of this paper, and that absence is reported here rather than papered over with a plausible figure.

CHAPTER NINE

Widening access without scaling bias

A system deployed across a biased pathway makes the pathway faster. Whether it makes it fairer depends entirely on decisions taken before anything is switched on.

The optimistic reading of this technology is that it reaches players the current process misses. The pessimistic reading is that it industrialises the filters described in Chapter 2. Both are available, and which one occurs is settled by design choices rather than by the technology.

Three design decisions do most of the work. The first is refusing to train on selection outcomes, for the reason set out in Chapter 3. The second is recording context alongside every observation, so that a comparison between an August-born late maturer and a September-born early maturer can be made honestly rather than accidentally. The third is measuring the composition of the observed pool rather than only the quality of the players signed, because a process can be improving on quality while narrowing on reach.

9.1 The female pathway, and what the evidence does not cover

The wider sport-science literature carries a marked sex skew, with an analysis of 5,261 publications covering roughly 12.5 million participants finding female participants made up around 34 per cent of the total.[42] PEER-REVIEWED A review covering two decades of talent identification in football named the female pathway explicitly as under-researched.[12] SURVEY

A MEASURE WORTH ADOPTING

Report the birth-month and geographic distribution of everyone observed, not only of those signed. A pathway that will not measure its own reach cannot claim to be widening it.

The practical consequence is direct and rarely stated by suppliers. Any model in this area has been built and tested on boys. Deployed in a girls' pathway it is operating outside the population it was validated on, and it should be required to say so on screen rather than in a technical appendix.

9.2 Cost, transport and the Saturday morning

Access to the pathway is mediated by things no scout observes. Whether a family can reach a fixture 40 minutes away, whether the kit and subscription are affordable, whether a parent can attend, and whether anybody in the family already knows how the system works. The birthplace literature establishes that geography shapes who is seen,[7,8] REVIEW and the small English study on socioeconomic position suggests the same forces reach into coach judgement itself.[9] PEER-REVIEWED

Technology addresses a narrow slice of this and should not pretend otherwise. It can translate communication into the languages a community actually speaks. It can make the process legible to a family with no prior contact with football. It can record where observations happen, which makes a geographic gap visible to somebody who can act on it. It cannot pay for petrol, and a club that presents a platform as an inclusion strategy is avoiding the more expensive conversation.

CHAPTER TEN

Safeguarding children and families

Everything in this paper concerns children, which raises the bar on what may be built and lowers the tolerance for getting it wrong.

This chapter describes law, regulator guidance and safeguarding standards, and the notice at the front of this paper applies to all of it. A club should take its own advice.

The Child Protection in Sport Unit publishes standards for safeguarding and protecting children in sport, which are the framework most governing bodies build their own requirements on.[35] CONSENSUS Those standards say nothing about artificial intelligence, and when this paper was written no safeguarding body had published guidance that did. The requirements therefore have to be assembled from data protection law, which is adequate to the task. Article 35 of the UK GDPR requires a data protection impact assessment before high-risk processing begins, and names systematic and extensive automated evaluation of personal aspects, where significant decisions rest on it, as exactly such a case.[36] CONSENSUS Article 22 gives a young person the right not to be subject to a decision taken solely by automated means where it affects them significantly, together with rights to human intervention, to make representations and to contest the outcome. And because a processor may act only on the controller's documented instructions, a club that has not asked in the contract whether a supplier trains its models on club or player data has not established that it does not.

On data protection, the Age Appropriate Design Code applies to services likely to be accessed by children in the United Kingdom and sets 15 standards, among them the best interests of the child, impact assessments, transparency, default settings, data minimisation and restrictions on profiling and on nudge techniques.[37] CONSENSUS It is a statutory code rather than optional guidance, and its

BEFORE, NOT AFTER

The impact assessment is described as something to complete before deployment. A club running one after a pilot has already answered the question it was meant to ask.

restriction on profiling is the provision most directly relevant to anything that would score a young player.

10.1 Two things the industry commonly gets wrong

Two points of precision are worth stating, because both are frequently misdescribed in sales conversations. Article 9 of the UK General Data Protection Regulation covers data revealing health and biometric data used for identification. Routine football performance metrics are not automatically special category data on that definition,[38] CONSENSUS and a supplier claiming blanket special-category status for all player data is overstating the position, as is a club that assumes the reverse. On my own reading the boundary sits where a metric reveals health status, and a club should verify that with its own adviser rather than with this paper.

Second, section 80 of the Data (Use and Access) Act 2025 will replace Article 22 of the UK GDPR with new provisions on automated decision-making, defining meaningful human involvement and retaining stricter conditions where special category data drives a solely automated significant decision.[39] CONSENSUS It had not been commenced when this paper was written, so the existing Article 22 is the operative law. The section also contains no child-specific provisions, which surprises people who assume otherwise. Children's additional protection in this context comes through the Age Appropriate Design Code rather than through that section.

10.2 Who the record belongs to

A charter of player data rights developed in professional football sets out eight rights over performance and personal data, including access, portability, rectification and erasure, citing survey evidence that 80 per cent of professional players surveyed wanted access to their own performance data.[40] CONSENSUS That charter addresses professional adults. No equivalent instrument covers academy minors, which is a gap the sport should close in its own policies rather than wait for.

There is also a live legal dimension. A group action brought by current and former players concerning the processing of performance and tracking data has been running since 2020 and is documented in legal and trade commentary rather than in academic literature.[41] PRACTICE It is cited here as a real dispute that clubs should be aware of rather than as evidence of anything, and its eventual outcome may matter more to this field than any technical development.

THE REFUSALS, RESTATED

- No score, ranking or predicted potential attached to a child, in any interface, for any user.
- No model trained on the club's own past selection decisions.
- No interpretation of pain, injury, growth or body composition.
- No direct communication between a system and a child without an adult in the loop.
- No deployment before a data protection impact assessment has been completed and reviewed.

CHAPTER ELEVEN

Transitions and career durability

The pathway's transitions are where players are lost, and they are the points at which a record is worth most.

A young footballer passes through a series of transitions, each of which resets the information around them. A new age group, a new coach, a loan, a release, a move to a different club, an attempt to return after time away. At every one of those points the accumulated understanding of the player is at risk, and at most of them it is lost.

The dropout literature identifies the recurring reasons that children leave organised football, and a systematic review specific to the sport catalogues them.[27] REVIEW A meta-analysis of prospective studies across team sports provides the more methodologically rigorous synthesis.[29] REVIEW Enjoyment, perceived competence and relatedness recur across both, which is to say that most of what determines whether a child stays is created by a coach on a Tuesday rather than by a system.

Where a system helps is in noticing and in continuity. A player whose attendance halves after Christmas is signalling something, and it is more often about enjoyment, transport or a falling out than about football. A record that persists across a change of coach means the new coach inherits a player rather than a stranger. Neither requires analysis.

WHAT A SYSTEM CONTRIBUTES

Noticing. A player whose attendance halves after Christmas is telling the club something, and no club currently looks.

11.1 Alternative pathways as a designed outcome

Most children in an academy will not have a professional career, and a system built only around the ones who do has failed the great majority of its users. The alternative pathways are real and largely undersupported. Community football, coaching, officiating, the university route, and simply remaining a participant for the next 40 years are all outcomes a club should be able to point a released player towards.

My own view, which goes beyond the evidence, is that a club's re-entry and alternative-pathway rate is a better measure of its academy than its conversion rate, because conversion is dominated by factors the club does not control and re-entry is dominated by factors it does.

CHAPTER TWELVE

Value for clubs, sponsors and communities

The commercial case is real, and it is smaller and slower than the sector's usual pitch.

Four kinds of value follow from the argument in this paper, and they are worth separating because they accrue to different people over different periods.

For an academy, the value is coverage and reduced waste. Observing more players with the same staff, and retaining a record of those not taken, increases the chance of finding a player who would otherwise have been missed and reduces the chance of investing years in one who was selected mainly for being born in September. Both are plausible on the evidence in Chapter 2 and neither has been demonstrated by a controlled study.

For a community club, the value is being seen. A grassroots club contributing observations to a shared regional record acquires a standing relationship with the professional game that it currently lacks, and its players acquire a route that does not depend on a scout happening to attend the right fixture.

For a sponsor or an owner, the value is credible social impact rather than a performance return. A pathway that can demonstrate a wider observed pool, a documented re-entry route and published measures on reach is doing something a community programme cannot easily claim and cannot easily be accused of inventing.

For a governing body, the value is evidence. A federated record across a region generates the first honest picture of who the pathway sees and who it loses, which is a question the sport currently answers by anecdote.

STATE IT AS A HYPOTHESIS

No controlled study has shown that a talent-identification system improves pathway outcomes. Anyone claiming a return here is describing a hope, and should say so.

HOW TO STATE THE ECONOMICS HONESTLY

The value of finding one additional academy-standard player, or of retaining one late developer, can be modelled and should be, provided every such model is labelled as a scenario. There is no published evidence that any system produces those outcomes. A club presenting a projected return from this technology is presenting an assumption dressed as a forecast, and should be asked which of its inputs are measured.

CHAPTER THIRTEEN

A federated talent ecosystem

A single national database of children is both a governance problem and a reason clubs will decline to take part. The shape that works is federated.

The obvious architecture for a wider net is a central register of every young player in the country. It is also the wrong one, for two reasons that reinforce each other. It concentrates children's data in a single place, which is precisely what safeguarding guidance counsels against, and it asks community clubs to hand over their only real asset to organisations that have not historically given much back.

GRASSROOTS-TO-ACADEMY INFORMATION ARCHITECTURE

Community club
Owns its own records. Sees its own players, and what it chooses to share.

shares · returns

County or regional body
Holds the shared standard, the consent framework and the escalation route.

Academy
Sees what has been shared, adds its own observations, and writes decisions back.

The four rules

1. A club never loses control of its own record by contributing to a shared one.
2. Nothing is shared about a child without consent that a parent understood.
3. Every decision written into the record names the person who made it.
4. A player and their family can see everything held about them.

The federated shape matters. A single central database of children is both a governance problem and a reason clubs will decline to take part.

FIGURE 11 A federated architecture across three tiers, and the four rules that make it acceptable to the clubs at the bottom of it.

Athleet.AI framework. ATHLEET.AI

The federated shape puts the standard, the consent framework and the escalation route at regional level, leaves ownership of each record with the organisation that created it, and moves only what has been consented to. A community club keeps its own record whether or not it shares anything. An academy sees what has been shared, adds its own observations, and writes decisions back so that the club that developed a player learns what happened to them.

13.1 Documentation, and the questions it answers

Two practices from machine learning transfer directly and cost almost nothing. Model cards document a system's intended use, its performance across subgroups and its known limitations.[45] PEER-REVIEWED Datasheets document a dataset's motivation, composition, collection and maintenance.[46] PEER-REVIEWED In a youth context both should be a condition of purchase, and the subgroup section of a model card is where a supplier has to state that their system was built and tested on boys.

The argument for simplicity applies with unusual force here. An influential position in machine learning holds that post-hoc explanations of opaque models can mislead in high-stakes settings, and that inherently interpretable models should be preferred where accuracy permits.[44] PEER-REVIEWED When the stakes concern a 13-year-old and the person interpreting the output is a volunteer, a transparent rule beats a sophisticated model with a plausible explanation attached.

WHY IT HAS TO BE FEDERATED

A grassroots club's record is its only asset in this relationship. An architecture that requires it to be surrendered will be declined, and should be.

13.2 What a governing body should hold

Three assets belong at the level above the club. The shared observation standard, so that a note written in one place means the same thing in another. The consent and safeguarding framework, so that no club has to construct one alone. And the evidence base, meaning the first honest regional picture of who the pathway observes, who it signs, who it releases and who returns.

A systematic review of 70 studies on talent identification in football concluded that the advantages associated with selection are technical, tactical, anthropometric, physiological and psychological, and that they change non-linearly with age, maturity and position.[11] REVIEW No single measurement at a single moment is sufficient, which is an argument for a longitudinal record rather than for a better test.

CHAPTER FOURTEEN

Implementation roadmap

A regional pilot with an independent evaluator, measures on reach rather than only on quality, and a published result whichever way it goes.

The evidence base for artificial intelligence in talent identification is, on its own reviewers' assessment, exploratory. What the sector needs next is an evaluation rather than a rollout.

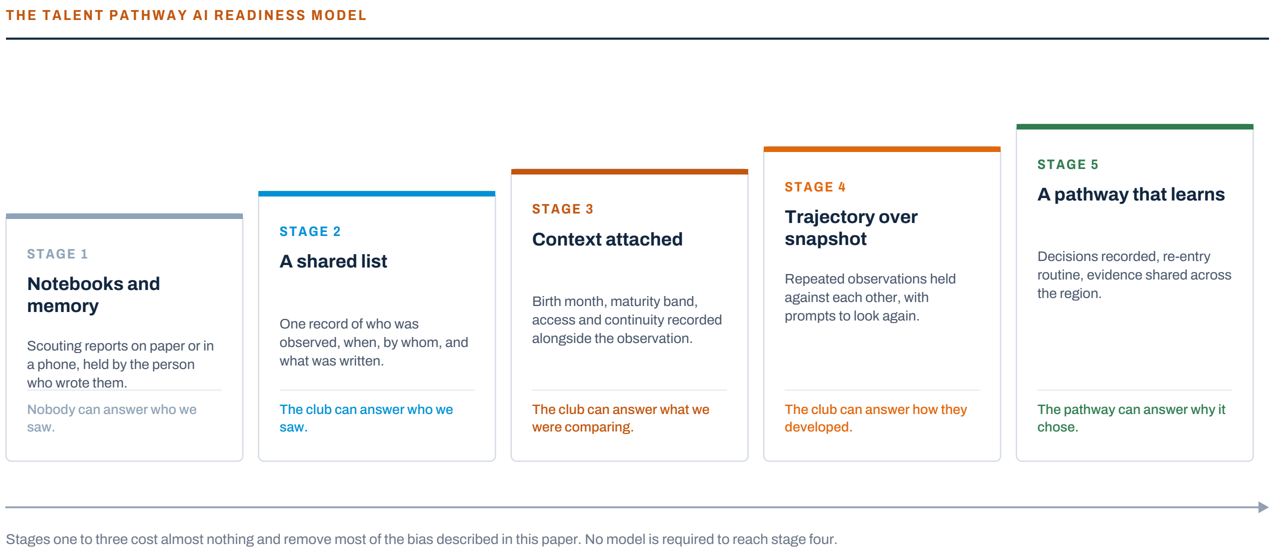


FIGURE 12 The talent pathway readiness model. Stages one to three cost almost nothing and remove most of the bias described in Chapter 2.

Athleet.AI framework. ATHLEET.AI

14.1 Where assistance may act

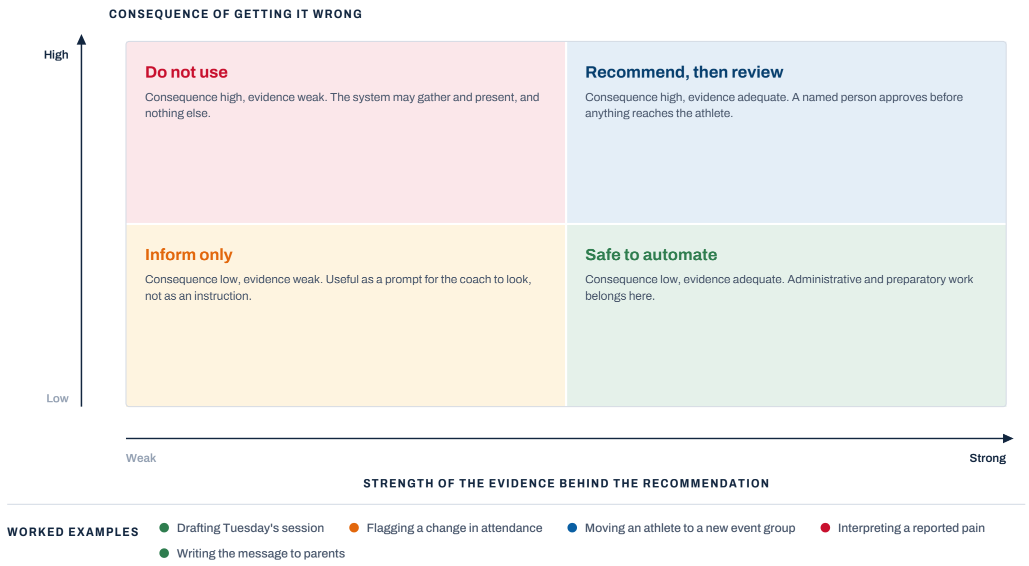


FIGURE 13 Consequence of error against strength of evidence, with youth-football examples. The upper-left quadrant is larger here than in any other paper in this series, because the consequence is higher and the evidence is thinner.

Athleet.AI framework, used across the whole series. ATHLEET.AI

14.2 A regional talent and development pilot

PROPOSED STRUCTURE

- **Before anything begins.** A data protection impact assessment completed and reviewed. A named safeguarding officer with authority to halt the pilot. A written statement of the refusals in Chapter 10, agreed by every participating organisation.
- **Season one, first half.** Records only. One academy and a network of community clubs recording observations with date, observer, birth month and maturity band. No assistance of any kind, and a baseline taken on players observed, follow-up rate and pool composition.
- **Season one, second half.** Capture, retrieval and prompting added. Scouts dictate rather than type. The system surfaces prior observations and asks about players flagged and not revisited. Still no scores and no rankings.
- **Season two.** Re-entry routine established, with released players remaining in the shared record by consent and a scheduled reassessment. An independent evaluation partner reports at the end.

The measures should expose failure rather than demonstrate success. Players observed and follow-up rate test capacity. The composition of the observed pool tests reach. Retention tests whether anything improved for those already inside, and re-entry rate tests whether the pathway acquired a memory.

14.3 What good looks like in three years

A region that has done this well will look much as it does now. Most of the children will still not become professional footballers, because that is arithmetic. What will have changed is that the club can say who it observed rather than only who it signed, that a judgement made at 12 is still readable at 16, and that a released player has a route back. The strongest thing this technology can do for a

PUBLISH EITHER WAY

An evaluation published only when it flatters the technology is a case study. The distinction matters more in youth sport than anywhere else in this series.

young footballer is to make sure the sport can still find them after deciding, wrongly, that they were not good enough.

END MATTER

Pathway readiness checklist

Sixteen things to settle before anything is deployed. The first eight cost nothing and remove most of the bias in this paper.

- [ ] **Every observation dated and attributed**
  Who watched, when, in what fixture, in what conditions.
- [ ] **Birth month recorded with every judgement**
  So that a comparison can be read fairly two years later.
- [ ] **Maturity recorded as a band, never as a number**
  The estimation error is large and its direction is systematic.
- [ ] **The record kept when a player is not taken**
  This is the single change that would most improve the English pathway.
- [ ] **A written rationale for every retain and release decision**
  Including what the family was told and when it will be revisited.
- [ ] **A scheduled reassessment for released players**
  At 15, at 16, at 18, prompted by the record rather than by chance.
- [ ] **Total weekly football across all organisations**
  Asked as a question each term and written down. No sensor required.
- [ ] **Composition of the observed pool reported**
  Birth month and geography of everyone seen, not only of those signed.
- [ ] **A data protection impact assessment completed before deployment**
  Reviewed, and repeated when the system changes.
- [ ] **A named safeguarding officer with authority to halt**
  Not consulted afterwards, and able to stop without a committee meeting.
- [ ] **The refusals written into policy**
  No score, no ranking, no prediction, no training on past selections, no direct contact with a child.
- [ ] **A contractual answer on third-party model training**
  In the contract rather than the sales meeting.
- [ ] **A route for a family to see, correct and remove**
  Operable by the club, in plain language, without contacting the supplier.
- [ ] **A model card from any supplier**
  Stating the population the system was built and tested on, including its sex.
- [ ] **An honest count of useless prompts**
  Reported alongside the useful ones, or the tool will be abandoned quietly.
- [ ] **A written definition of a disappointing result**
  Agreed before the pilot starts, by the people who will want to redefine success at the end.

END MATTER

Glossary

Terms used in a specific sense in this paper, defined once and held to throughout.

Bio-banding
Grouping young players by biological maturity rather than chronological age for training or competition.

Birthplace effect
The over-representation among professional players of those raised in communities of a particular size, associated with density and proximity rather than facility quality.

Deselection
The decision not to retain a player in a pathway. Used in this paper in preference to release when the emphasis is on the judgement rather than the event.

Development trajectory
The rate and shape of a player's improvement over time, which is a different signal from their current level and invisible in a single observation.

Federated architecture
An arrangement in which each organisation retains its own records and shares only what has been consented to, with the shared standard held above them.

Label
The outcome a model is trained to predict. In talent identification the only readily available labels are past human selection decisions, which is the central problem described in Chapter 3.

Maturity band
A coarse category describing how far through biological maturation a young player is, used in this paper in preference to any numeric estimate.

Peak height velocity
The point of fastest growth in adolescence. Estimated non-invasively with substantial individual error and a systematic direction to that error.

Re-entry
A designed route by which a player not selected can be reconsidered later, on the strength of a surviving record rather than a chance sighting.

Relative age effect
The advantage held by children born early in a selection year, largest in the youngest age groups and washing out by adulthood.

Responsible AI
Systems designed so that their limits are visible, their outputs are contestable, and accountability rests with an identified person.

END MATTER

Method and evidence grading

How the sources were selected and checked, what could not be verified, and what the reader should distrust.

Sources were gathered from the peer-reviewed literature, governing-body publications, regulator guidance and primary legislation, with priority given to systematic reviews and meta-analyses over single studies. Every reference below was checked against the publishing journal, publisher or issuing body, and verified independently a second time before publication. Where a figure could not be traced to its original source it was omitted rather than approximated.

What could not be verified

Four things were searched for and not confirmed, and their absence shapes this paper. No England-wide academy-to-professional conversion rate appears in the peer-reviewed literature, so the widely circulated round numbers are not used and a Spanish cohort is cited and labelled as Spanish. The Football Association report behind the participation figures in Chapter 1 could not be located, so those figures are cited through a parliamentary library briefing and flagged in the text as being at one remove. No peer-reviewed account of the players' data group action exists, so legal and trade commentary is cited and labelled as such. And no evaluated deployment of an artificial intelligence tool in an English talent pathway appears in the published literature.

Two points of law commonly stated wrongly

Article 9 of the UK General Data Protection Regulation covers data revealing health and biometric data used for identification. Routine performance metrics do not automatically fall within it, and this paper does not claim they do. Section 80 of the Data (Use and Access) Act 2025 contains no child-specific provisions, it had not been commenced when this paper was written, and children's additional protection in this area comes through the Age Appropriate Design Code. Both points were checked during preparation because both are commonly asserted the other way.

What to distrust in this paper

Three things. The capacity multiplier in Chapter 6, the trajectory curves in Chapter 4 and the growth bands in Chapter 5 are illustrative models rather than measured data, and are labelled as such in their captions. The release literature rests on studies of four players and eight parents, which establish mechanism rather than prevalence. And the geographic selection finding in Chapter 2 comes from a cross-sport programme in another country and is cited as corroborating evidence rather than as a football result.

Conflict of interest

The author is the founder of Athleet.AI, which builds software in the area this paper examines. The paper argues against building the product most obviously saleable into this market, which is a system that ranks young players. Readers should weigh the argument accordingly and are encouraged to test the citations rather than the author.

END MATTER

About the author

Paul Forrest is the founder and Chief AI Officer of Athleet.AI. He works with organisations on the deployment of artificial intelligence, having begun in business process reengineering in the 1990s and moved into machine learning and language model work from 2014, building small domain-specific models in preference to general-purpose ones.

He is a qualified coach and coaches young athletes as a volunteer at club level. Watching children sorted at speed, on evidence nobody records, is the reason this paper exists. He is the author of *Tales from AI Development Hell* and *AI Made the Decision: Who Answers When Nobody Owns the Outcome?*, and holds an adjunct professorship teaching digital acumen and digital entrepreneurship.

END MATTER

References

Forty-six sources, each verified twice against the publishing journal, publisher or issuing body.

The hidden filters in the pathway

1. **Helsen, W. F., Van Winckel, J., & Williams, A. M.** (2005). The relative age effect in youth soccer across Europe. *Journal of Sports Sciences*, 23(6), 629–636. doi: 10.1080/02640410400021310
2. **Cobley, S., Baker, J., Wattie, N., & McKenna, J.** (2009). Annual age-grouping and athlete development: a meta-analytical review of relative age effects in sport. *Sports Medicine*, 39(3), 235–256. doi:10.2165/00007256-200939030-00005
3. **Kelly, A. L., Wilson, M. R., Gough, L. A., Knapman, H., Morgan, P., Cole, M., Jackson, D. T., & Williams, C. A.** (2020). A longitudinal investigation into the relative age effect in an English professional football club: exploring the 'underdog hypothesis'. *Science and Medicine in Football*, 4(2), 111–118. doi: 10.1080/24733938.2019.1694169
4. **Doncaster, G., Kelly, A. L., McAuley, A. B. T., Cain, A., Partington, M., Nelson, L., & O'Gorman, J.** (2024). Relative age effects and the Premier League's Elite Player Performance Plan (EPPP): a comparison of birthdate distributions within and between age groups. *Journal of Science in Sport and Exercise.* Advance online publication, 31 May 2024. doi:10.1007/s42978-024-00285-w
5. **Curnyn, S., Leslie, M., Palmer, D., Williams, S., & Cumming, S.** (2025). The influence of relative age and biological maturation on player selection in the Scottish Football Association's Club Academy Scotland. *Journal of Sports Sciences*, 43(18), 1980–1991. doi:10.1080/02640414.2025.2527436
6. **Sweeney, L., Cumming, S. P., MacNamara, Á., & Horan, D.** (2023). A tale of two selection biases: the independent effects of relative age and biological maturity on player selection in the Football Association of Ireland's national talent pathway. *International Journal of Sports Science & Coaching*, 18(6), 1992–2003. doi:10.1177/17479541221126152
7. **Côté, J., Macdonald, D. J., Baker, J., & Abernethy, B.** (2006). When "where" is more important than "when": birthplace and birthdate effects on the achievement of sporting expertise. *Journal of Sports Sciences*, 24(10), 1065–1073. doi:10.1080/02640410500432490
8. **Hernández-Simal, L., Calleja-González, J., Lorenzo Calvo, A., & Aurrekoetxea-Casaus, M.** (2024). Birthplace effect in soccer: a systematic review. *Journal of Human Kinetics*, 94, 227–242. doi:10.5114/jhk/186935
9. **Kelly, A. L., Williams, C. A., Jackson, D. T., Turnnidge, J., Reeves, M. J., Dugdale, J. H., & Wilson, M. R.** (2024). Exploring the role of socioeconomic status and psychological characteristics on talent development in an English soccer academy. *Science and Medicine in Football*, 8(3), 251–259. doi: 10.1080/24733938.2023.2213191
10. **Bradley, B., Johnson, D., Hill, M., McGee, D., Kana-ah, A., Sharpin, C., Sharp, P., Kelly, A., Cumming, S. P., & Malina, R. M.** (2019). Bio-banding in academy football: player's perceptions of a maturity-matched tournament. *Annals of Human Biology*, 46(5), 400–408. doi:10.1080/03014460.2019.1640284

What the talent identification evidence supports

1. **Sarmento, H., Anguera, M. T., Pereira, A., & Araújo, D.** (2018). Talent identification and development in male football: a systematic review. *Sports Medicine*, 48(4), 907–931. doi:10.1007/s40279-017-0851-7
2. **Williams, A. M., Ford, P. R., & Drust, B.** (2020). Talent identification and development in soccer since the millennium. *Journal of Sports Sciences*, 38(11–12), 1199–1210. doi:10.1080/02640414.2020.1766647
3. **Bergkamp, T. L. G., Meijer, R. R., den Hartigh, R. J. R., Frencken, W. G. P., & Niessen, A. S. M.** (2022). Examining the reliability and predictive validity of performance assessments by soccer coaches and scouts: the influence of structured collection and mechanical combination of information. *Psychology of Sport and Exercise*, 63, 102257. doi:10.1016/j.psychsport.2022.102257
4. **Bergkamp, T. L. G., Frencken, W. G. P., Niessen, A. S. M., Meijer, R. R., & den Hartigh, R. J. R.** (2022). How soccer scouts identify talented players. *European Journal of Sport Science*, 22(7), 994–1004. doi:10.1080/17461391.2021.1916081
5. **Barth, M., Güllich, A., Macnamara, B. N., & Hambrick, D. Z.** (2024). Quantifying the extent to which junior performance predicts senior performance in Olympic sports: a systematic review and meta-analysis. *Sports Medicine*, 54(1), 95–104. doi:10.1007/s40279-023-01906-0
6. **Güllich, A., Barth, M., Macnamara, B. N., & Hambrick, D. Z.** (2023). Quantifying the extent to which successful juniors and successful seniors are two disparate populations: a systematic review and synthesis of findings. *Sports Medicine*, 53(6), 1201–1217. doi:10.1007/s40279-023-01840-1
7. **Barth, M., Güllich, A., Macnamara, B. N., & Hambrick, D. Z.** (2022). Predictors of junior versus senior elite performance are opposite: a systematic review and meta-analysis of participation patterns. *Sports Medicine*, 52(6), 1399–1416. doi:10.1007/s40279-021-01625-4
8. **Fortin-Guichard, D., Huberts, I., Sanders, J., van Elk, R., Mann, D. L., & Savelsbergh, G. J. P.** (2022). Predictors of selection into an elite level youth football academy: a longitudinal study. *Journal of Sports Sciences*, 40(9), 984–999. doi:10.1080/02640414.2022.2044128
9. **Baker, J., Cattle, A., McAuley, A., Kelly, A., & Johnston, K.** (2025). Will artificial intelligence solve the riddle of athlete development? A critical review of how AI is being used for athlete identification, selection and development. *Psychology of Sport and Exercise.* Advance online publication, 15 September 2025. doi:10.1016/j.psychsport.2025.102978
10. **Sartor, F., Rosito, S. A., Raimo, M., Beato, M., & Bertollo, M.** (2025). Artificial intelligence in sport psychology: implications for the identification and development of talent. *Psychology of Sport and Exercise.* Advance online publication, 15 September 2025. doi:10.1016/j.psychsport.2025.102977

Academy progression, release and dropout

1. **Moran, J., Cervera Raga, V., Jones, B., Tallent, J., Howe, L., Clemente, F. M., Arede, J., & Freeman, P.** (2024). The identification and development of young talent in Spanish soccer academies: a 10-year multi-study follow-up. *International Journal of Sports Science & Coaching*, 19(5), 1984–1994. doi: 10.1177/17479541241254767
2. **Balassiano, V., & Bullough, S.** (2021). Measures of academy productivity in English Championship clubs. *Team Performance Management*, 27(3/4), 332–350. doi:10.1108/TPM-09-2020-0083
3. **McGlinchey, T. R., Saward, C., Healy, L. C., & Sarkar, M.** (2022). "From everything to nothing in a split second": elite youth players' experiences of release from professional football academies. *Frontiers in Sports and Active Living*, 4, 941482. doi:10.3389/fspor.2022.941482
4. **Giles, L. E. L., Harwood, C. G., & Rongen, F.** (2025). The retain-release process in elite youth football academies: parent perceptions and recommendations for minimising harm. *Journal of Applied Sport Psychology*, published online, 1–21. doi:10.1080/10413200.2025.2543759
5. **Sothern, N. A., & O'Gorman, J.** (2021). Exploring the mental health and wellbeing of professional academy footballers in England. *Soccer & Society*, 22(6), 641–654. doi:10.1080/14660970.2021.1952693
6. **O'Gorman, J., Partington, M., Potrac, P., & Nelson, L.** (2021). Translation, intensification and fabrication: professional football academy coaches' enactment of the Elite Player Performance Plan. *Sport, Education and Society*, 26(3), 309–325. doi:10.1080/13573322.2020.1726313
7. **Temple, V. A., & Crane, J. R.** (2016). A systematic review of drop-out from organized soccer among children and adolescents. *Soccer & Society*, 17(6), 856–881. doi:10.1080/14660970.2015.1100901
8. **Doyle, J. R., & Bottomley, P. A.** (2018). Relative age effect in elite soccer: more early-born players, but no better valued, and no paragon clubs or countries. *PLOS ONE*, 13(2), e0192209. doi:10.1371/journal.pone.0192209
9. **Back, J., Johnson, U., Svedberg, P., McCall, A., & Ivarsson, A.** (2022). Drop-out from team sport among adolescents: a systematic review and meta-analysis of prospective studies. *Psychology of Sport and Exercise*, 61, 102205. doi:10.1016/j.psychsport.2022.102205

Coach and scout capacity

1. **House of Lords Library** (2024). *Non-league and grassroots football: what is the state of play?* London, House of Lords Library, October 2024. Reports figures attributed to the Football Association of 15.7 million people playing football in England, more than a million volunteers and roughly 18,000 accredited clubs. The underlying Football Association report could not be located during preparation, so these figures are cited at one remove and should be treated accordingly.
2. **Lüdin, D., Donath, L., & Romann, M.** (2023). Disagreement between talent scouts: implications for improved talent assessment in youth football. *Journal of Sports Sciences*, 41(8), 758–765. doi:10.1080/02640414.2023.2239614
3. **Mann, D. L., Fortin-Guichard, D., & Müller, D.** (2025). A framework of cognitive biases that might influence talent identification in sport. *International Review of Sport and Exercise Psychology*, published online, 1–28. doi:10.1080/1750984X.2025.2556393
4. **Lawlor, C., Rookwood, J., & Wright, C.** (2021). Player scouting and recruitment in English men's professional football: opportunities for research. *Journal of Qualitative Research in Sports Studies*, 15(1), 57–76.
5. **Dempsey, N. M., Richardson, D. J., Cope, E., & Cronin, C. J.** (2021). Creating and disseminating coach education policy: a case of formal coach education in grassroots football. *Sport, Education and Society*, 26(8), 917–930. doi:10.1080/13573322.2020.1802711

Safeguarding, children's data and regulation

1. **NSPCC Child Protection in Sport Unit** (second edition; first published 2002, revised 2018). *Standards for Safeguarding and Protecting Children in Sport.* Leicester, CPSU. Ten standards.
2. **United Kingdom General Data Protection Regulation**, Articles 22, 28 and 35. The right not to be subject to a decision based solely on automated processing where it produces legal or similarly significant effects; the requirement that a processor act only on the controller's documented instructions; and the requirement to complete a data protection impact assessment prior to high-risk processing, including systematic and extensive automated evaluation on which significant decisions are based.
3. **Information Commissioner's Office** (2020). *Age Appropriate Design Code: A Code of Practice for Online Services.* Wilmslow, ICO. Published 27 January 2020; in force 2 September 2020 after a 12-month transition. Fifteen standards.
4. **United Kingdom General Data Protection Regulation**, Article 9. Processing of special categories of personal data, covering data revealing health and biometric data used for unique identification. Routine performance metrics do not fall within it automatically.
5. **Data (Use and Access) Act 2025** (c. 18), section 80. Will replace Article 22 of the UK GDPR with new Articles 22A to 22D on automated decision-making. Royal Assent 19 June 2025; section 80 not commenced at the date of this paper. Contains no child-specific provisions.
6. **FIFPRO & FIFA** (2022). *Charter of Player Data Rights.* Hoofddorp, FIFPRO. Published 19 September 2022, setting out eight rights over performance and personal data.
7. **Project Red Card** (2020 onwards). Group action by current and former professional footballers concerning the processing of performance and tracking data. Documented in legal and trade commentary; no peer-reviewed account exists.
8. **Cowley, E. S., Olenick, A. A., McNulty, K. L., & Ross, E. J.** (2021). "Invisible sportswomen": the sex data gap in sport and exercise science research. *Women in Sport and Physical Activity Journal*, 29(2), 146–151. doi:10.1123/wspaj.2021-0028

Algorithmic bias, interpretability and documentation

1. **Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S.** (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, 366(6464), 447–453. doi:10.1126/science.aax2342
2. **Rudin, C.** (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, 1(5), 206–215. doi:10.1038/s42256-019-0048-x
3. **Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., Spitzer, E., Raji, I. D., & Gebru, T.** (2019). Model cards for model reporting. *Proceedings of the Conference on Fairness, Accountability, and Transparency*, 220–229. doi:10.1145/3287560.3287596
4. **Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Daumé III, H., & Crawford, K.** (2021). Datasheets for datasets. *Communications of the ACM*, 64(12), 86–92. doi:10.1145/3458723

WHERE THIS GOES NEXT

A regional talent and development pilot

The argument in this paper is testable, and the only honest way to test it is across a professional academy and a network of community clubs, with an independent evaluator and a published result.

Athleet.AI works with academies, community clubs and county associations to run a two-season cohort pilot using the design in Chapter 14. The coach, scout, player and parent experiences are parts of one governed environment rather than separate applications, every output passes through a named person, and the refusals in Chapter 10 are written into the deployment rather than left to a setting.

What a pilot involves

- One academy, a network of community clubs and an independent evaluation partner.
- A first half-season of records only, so the gain from recording properly is separated from any gain from the technology.
- A data protection impact assessment completed before anything begins, and a named safeguarding officer with authority to halt.
- Measures on players observed, follow-up rate, the composition of the observed pool, retention and re-entry.
- No score, no ranking and no prediction attached to any child, at any point, for any user.

Athleet.Ai

Casting the Net Wider · Paper Four of Seven · Athleet.AI Performance Intelligence Whitepaper Series
Paul Forrest · October 2025 · Version 1.0