



The Measured Child

Datafication, Consent and Autonomy Across the Athlete Pathway

About this paper

This paper examines what happens to the data collected about children in selective sport pathways once artificial intelligence (AI) tools can turn that data into profiles and predictions. It is written for the boards, safeguarding leads, performance directors and data protection advisers of clubs and national governing bodies, and for the coaches who carry out the measuring. It asks one question. How can a pathway that measures athletes from early adolescence keep the child's consent meaningful, and keep control over what is inferred about them, across the years in which that child becomes an adult?

The paper belongs to a strand of working papers whose single argument is that AI in coaching and athlete development must respect the non-linear, human and developmental nature of athletic progress. Children are where that argument is tested hardest, because the measured person is still changing, has the least power in the room, and will live longest with whatever the data are later taken to mean. The paper touches data protection law and the positions of regulators. It reports what those sources say and attributes each conclusion to its source, and it offers no view on what the law requires of any particular organisation.

Editorial balance. Roughly two fifths of the paper summarises published evidence and official guidance, about a third is my own analysis of how that material applies to youth pathways, and the remainder, concentrated in the final chapter, sets out a framework of my own that is proposed for discussion and has not been tested.

Legal and regulatory notice. Parts of this paper describe legislation, regulation, regulator activity or court decisions. 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. Statements about what a source says are attributed to that source. Where I draw a practical implication that a source does not state, the text says so.

HOW TO READ THE EVIDENCE TAGS

Every substantive claim carries a tag stating what kind of evidence supports it. Two tags are combined where a claim rests on both.

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

REVIEW A systematic review, meta-analysis or narrative review.

CONSENSUS A governing body position, consensus statement or official guidance.

PREPRINT A working paper or preprint not yet peer reviewed.

INDUSTRY A company announcement, press report or industry source.

SCHOLARLY A scholarly book, edited volume or law journal article.

PRACTICE Practitioner opinion, including the author's own.

EVIDENCE GAP The available evidence does not answer the question.

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

CONTENTS

Contents

FRONT MATTER

A note from the author	5
Executive summary	6

PART ONE · THE MEASURED ATHLETE

1 The screening day	9
2 From monitoring to datafication	11

PART TWO · WHY THE CHILD IS DIFFERENT

3 Why children are different	14
4 Inferred data as the real risk	16
5 Consent that cannot be informed	18

PART THREE · THE ADULTHOOD REVIEW

6 A framework for the measured child	21
--------------------------------------	----

END MATTER

References	24
About the author	25

ILLUSTRATIVE MATERIAL

Every organisation, person and scenario described as illustrative in this paper is invented. Published research and official documents are cited and numbered, and invented material is never presented as evidence.

Research closed 17 April 2023. Sources and events are as they stood at that date.

CITATION

Forrest, P. (2023) *The Measured Child: Datafication, Consent and Autonomy Across the Athlete Pathway*. Athleet.AI Working Papers, AI in Coaching and Athlete Development, Paper 3. Athleet.AI, 19 April 2023. Available at: athleet.ai.

Comments and corrections are welcome and will be acknowledged in later versions.

PAPERS IN THIS STRAND

- Paper one. *Predicting the Unpredictable*. Why Algorithmic Talent Identification Struggles With Non-Linear Athlete Development. 25 January 2023.
- Paper two. *The Coach in the Loop*. Tacit Knowledge, Professional Judgement and the Risk of Deskilled Coaching. 27 February 2023.
- **Paper three. *The Measured Child*. Datafication, Consent and Autonomy Across the Athlete Pathway. 19 April 2023.**
- Paper four. *Maturation, Bias and the Training Set*. Developmental Fairness in AI-Supported Athlete Development. 24 May 2023.
- Paper five. *Load, Readiness and False Precision*. The Limits of Algorithmic Training Prescription in Long-Term Endurance Development. 21 June 2023.
- Paper six. *Democratised or Stratified?* AI, Competitive Fairness and the

Grassroots Coaching Gap. 26 July 2023.

A NOTE FROM THE AUTHOR

Why I stopped asking parents to sign the form without reading it

For a long stretch of my coaching life, the consent form was the last thing that happened before the testing itself began. A parent signed it on a clipboard by the door, usually while talking to their child, and the session moved on to sprint gates and jump mats. I did not think of that signature as a decision anyone had made. It was a ticket.

What changed my view was a spreadsheet. Some years ago I was asked to help tidy a set of historical testing records, and I found myself looking at the numbers of athletes I had coached as young teenagers who were now adults with jobs and children of their own. Some of the columns were ones I had added. Others had been derived later by people I had never met, and a few of them (a maturity estimate, a note about "attitude") were things I would not have wanted written about me at that age. None of those athletes had any idea the file still existed.

In my advisory work since, I have seen the same pattern in more polished form. Organisations collect in good faith, store by default and reuse because the data are there. The arrival of AI tools that can read years of records at once, and produce confident scores from them, makes that pattern more consequential, and it is happening in sport faster than the guidance for children has moved.

I am not a lawyer, and this paper is not legal advice. It is an attempt to set out, for the people who run pathways, why the child's data deserve different treatment from the adult's, and what a pathway could do about it before a regulator or a former athlete asks. The proposal in the final chapter is mine, and I expect it to be improved by people who know the law and the sports better than I do.

Paul Forrest
Founder, Athleet.AI
19 April 2023

TWO PAGES THAT CIRCULATE ON THEIR OWN

Executive summary

Youth sport pathways now measure athletes from early adolescence, and AI tools can turn those measurements into lasting predictions about maturity, injury and potential that the child never agreed to and may never see. Consent given at 13 cannot cover uses nobody can yet describe, so pathways should collect less, treat inferred data as a separate and more tightly limited category, and give every athlete a point at adulthood to review and withdraw what was gathered about them as a child.

The ethics of athlete monitoring has been studied in adult professional sport, and the ethics of children's data has been addressed in general digital policy. The two have barely met. This paper brings them together. Drawing on research into adult monitoring, the science of growth and maturation, and the guidance that regulators and international bodies have issued on children and AI, it argues that the greatest risk to young athletes lies in what is inferred from their data, and that current consent practice in pathways cannot carry the weight placed on it.

1

Adult sport already shows that athlete monitoring can compromise privacy, autonomy and well-being.

Karkazis and Fishman describe how biometric tracking in United States professional leagues raises problems of data validity, privacy, data security, conflicts of interest and coercion.¹ **PEER-REVIEWED** A Foucauldian analysis of wearable GPS in professional rugby league argues that routine tracking can become a normalised form of discipline that athletes turn on themselves.² **PEER-REVIEWED** Children in pathways face the same mechanisms with fewer defences.

3

Inferred data shape decisions about a child's future, yet they receive the least scrutiny of anything a pathway holds.

Wachter and Mittelstadt argue that data protection frameworks give inferences weaker protection than the data used to produce them.⁵ **SCHOLARLY** In a pathway, an inferred "potential" score can shape selection and contracts long after the raw sprint time has lost relevance. **PRACTICE**

5

Anonymisation offers little protection in small, visible elite cohorts.

A modelling study estimated that 99.98% of Americans would be correctly re-identified in any dataset using 15 demographic attributes.⁷ **PEER-REVIEWED** A regional squad of 20 teenagers, whose ages and events are public, is far easier to re-identify than a national population. **PRACTICE**

2

A young athlete's body is still changing, so any measurement describes a moment and predicts poorly.

Children of the same chronological age can differ widely in biological maturity, and the timing and tempo of maturation vary from child to child.³

SCHOLARLY A review of maturity assessment in youth athletes cautions that common prediction methods carry errors large enough to matter for individual decisions.⁴ **REVIEW**

4

Consent from a child, or from a parent on the child's behalf, cannot be informed about uses that do not yet exist.

The General Data Protection Regulation (GDPR) recognises that children may be less aware of the risks of processing.⁶ **CONSENSUS** In a selective pathway, declining to be measured may carry a cost to selection, which leaves the choice less free than the signature suggests. **PRACTICE**

6

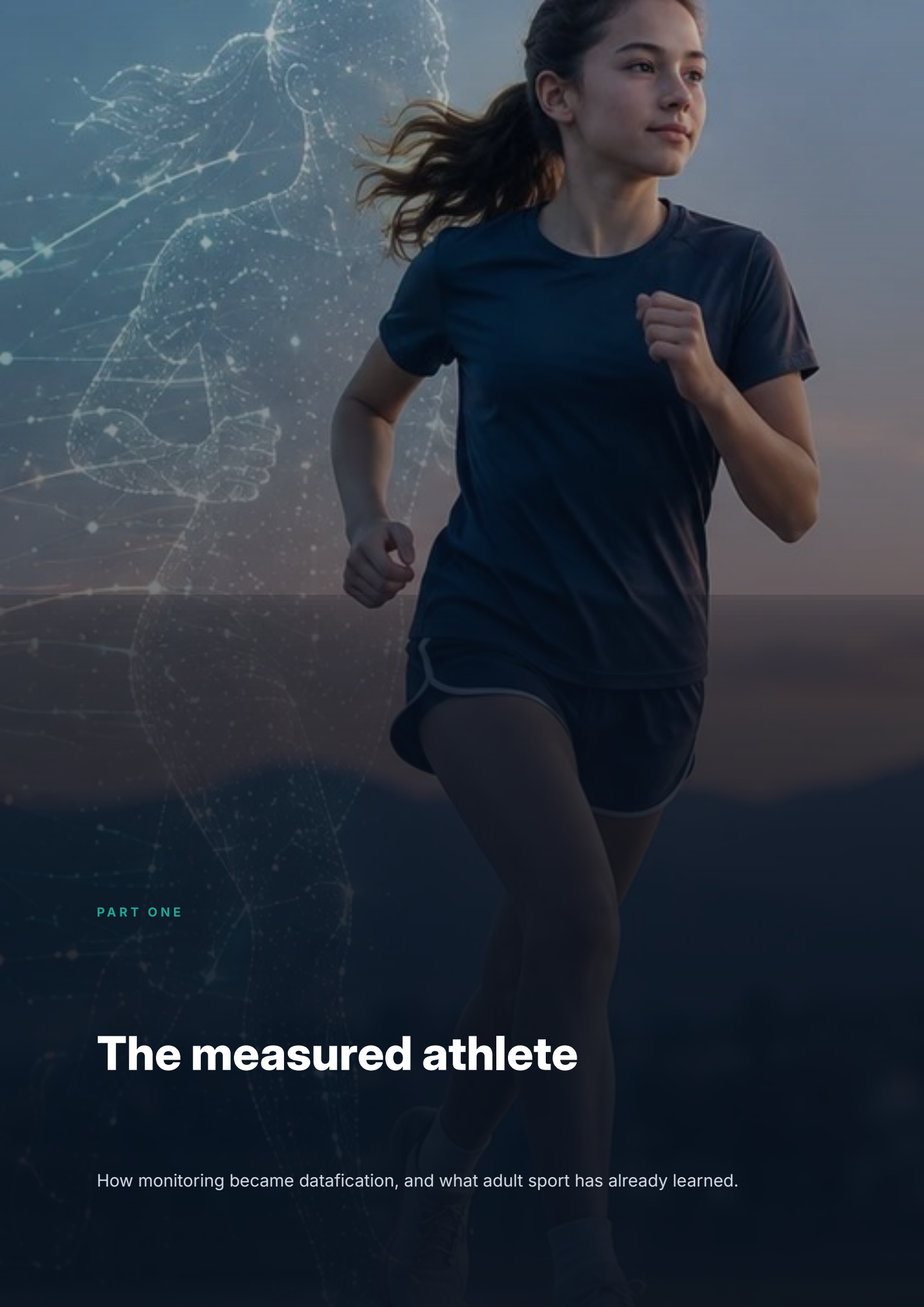
Regulators are beginning to act on AI and children, but nothing yet addresses youth athlete data directly.

Italy's data protection authority restricted the Replika chatbot in February 2023 and ChatGPT in March 2023, citing among other things the risks to minors.^{8 9} **CONSENSUS** The FIFPRO Charter of Player Data Rights covers professional footballers only.¹⁰ **INDUSTRY** No comparable charter for youth athletes has been published. **EVIDENCE GAP**

What follows for a club or governing body

Pathway boards should treat youth athlete data as a long-lived liability as well as a coaching asset. The practical response proposed here has four parts. Collect only what a named coaching purpose requires, hold inferred data apart from measurements with stricter limits on who sees them and for how long, set retention periods that end, and build an adulthood review into the pathway so that each athlete can see and withdraw what was gathered about them before they turned 18.

All of it can be done with existing technology, given a governing body willing to write the rules before a commercial AI tool writes them by default, and a data protection adviser engaged early enough to test each step against the organisation's own obligations.



PART ONE

The measured athlete

How monitoring became datafication, and what adult sport has already learned.

CHAPTER ONE

The screening day

A child measured at 13 may still be carrying that measurement, in altered and enlarged form, at 21.

Talent pathways in many Olympic and professional sports now begin formal testing in early adolescence, and the data gathered on those first days rarely disappears.

PRACTICE What has changed is scale. Analytics tools, and more recently general-purpose AI systems, can read a child's history of measurements and return a score, a flag or a ranking that nobody in the room calculated and few could explain.

ILLUSTRATIVE SCENARIO**A first screening day and what follows**

Maya is 12 when her athletics coach nominates her for a regional pathway, and she turns 13 the week before its screening day. She is measured for standing and sitting height and body mass, runs a 20 m sprint, completes a jump test and fills in a short questionnaire about sleep and mood. Her father is asked for both parents' heights so that a predicted adult height can be calculated. He signs a form on a clipboard that mentions "performance monitoring, athlete development and research".

Over the next four years the pathway adds GPS training loads, a daily wellness app and a subscription analytics tool. The tool combines her records into a maturity-adjusted "potential" rating and an injury risk flag, and the pathway shares both with a partner academy when she is released at 16. At 21 Maya applies for a performance support role in sport. She has never seen the rating, does not know the partner academy holds it, and has no idea that the wellness entries she typed as a tired 14-year-old still sit on a vendor's server.

Nothing in the scenario is exotic, and nobody in it acts in bad faith. The coach nominated a promising athlete, the pathway measured what pathways measure, the father signed what he was given, and the vendor did what its product was sold to do. My concern is with what that ordinary sequence produces. A record created for coaching becomes a profile, the profile acquires inferences, and the inferences travel to people and purposes that no one at the screening day could have named.

1.1 The question this paper asks

The paper asks how a pathway can keep consent meaningful, and keep control over what is inferred, across the span from early adolescence to adulthood. Two bodies of work bear on it. The first is the ethics of monitoring in adult professional sport, where researchers have documented privacy, ownership and coercion problems in some depth.¹ **PEER-REVIEWED** The second is children's digital rights, where regulators and international bodies have set out detailed expectations for services that children use,

largely with online platforms and apps in mind.^{11 12} **CONSENSUS** Youth pathways sit in the space between them, and to my knowledge no governing body has yet published guidance written for that space. **EVIDENCE GAP**

My answer, developed across the chapters that follow, is that the most serious risk to the measured child lies in the inferences drawn from their data, that the consent pathways currently rely on cannot bear the weight placed on it, and that the remedy is structural. Pathways should collect less, separate inferred data from measured data, set limits that expire, and give each athlete a formal review at adulthood.

1.2 How to read what follows

Because the paper draws on legislation, regulators' decisions and official guidance as well as research, its claims differ in strength, and a reader deciding what to act on needs to know which kind of claim is in front of them. Firm conclusions are reported as the source's own. Where a source says the answer depends on the facts, I carry that qualification across. Where I draw a practical implication that the source does not state, I say that the view is mine.

A related convention runs through the paper. The evidence tags after each claim distinguish peer-reviewed research, reviews, official guidance, industry material, practitioner observation and gaps in the evidence, and the final chapter's proposal carries its own tag because it is a framework I am proposing and have not tested.

TO CONFIRM

Margin notes headed in this way mark a point where an organisation's own legal position would need checking by its data protection adviser or lawyer. They say what to check, and they do not suggest the answer.

CHAPTER TWO

From monitoring to datafication

Monitoring was once something a coach did to support a session. It has become a continuous record that exists whether or not anyone is coaching.

The shift matters because a record outlives its purpose. When a coach timed a sprint with a stopwatch and wrote the result in a notebook, the number served that week's decision and then faded. A sensor feed stored on a platform, by contrast, accumulates, and its value to others grows with its length.

2.1 The continuously measured subject

Lupton's sociology of self-tracking describes how people come to understand themselves through data, and it distinguishes modes of tracking by who initiates them and who benefits.¹³ **SCHOLARLY** Tracking that a person chooses for their own ends sits at one end of her account. At the other sit tracking that institutions impose and tracking whose data are exploited by others for commercial or institutional gain. Read against that scheme, a youth pathway's monitoring programme is imposed from the start, because selection depends on taking part, and it can drift towards exploitation when records are shared with partners or vendors. That reading is mine, and it applies Lupton's categories to a setting she did not examine.

Datafication, as I use the term here, means the conversion of an athlete's training, body and mood into records that can be stored, combined and analysed without the athlete present. The term matters because it names a change in who holds the athlete's story. A coach remembers a teenager's difficult winter in context. A dataset holds the dip in wellness scores and the missed sessions, and any system that reads it later will see only those.

2.2 What adult sport has already learned

Karkazis and Fishman examined biometric tracking across professional leagues in the United States. They identified a cluster of ethical problems that the technology's promoters had largely passed over, and their analysis remains the most thorough treatment of the subject I know.¹ **PEER-REVIEWED** They question whether many devices measure what they claim to, they ask who owns and controls the data, and they point to the risk that teams use physiological records in contract decisions. Their account also raises coercion, since an athlete employed by a club has limited room to refuse monitoring the club considers standard.

Jones, Marshall and Denison reach a complementary conclusion from a different direction.² **PEER-REVIEWED** Drawing on Foucault, they argue that wearable GPS in professional rugby league functions as a disciplinary technology, one that players come to

accept as normal and then apply to themselves, with possible consequences for their health and wellbeing that the performance rationale does not capture. Their analysis concerns adult professionals with contracts, agents and unions.

A young athlete has none of those supports, and the transfer of these findings to children is my inference. I think it is a cautious one. If an adult professional with representation finds it hard to decline monitoring, a 13-year-old hoping to be selected will find it harder, and the normalising effect Jones and colleagues describe may take deeper root in someone whose sense of self is still forming. The evidence to test that directly does not yet exist. **EVIDENCE GAP**

ADULTS, NOT CHILDREN

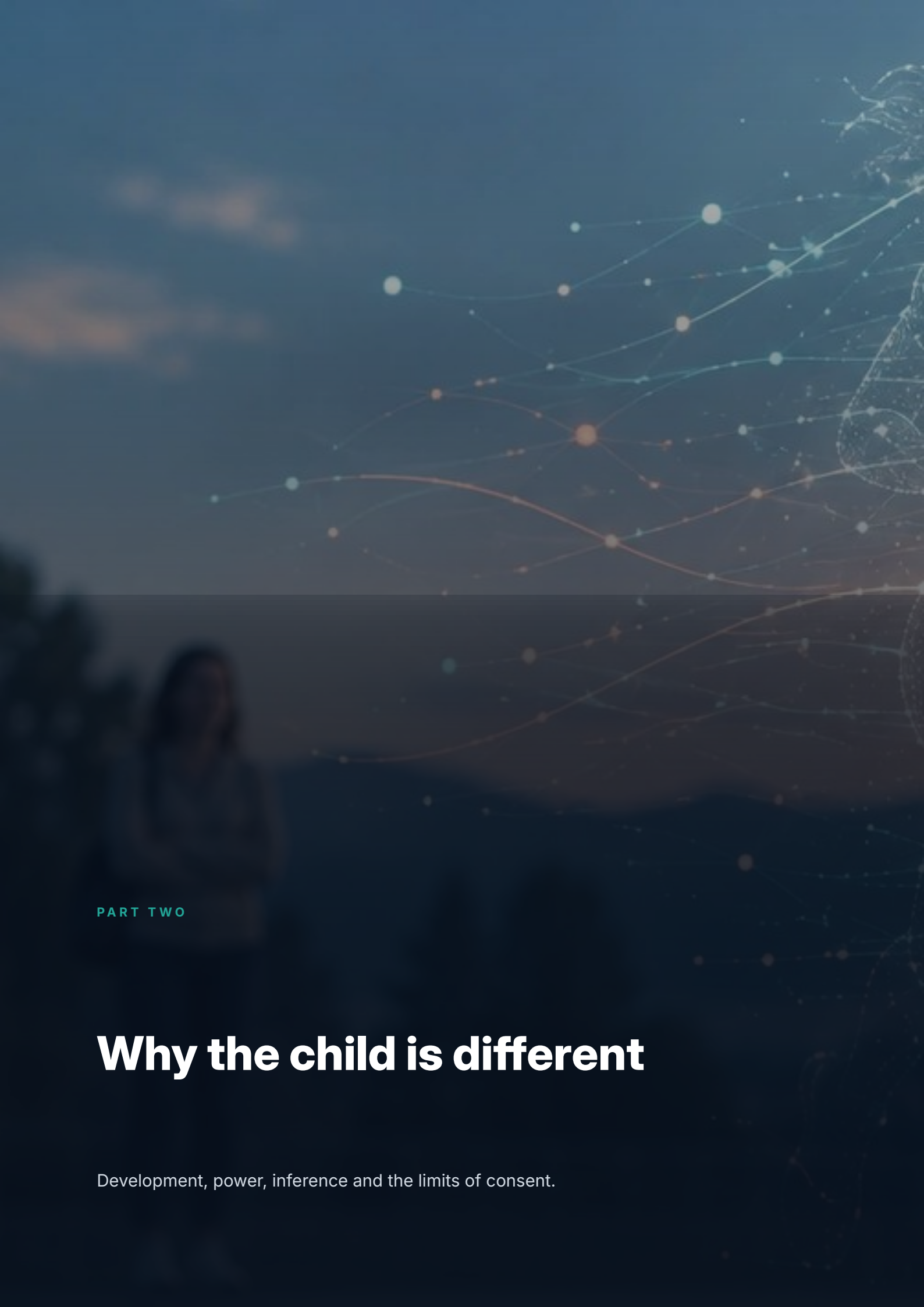
Both studies concern adult professionals. I have found no comparable empirical work on how monitoring is experienced by athletes in youth pathways, which is itself a finding.

2.3 What changes when AI arrives

Until recently, the inferences drawn from athlete data were mostly produced by people with sport science training who could be asked how they reached them. General-purpose AI systems change that in two ways. They lower the skill needed to produce an apparently authoritative analysis, and they obscure the reasoning behind it. Sport science journals have begun to examine the risks of generative AI within their own fields, with one analysis setting out both the prospects and the threats that ChatGPT poses to academic writing.¹⁴ **PEER-REVIEWED** The same fluency that worries journal editors could, in my view, produce confident reports on individual children from data the tool does not understand.

Regulators have started to respond to AI systems that reach children. On 2 February 2023 Italy's data protection authority, the Garante, ordered a temporary limitation on the processing of Italian users' data by the Replika chatbot, citing risks to minors and to emotionally vulnerable people.⁸ **CONSENSUS** On 30 March 2023 the same authority ordered a temporary limitation on ChatGPT, pointing among other concerns to the absence of any mechanism to verify users' ages.⁹ **CONSENSUS** In the United Kingdom, the government's AI regulation white paper, published on 29 March 2023, proposes five cross-sector principles and relies on existing regulators, among them the Information Commissioner's Office (ICO), to apply them within their remits.¹⁵ **CONSENSUS**

None of these actions concerns sport. Their relevance lies in the direction they indicate. Authorities are willing to act quickly when an AI service handles children's data without adequate safeguards, and the white paper would have the ICO, which already has a detailed code on children, apply the new principles within its remit. My own view, which goes beyond what either the Garante or the white paper says, is that pathways adopting AI analytics should expect scrutiny of their youth data practices to increase, and should prepare before it arrives.



PART TWO

Why the child is different

Development, power, inference and the limits of consent.

CHAPTER THREE

Why children are different

The case for treating youth athlete data differently rests on facts about children that no adult framework was built to handle.

Children change quickly and have little power in selective settings, and both facts are well documented in their own fields. Taken together, they explain why the ethics of adult monitoring cannot simply be scaled down to fit a 13-year-old.

3.1 A body still changing

Malina, Bouchard and Bar-Or's standard text on growth and maturation sets out how children of the same chronological age can differ substantially in biological maturity, and how the timing and tempo of the adolescent growth spurt vary from one individual to the next.³ **SCHOLARLY** A measurement taken at 13 therefore describes a particular child at a particular point in a process whose pace is unknown. An early maturer may look stronger and faster than peers for reasons that have little to do with lasting ability, and a late maturer may look the reverse.

A later review by Malina and colleagues examines the methods sport uses to estimate maturity in youth athletes, from skeletal age to predictions based on height and body dimensions.⁴ **REVIEW** The authors caution that the more convenient prediction methods carry errors, and that those errors appear to be greatest for the early and late maturers whose classification matters most. For the purposes of this paper the implication is simple enough. Any AI tool that ingests these estimates inherits their uncertainty, and a score built on them may present a guess as a finding.

The International Olympic Committee's consensus statement on youth athletic development reinforces this from a policy angle.¹⁶ **CONSENSUS** It urges sport organisations to put the health and long-term development of young athletes ahead of early results, and to recognise wide individual variation in growth and maturation when making decisions about them.

3.2 An unequal room

The second fact is power. A pathway exists to select, and a child who wants to be selected has every reason to comply with whatever the pathway asks. Parents are in a similar position. A parent who hesitates over a consent form may reasonably wonder whether hesitation will be noticed, and in a small programme it may well be. **PRACTICE** I have watched this dynamic play out often enough to be confident it is common, though I know of no study that measures it in sport. **EVIDENCE GAP**

The imbalance runs through time as well as through the room. The adults who design a monitoring programme will have moved on long before its records stop being used, while the child will carry the consequences for years. A decision made by a coach in 2023 about what to store is, in effect, a decision about what a 25-year-old will find in 2035.

3.3 What the policy frameworks say

Several frameworks already treat children's data as a special case, though none was written with sport pathways in mind. Recital 38 of the GDPR states that children merit specific protection of their personal data because they may be less aware of the risks, consequences and safeguards involved.⁶ **CONSENSUS** The ICO's Age Appropriate Design Code sets out 15 standards for online services likely to be accessed by children, with conformance expected from 2 September 2021.¹¹ **CONSENSUS** Among them, the code expects services to put the best interests of the child first, to collect and retain only the minimum personal data needed, and to switch profiling off by default unless there is a compelling reason otherwise. Another standard asks services not to use children's personal data in ways shown to be detrimental to their wellbeing.¹¹ **CONSENSUS** A sport pathway is not the kind of service the code was drafted around. Even so, the logic of that standard carries across, in my view, to any inference that could narrow the opportunities offered to a child on the strength of an uncertain prediction.

UNICEF's policy guidance on AI for children sets out nine requirements for child-centred AI, including protecting children's data and privacy and supporting their development and wellbeing.¹² **CONSENSUS** The UNESCO Recommendation on the Ethics of Artificial Intelligence, adopted in November 2021, likewise includes privacy and data protection among the principles it says should be respected throughout an AI system's life cycle.¹⁷ **CONSENSUS**

TO CONFIRM

Whether a pathway's athlete management platform, wellness app or analytics tool falls within the scope of the Age Appropriate Design Code depends on how the service is provided and to whom. Each organisation's data protection adviser should confirm scope for each system it uses.

Framework	Status	What it offers a pathway
GDPR, Recital 38	EU regulation, with a UK version	States that children merit specific protection
ICO Age Appropriate Design Code	UK statutory code of practice	Best interests, data minimisation, profiling off by default
UNICEF Policy Guidance on AI for Children 2.0	International guidance	Nine requirements for child-centred AI
UNESCO Recommendation on the Ethics of AI	Adopted by member states, non-binding	Privacy across the AI life cycle
FIFPRO Charter of Player Data Rights	Players' union charter	Rights for adult professional footballers only

TABLE 1 Selected frameworks bearing on youth athlete data, as at April 2023. Sources: GDPR (2016); ICO (2020); UNICEF (2021); UNESCO (2021); FIFPRO (2022).

What the table shows is a gap. General children's frameworks do not speak to the specific practices of sport, and the one sport-specific charter, which FIFPRO launched on 19 September 2022 after developing it with FIFA, sets out eight rights for professional footballers and says nothing about academy or youth players.¹⁰ **INDUSTRY** Whether that silence reflects a judgement that youth players are covered elsewhere, or simply that nobody has yet addressed them, I cannot tell from the published material.

CHAPTER FOUR

Inferred data as the real risk

Measurements describe a child. Inferences make claims about who that child will become, and those claims travel further.

Most discussion of athlete data concentrates on what is collected. This chapter argues that the larger risk lies in what is inferred, because inferred data are harder for the athlete to see, harder to correct and more likely to drive decisions.

4.1 From measurement to prediction

A pathway's records can be sorted into three kinds. Measured data are direct observations such as a sprint time or a height. Derived data are calculations from those observations, such as a training load ratio. Inferred data are predictions or judgements about the athlete, such as an estimated maturity status, an injury risk flag or a potential rating, and they are increasingly produced by software.

The GDPR defines profiling as automated processing that evaluates personal aspects of a person, in particular to analyse or predict matters such as their performance at work, health, reliability or behaviour.⁶ **CONSENSUS** An AI tool that turns a child's records into a potential rating fits that description closely on my reading, though whether a given tool amounts to profiling in the legal sense is a question for each organisation's advisers.

Wachter and Mittelstadt argue, in a 2019 legal analysis, that European data protection law offers people relatively little control over inferences drawn about them, compared with the data those inferences come from.⁵ **SCHOLARLY** They propose a right to reasonable inferences, under which organisations would have to justify in advance why a given inference is acceptable, relevant and reliable. It has not been adopted. It does, however, name the problem this chapter describes with some precision, and it offers a test (is this inference reasonable, and can the organisation show why?) that a pathway could apply to its own tools without waiting for any change in the law.

TO CONFIRM

Whether a pathway's use of an analytics tool counts as profiling, and what follows if it does, needs checking against the organisation's own processing. The answer may differ between tools.

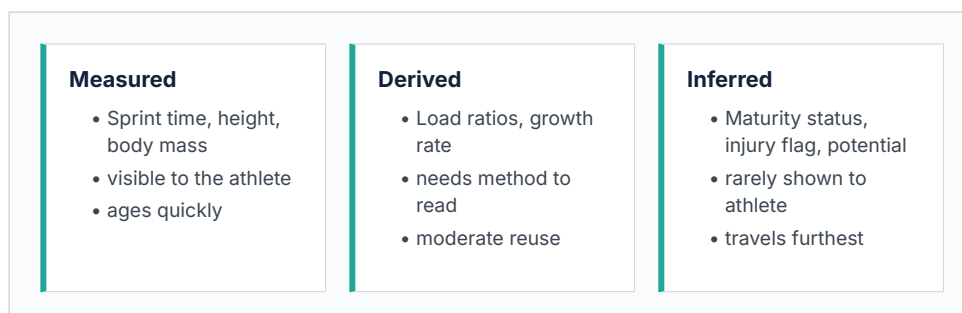


FIGURE 1 Three kinds of youth athlete data carry different risks. Source: author's analysis.

4.2 Where inferences travel

The three kinds of data differ in how far they spread. A sprint time recorded at 13 has little use to anyone five years later. An inferred label such as "late maturer, high potential" or "injury prone" keeps its apparent meaning, and it can move into selection meetings, into files shared with partner clubs, into contract discussions and, conceivably, into commercial scouting products. **PRACTICE** Karkazis and Fishman's concern about physiological data entering contract decisions in adult sport suggests that this path is not hypothetical.¹ **PEER-REVIEWED**

A further route runs through the vendors themselves. Some analytics providers pool records from many client organisations to build or refine their models, which would make one child's history part of a product used to assess other children. **PRACTICE** Whether a particular vendor does so, and on what terms, can only be determined from its contract and documentation, and I have seen few pathway agreements that ask the question directly. A board that cannot answer it for each tool in use has, in my view, found its first task.

The consequences of an inference also outlast its accuracy. Given the uncertainty in maturity estimates described in the previous chapter, a potential rating built at 13 may be wrong in ways that only become visible at 17.⁴ **REVIEW** By then the rating may already have shaped the coaching the athlete received, the opportunities offered and the way adults spoke about them, which can make an inaccurate prediction partly self-fulfilling (a mechanism familiar to anyone who has worked on relative age effects, though I know of no study that tests it for AI-generated ratings). **EVIDENCE GAP**

4.3 Why the inferred category needs its own rules

Controls should follow the risk. My own view is that inferred data about under-18s should be held separately from measured data, available to fewer people, kept for shorter periods and never shared outside the organisation without a fresh and specific decision. That view goes beyond what any of the sources cited here says, although it is consistent with the ICO code's expectation that profiling be off by default for children.¹¹ **CONSENSUS**

There is a practical difficulty. Many commercial platforms do not separate the three kinds of data, and some present inferred scores alongside measurements as if they were of the same standing. **PRACTICE** Separation may therefore require pathways to ask vendors for changes, or to decide that some features should stay switched off for young athletes.

CHAPTER FIVE

Consent that cannot be informed

A signature on a screening day is asked to authorise uses that nobody present can describe.

In my experience, consent is the mechanism pathways most often point to when asked how they justify collecting data on children. **PRACTICE** This chapter argues that it cannot bear that weight, for two reasons, and that anonymisation, the usual fallback, offers less protection in elite youth sport than organisations may assume.

5.1 Agreeing to the unknown

Informed consent presupposes that the person agreeing knows what they are agreeing to. When data are collected in 2023 and may be analysed in 2028 by a tool that does not yet exist, for a purpose that has not yet been defined, that presupposition fails. A longer form cannot fix that failure. A form that tried to list every possible future use would be either so broad as to mean nothing or so long that no parent would read it on a clipboard by the door.

The GDPR itself recognises that children may be less aware of the risks and consequences of processing.⁶ **CONSENSUS** Where the consent is given by a parent, a further question arises about whose interests the signature represents. A parent may value selection for their child more highly than the child's future privacy, and neither view is unreasonable, but the child is the person who will live with the record. The observation is mine, and I offer it without any suggestion that parents act improperly.

A more workable approach treats consent as a process that continues through the pathway. The pathway would return to the family, and increasingly to the athlete, whenever a new tool or a new use is introduced, and the athlete's own view would carry more weight each year as their understanding grows. **ATHLEET.AI** This is harder to administer than a single form, and it will not remove the problem of unknown future uses, but it shortens the distance between the decision and the use it authorises.

5.2 Refusal with a price

The GDPR defines consent as freely given, among other conditions.⁶ **CONSENSUS** In a selective pathway that freedom is often limited in practice. If a child who declines monitoring cannot be compared with peers, a coach may find it harder to justify selecting them, and both the child and the parent will understand that. **PRACTICE** Karkazis and Fishman make a comparable point about adult professionals, whose employment gives them limited room to refuse.¹ **PEER-REVIEWED** In youth sport the pressure is softer and possibly stronger, because it operates through hope.

TO CONFIRM

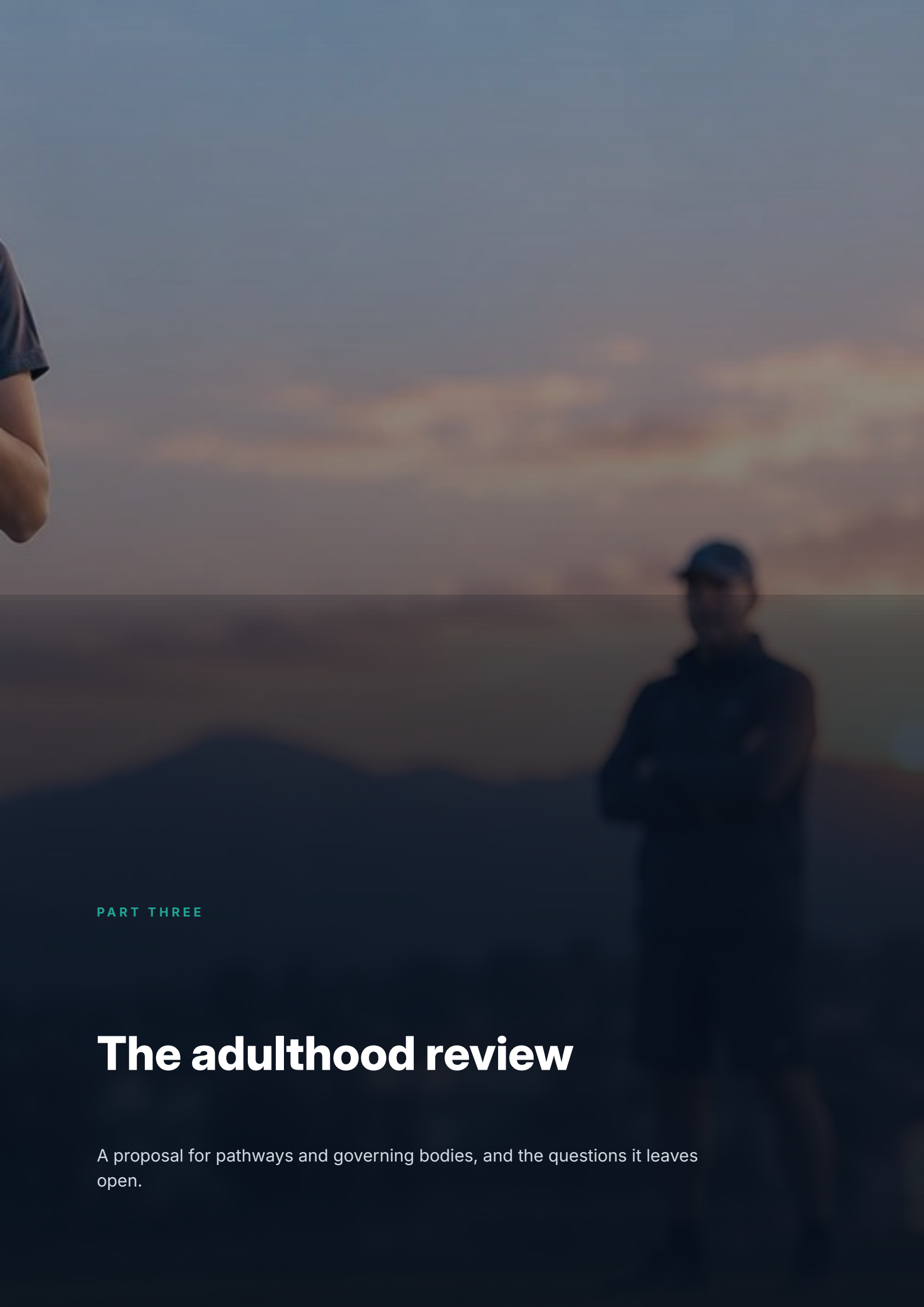
Whether consent, or some other basis, is the appropriate basis for processing youth athlete data at all is a legal question for each organisation. The analysis here concerns the practical quality of consent and leaves its legal status to advisers.

A choice with consequences is still a choice, but it should not be described as free. Pathways that rely on consent should be honest about the pressure their selection role creates, and should design their data practices so that declining non-essential collection carries no cost to selection. Where a measure is essential to coaching or safeguarding, the pathway should say so openly and present it as a requirement.

5.3 Why anonymisation will not rescue the cohort

A common response to privacy concerns is to anonymise data before sharing or analysing them. Rocher, Hendrickx and de Montjoye developed a method for estimating how likely a person is to be re-identified in an incomplete dataset, and applied it across 210 populations.⁷ [PEER-REVIEWED](#) Using their model, they estimated that 99.98% of Americans would be correctly re-identified in any dataset using 15 demographic attributes, and concluded that even heavily sampled anonymised datasets are unlikely to meet the standards for anonymisation set by the GDPR.

Elite youth cohorts are a far easier case than a national population. A regional squad may have 20 members, whose ages, events, clubs and competition results are publicly listed, and whose growth and performance trajectories are known to coaches, parents and rivals. Removing names from such a dataset does little, in my view, because a recognisable profile (the tallest 14-year-old hurdler in the region, say) is identifying in itself. Pathways that plan to share youth data with researchers or vendors on the strength of anonymisation should test that assumption carefully before relying on it.



PART THREE

The adulthood review

A proposal for pathways and governing bodies, and the questions it leaves open.

CHAPTER SIX

A framework for the measured child

The remedy lies in structure. A pathway can decide what it collects, how it treats inferences, when records end and when the athlete gets a say.

The proposal in this chapter is my own, offered as a framework for discussion. It has not been tested in any pathway, and every part of it would need adapting to the organisation's legal obligations and to the sport concerned. **ATHLEET.AI**

6.1 Four commitments

The framework rests on four commitments that a pathway board could adopt as policy. **ATHLEET.AI**

1. Collect only what a named coaching or safeguarding purpose requires, and record that purpose against each data item at the point of collection.
2. Hold inferred data apart from measured and derived data, restrict access to it and never share it outside the organisation without a fresh, specific decision.
3. Set retention periods for each category that end, with inferred data held for the shortest time.
4. Give every athlete measured before 18 a formal review of their record once they reach adulthood.

The first commitment follows the data minimisation and storage limitation principles in Article 5 of the GDPR, which require personal data to be limited to what is necessary for their purposes and kept in identifiable form for no longer than necessary.⁶

CONSENSUS The second reflects the earlier argument about inferred data and the ICO code's expectation on profiling.¹¹ **CONSENSUS** The third and fourth are where the framework goes furthest beyond current practice.

Retention needs a default. A period that must be argued for each time tends to become indefinite, so my suggestion is that inferred data about an athlete lapse when they leave the pathway or reach 18, whichever comes first, unless a documented safeguarding or medical reason requires them to be kept. **ATHLEET.AI** Measured data can be held longer, since they support coaching education and research once properly protected, but the same principle applies and each period should have an end date written into policy. The exact periods are for each governing body to set with its advisers, and they may differ between sports with early and late peak ages.

1 Purpose at collection

Each item is tied to a named coaching or safeguarding purpose.

2 Inferences held apart

Predictions and ratings sit in a separate, restricted store.

3	Expiry by default	Records lapse unless a documented reason keeps them.
4	Notice at 18	The athlete is told what exists and offered a review.
5	Review and withdraw	The athlete sees the record and can ask for deletion.
6	Confirm and close	The pathway confirms what was kept, deleted or corrected.

FIGURE 2 The adulthood review in outline. Source: author's framework.

6.2 The review at 18

Recital 65 of the GDPR notes that the right to erasure is relevant in particular where a person gave consent as a child without being fully aware of the risks, and that they should be able to exercise the right even though they are no longer a child.⁶

CONSENSUS The adulthood review builds on that idea by making the opportunity active. The pathway would contact each athlete at 18, show them what it holds, including every inference, explain who has received it, and ask whether they want any of it corrected or removed.

ILLUSTRATIVE SCENARIO

The review in practice

Sam is measured in a pathway from 13 to 16 and then released. A month after Sam turns 18, the governing body sends a short message with a link to a plain-language summary of the record. It lists the measurements, two maturity estimates and a potential rating produced by an analytics tool, and it names the partner academy that received the rating. Sam asks for the rating and the wellness entries to be deleted and agrees that the pathway may keep the sprint times for a further two years for use in coach education. The governing body confirms the deletions, and asks the partner academy to do the same.

Several practical problems remain. Many athletes will have left the sport by 18, and contact details will be out of date. Some will ignore the message. Records may have been shared with organisations that the pathway cannot compel. My own view is that none of these problems defeats the proposal, and that a review which reaches most athletes is far better than the present default, which reaches none. A governing body could reduce the contact problem by collecting a contact route at the time of release, and could reduce the sharing problem by writing onward deletion into any data sharing agreement from the outset.

6.3 What a governing body should do first

A governing body that wanted to act on this paper should begin with an inventory, and the inventory should list the systems that hold data on under-18s, the categories each holds, whether any system produces inferences, and who receives them. Once the inventory exists, the body can decide which inferred outputs to switch off for young athletes, what retention period each category should carry, and how the review at 18 will reach people. Advisers belong in the room from day one. Several of the questions flagged in this paper turn on the organisation's own processing and cannot be resolved by a general framework.

Contracts come next. Commercial AI tools will otherwise set the defaults, and a vendor's default is designed around its product and its own interest in data, which may include using a client's records to improve the tool for everyone else. National bodies are better placed than individual clubs to negotiate those terms, because they can negotiate once for a whole sport.

6.4 Conclusion and open questions

No child who steps onto a testing mat at 13 can agree to what may be done with the result over the following decade, however carefully the form is worded and however well the adults around them mean. AI tools make that decade more eventful, because they can turn an ordinary record into a lasting prediction at almost no cost. The frameworks that exist for adult sport and for children's digital services each cover part of the problem, and neither was written for the measured child. The adulthood review is one way to close the gap. It will change on contact with practice.

Some questions remain that I cannot answer from the evidence available. Should governing bodies set a minimum age below which no predictive inference is produced at all? Who should hold youth pathway data, whether the club, the governing body or a trusted third party with no interest in selection? How would a review work for athletes who left the sport at 14? And what would young athletes themselves say, if anyone asked them, about which parts of their record they would want to exist when they are 30?

END MATTER

References

1. Karkazis, K. and Fishman, J.R. (2017) Tracking U.S. professional athletes: the ethics of biometric technologies. *American Journal of Bioethics* 17(1), 45-60.
2. Jones, L., Marshall, P. and Denison, J. (2016) Health and well-being implications surrounding the use of wearable GPS devices in professional rugby league: a Foucauldian disciplinary analysis of the normalised use of a common surveillance aid. *Performance Enhancement and Health* 5(2), 38-46.
3. Malina, R.M., Bouchard, C. and Bar-Or, O. (2004) *Growth, Maturation, and Physical Activity*. 2nd edn. Champaign, IL: Human Kinetics.
4. Malina, R.M., Rogol, A.D., Cumming, S.P., Coelho e Silva, M.J. and Figueiredo, A.J. (2015) Biological maturation of youth athletes: assessment and implications. *British Journal of Sports Medicine* 49(13), 852-859.
5. Wachter, S. and Mittelstadt, B. (2019) A right to reasonable inferences: re-thinking data protection law in the age of big data and AI. *Columbia Business Law Review* 2019(2), 494-620.
6. European Parliament and Council of the European Union (2016) Regulation (EU) 2016/679 (General Data Protection Regulation). *Official Journal of the European Union* L 119, 4 May 2016, 1-88.
7. Rocher, L., Hendrickx, J.M. and de Montjoye, Y.-A. (2019) Estimating the success of re-identifications in incomplete datasets using generative models. *Nature Communications* 10, 3069.
8. Garante per la protezione dei dati personali (2023) Provvedimento del 2 febbraio 2023 [order concerning the Replika chatbot].
9. Garante per la protezione dei dati personali (2023) Provvedimento del 30 marzo 2023 [limitation order concerning ChatGPT].
10. FIFPRO (2022) Charter of Player Data Rights launched for professional footballers. 19 September 2022.
11. Information Commissioner's Office (2020) *Age Appropriate Design: A Code of Practice for Online Services*. Conformance expected from 2 September 2021.
12. UNICEF (2021) *Policy Guidance on AI for Children 2.0*. November 2021.
13. Lupton, D. (2016) *The Quantified Self: A Sociology of Self-Tracking*. Cambridge: Polity.
14. Dergaa, I., Chamari, K., Zmijewski, P. and Ben Saad, H. (2023) From human writing to artificial intelligence generated text: examining the prospects and potential threats of ChatGPT in academic writing. *Biology of Sport* 40(2), 615-622.
15. Department for Science, Innovation and Technology (2023) *A pro-innovation approach to AI regulation*. CP 815, 29 March 2023.
16. Bergeron, M.F., Mountjoy, M., Armstrong, N., Chia, M., Côté, J., Emery, C.A. et al. (2015) International Olympic Committee consensus statement on youth athletic development. *British Journal of Sports Medicine* 49(13), 843-851.
17. UNESCO (2021) *Recommendation on the Ethics of Artificial Intelligence*. Adopted 23 November 2021.

END MATTER

About the author

Paul Forrest is the founder of Athleet.AI. He advises boards and leadership teams on the adoption and governance of artificial intelligence, drawing on a background in business process reengineering and on hands-on work building machine learning systems, including small, domain-specific language models. He has served as a special adviser to UK government on the Global Entrepreneur Programme and works as a board member and business adviser.

He is also a qualified athletics coach who coaches sprint, hurdle and endurance athletes as a volunteer, and much of his thinking about how people learn, decide and improve comes from the track.

This paper is part of the Athleet.AI Working Papers, AI in Coaching and Athlete Development. The views expressed are the author's own.

athleet.ai

The Measured Child

Athleet.AI Working Papers, AI in Coaching and Athlete Development
Paper three of six · 19 April 2023