

Playing by the Rules

Of AI, data protection law, and sports

Pierre Dewitte

✉ pierre.dewitte@kuleuven.be

🐦 [@PiDewitte](https://twitter.com/PiDewitte)

Focus on injury forecasting



**Two questions
and a key
takeaway**



**Two friction
points and a
research
agenda**

An injury forecasting system for football players



An injury forecasting system for football players

- **Injury forecasting system** using training data and machine learning
 - Use case available [here](#) (short version) and [here](#) (academic paper)

Effective injury forecasting in soccer with GPS training data and machine learning

Alessio Rossi^{1*}, Luca Pappalardo^{1,2}, Paolo Cintia², F. Marcello Iaia³, Javier Fernández⁴, Daniel Medina⁵

Data scientists are predicting sports injuries with an algorithm

Machine learning can tell athletes when to train and when to stop.

An injury forecasting system for football players

- **Injury forecasting system** using training data and machine learning
 - Use case available [here](#) (short version) and [here](#) (academic paper)
 - Jersey equipped with GPS sensors measuring training data



An injury forecasting system for football players

- **Injury forecasting system** using training data and machine learning
 - Use case available [here](#) (short version) and [here](#) (academic paper)
 - Jersey equipped with GPS sensors measuring training data
 - Goal = train a model to predict injuries based on a series of variables

d_{TOT}	Distance in meters covered during the training session
d_{HSR}	Distance in meters covered above 5.5m/s
d_{MET}	Distance in meters covered at metabolic power
d_{IML}	Distance in meters covered by a player with a Metabolic Power is above 25.5W/Kg
$d_{IML/m}$	Distance in meters covered by a player with a Metabolic Power is above 25.5W/Kg per minute
d_{EXP}	Distance in meters covered above 25.5W/Kg and below 19.8Km/h
Acc_2	Number of accelerations above 2m/s ²
Acc_3	Number of accelerations above 3m/s ²
Dec_2	Number of decelerations above 2m/s ²
Dec_3	Number of decelerations above 3m/s ²
DSL	Total of the weighted impacts of magnitude above 2g. Impacts are collisions and step impacts during running
FI	Ratio between DSL and speed intensity
Age	age of players
BMI	Body Mass Index: ratio between weight (in kg) and the square of height (in meters)
Role	Role of the player
PI	Number of injuries of the players before each training session
Play time	Minutes of play in previous games
Games	Number of games played before each training session

Q#1 Material scope of application

Q#1 Does the GDPR **apply** to the injury forecasting system?

Q#1 Material scope of application

- **Material scope** (Art. 2(1) GDPR)

1. This Regulation applies to the processing of personal data wholly or partly by automated means and to the processing other than by automated means of personal data which form part of a filing system or are intended to form part of a filing system.

Q#1 Material scope of application

- **Material scope** (Art. 2(1) GDPR)
- A “**processing**” (Art. 4(2) GDPR)

(2) ‘processing’ means any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction;

As a practical example of application of these rules, let us consider a player’s training session with $PI^{(EWMA)} = 0.28$, $d_{HSR}^{(EWMA)} = 126.58$ and $d_{TOT}^{(MSWR)} = 1.66$, associated with an injury. This example is associated with rule 2 (Fig 4A), corresponding to the following decision path:

$$d_{HSR}^{(EWMA)} > 112.35 \rightarrow d_{TOT}^{(MSWR)} \leq 1.78 \rightarrow PI^{(EWMA)} > 0.03 \rightarrow PI^{(EWMA)} \leq 0.68 \\ \rightarrow \text{INJURY}$$

Q#1 Material scope of application

- **Material scope** (Art. 2(1) GDPR)
- A “**processing**” (Art. 4(2) GDPR)
- Of “**personal data**” (Art. 4(1) GDPR)

(1) ‘personal data’ means any information relating to an identified or identifiable natural person (‘data subject’); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person;

Q#1 Material scope of application

- **Material scope** (Art. 2(1) GDPR)
- A “**processing**” (Art. 4(2) GDPR)
- Of “**personal data**” (Art. 4(1) GDPR)

d_{TOT}	Distance in meters covered during the training session
d_{HSR}	Distance in meters covered above 5.5m/s
d_{MET}	Distance in meters covered at metabolic power
d_{HML}	Distance in meters covered by a player with a Metabolic Power is above 25.5W/Kg
$d_{HML/m}$	Distance in meters covered by a player with a Metabolic Power is above 25.5W/Kg per minute
d_{EXP}	Distance in meters covered above 25.5W/Kg and below 19.8Km/h
Acc_2	Number of accelerations above 2m/s ²
Acc_3	Number of accelerations above 3m/s ²
Dec_2	Number of decelerations above 2m/s ²
Dec_3	Number of decelerations above 3m/s ²
DSL	Total of the weighted impacts of magnitude above 2g. Impacts are collisions and step impacts during running
FI	Ratio between DSL and speed intensity
Age	age of players
BMI	Body Mass Index: ratio between weight (in kg) and the square of height (in meters)
Role	Role of the player
PI	Number of injuries of the players before each training session
Play time	Minutes of play in previous games
Games	Number of games played before each training session

Q#1 Material scope of application

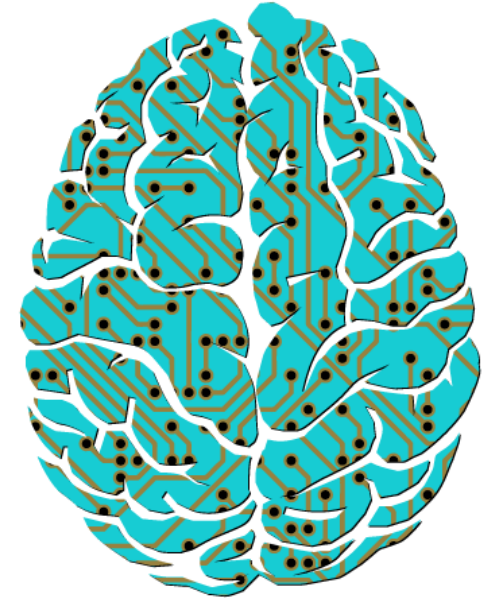
Layer	Controller	Processing	Personal data	Purpose	Means
#1a	[See Q#2]	Collection	Historical performance data	Creating training dataset A	By scraping statistics databases
#1b	[See Q#2]	Collection	Real-time performance data from players XYZ	Creating training dataset B	By equipping players XYZ with jerseys
#1c	[See Q#2]	<i>fp</i> : Use	Datasets A + B	Train injury forecasting system	By using a linear regression
#1d	[See Q#2]	<i>fp</i> : Use	Trained model (<i>debated</i>)	Fine-tune injury forecasting system	By using RLwHF
#2a	[See Q#2]	Collection	Real-time performance data from players UVW	Forecast injury in real-time	By equipping players UVW with jerseys + using the trained model
#2b	[See Q#2]	<i>fp</i> : Sharing	Historical performance data	Monetising the trove of data	By disclosing to third party

Q#2 Personal scope of application

Q#1 Does the GDPR **apply** to the injury forecasting system?

Q#2 Who is **responsible** for complying with the GDPR?

Q#2 Personal scope of application



Who is responsible for complying with the GDPR?

Q#2 Personal scope of application

Controller
(Art. 4(7) GDPR)

*Must comply with all principles
and rules contained in the
GDPR*

And

Processor
(Art. 4(8) GDPR)

*Must only comply with specific
obligations (e.g. security,
maintain a record of
processing, etc.)*

Q#2 Personal scope of application

- The notion of **controller** (Art. 4(7) GDPR)

(7) 'controller' means the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data; where the purposes and means of such processing are determined by Union or Member State law, the controller or the specific criteria for its nomination may be provided for by Union or Member State law;

Q#2 Personal scope of application

- The notion of **joint controller** (Art. 26(1) GDPR)

1. Where two or more controllers jointly determine the purposes and means of processing, they shall be joint controllers. They shall in a transparent manner determine their respective responsibilities for compliance with the obligations under this Regulation, in particular as regards the exercising of the rights of the data subject and their respective duties to provide the information referred to in Articles 13 and 14, by means of an arrangement between them unless, and in so far as, the respective responsibilities of the controllers are determined by Union or Member State law to which the controllers are subject. The arrangement may designate a contact point for data subjects.

Q#2 Personal scope of application

- The notion of **processor** (Art. 4(7) GDPR)

(8) 'processor' means a natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller;

Q#2 Personal scope of application

Layer	Controller	Processing	Personal data	Purpose	Means
#1a	Tech provider α	Collection	Historical performance data	Creating training dataset A	By scraping statistics databases
#1b	Tech provider β	Collection	Real-time performance data from players XYZ	Creating training dataset B	By equipping players XYZ with jerseys
#1c	Tech provider γ + football club δ (JC)	<i>fp</i> : Use	Datasets A + B	Train injury forecasting system	By using a linear regression
#1d	Tech provider γ + football club δ (JC)	<i>fp</i> : Use	Trained model (<i>debated</i>)	Fine-tune injury forecasting system	By using RLwHF
#2a	Football club ϵ	Collection	Real-time performance data from players 123	Forecast injury in real-time	By equipping players 123 with jerseys + using the trained model
#2b	Football club ϵ	<i>fp</i> : Sharing	Historical performance data	Monetising the trove of data	By disclosing to third party

Key takeaway

Adopt a **phase-oriented approach** when confronted to *complex* processing activities involving *multiple* actors

Algorithmic supply chains are *prime examples* of such complex and multi-stakeholder systems

FP#1 The peculiar status of athletes

FP#1 (Explicit) **consent** v. performance of a **contract**

FP#1 The peculiar status of athletes

- The processing of personal data must be based on a **lawful ground**
- **Six lawful grounds** to choose from (Art. 6(1) GDPR)
 1. Processing shall be lawful only if and to the extent that at least one of the following applies:
 - (a) the data subject has given consent to the processing of his or her personal data for one or more specific purposes;
 - (b) processing is necessary for the performance of a contract to which the data subject is party or in order to take steps at the request of the data subject prior to entering into a contract;
 - (c) processing is necessary for compliance with a legal obligation to which the controller is subject;
 - (d) processing is necessary in order to protect the vital interests of the data subject or of another natural person;
 - (e) processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller;
 - (f) processing is necessary for the purposes of the legitimate interests pursued by the controller or by a third party, except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject which require protection of personal data, in particular where the data subject is a child.

FP#1 The peculiar status of athletes

- **Special categories of personal data (Art. 9(1) GDPR)**

1. Processing of personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation shall be prohibited.

FP#1 The peculiar status of athletes

- **Special categories of personal data** (Art. 9(1) GDPR)
- **Prohibition**, unless relying on one of the ten **exemptions** (Art. 9(2) GDPR)
 - Lawfulness and the need to specify an exemption = cumulative requirements

- (a) the data subject has given explicit consent to the processing of those personal data for one or more specified purposes, except where Union or Member State law provide that the prohibition referred to in paragraph 1 may not be lifted by the data subject;
- (b) processing is necessary for the purposes of carrying out the obligations and exercising specific rights of the controller or of the data subject in the field of employment and social security and social protection law in so far as it is authorised by Union or Member State law or a collective agreement pursuant to Member State law providing for appropriate safeguards for the fundamental rights and the interests of the data subject;
- (c) processing is necessary to protect the vital interests of the data subject or of another natural person where the data subject is physically or legally incapable of giving consent;
- (d) processing is carried out in the course of its legitimate activities with appropriate safeguards by a foundation, association or any other not-for-profit body with a political, philosophical, religious or trade union aim and on condition that the processing relates solely to the members or to former members of the body or to persons who have regular contact with it in connection with its purposes and that the personal data are not disclosed outside that body without the consent of the data subjects;
- (e) processing relates to personal data which are manifestly made public by the data subject;
- (f) processing is necessary for the establishment, exercise or defence of legal claims or whenever courts are acting in their judicial capacity;
- (g) processing is necessary for reasons of substantial public interest, on the basis of Union or Member State law which shall be proportionate to the aim pursued, respect the essence of the right to data protection and provide for suitable and specific measures to safeguard the fundamental rights and the interests of the data subject;
- (h) processing is necessary for the purposes of preventive or occupational medicine, for the assessment of the working capacity of the employee, medical diagnosis, the provision of health or social care or treatment or the management of health or social care systems and services on the basis of Union or Member State law or pursuant to contract with a health professional and subject to the conditions and safeguards referred to in paragraph 3;
- (i) processing is necessary for reasons of public interest in the area of public health, such as protecting against serious cross-border threats to health or ensuring high standards of quality and safety of health care and of medicinal products or medical devices, on the basis of Union or Member State law which provides for suitable and specific measures to safeguard the rights and freedoms of the data subject, in particular professional secrecy;
- (j) processing is necessary for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89(1) based on Union or Member State law which shall be proportionate to the aim pursued, respect the essence of the right to data protection and provide for suitable and specific measures to safeguard the fundamental rights and the interests of the data subject.

FP#1 The peculiar status of athletes

- General prohibition of **automated decision-making** (Art. 22 Recital 71)

Prohibition

1. The data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her.

Exemptions

2. Paragraph 1 shall not apply if the decision:

(a) is necessary for entering into, or performance of, a contract between the data subject and a data controller;

(b) is authorised by Union or Member State law to which the controller is subject and which also lays down suitable measures to safeguard the data subject's rights and freedoms and legitimate interests; or

(c) is based on the data subject's explicit consent.

Limitation when
SCPD are
processed

4. Decisions referred to in paragraph 2 shall not be based on special categories of personal data referred to in Article 9(1), unless point (a) or (g) of Article 9(2) applies and suitable measures to safeguard the data subject's rights and freedoms and legitimate interests are in place.

FP#1 The peculiar status of athletes

- “**Freely given**” consent (Art. 4(11) GDPR) is difficult to achieve
 - In employment context, power imbalance employee-employer
 - When the employee is an athlete, room for manoeuvre even more limited
- The “performance of a contract” option require a **necessity test**
 - The processing must be objectively necessary for the contract
 - Otherwise, *not even possible* to rely on that lawful ground
 - The processing is not objectively necessary if
 - It is possible to achieve the stated purpose *without* the said processing
 - There is a *less intrusive* way to achieve the stated purpose
 - Is the processing of performance data objectively necessary in a high-stakes environment such as competition?

FP#2 The peculiar nature of performance data

FP#2 **Facts v. predictions**, and the right to **rectification**

FP#2 The peculiar nature of performance data

- Personal data shall be **accurate** and kept **up-to-date** (Art. 5(1)d GDPR)
 - To be assessed in light of the stated purpose (Nowak)
- Corresponding data subject's right to request **rectification** (Art. 16 GDPR)
- What about **subjective** pieces of personal data used as input data?
- What about **wrong predictions** based on accurate personal data?
 - Accuracy as a legal principle v. statistical accuracy in AI
 - Problem when AI predictions become personal data
 - Clearly *label* predictions so as to not confuse them with facts
 - Allow data subjects to exercise their *rights* (e.g. rectification, explanation)

A research agenda

Lawfulness of **reuse of athlete's public historical performance data**

Impact of high-stakes contractual relationships on the **validity of consent** given by an athlete to its employer

Necessity of historical and real-time performance data for the **performance of the contract** between an athlete and its employer

Scope of the right to **rectification of performance predictions**

Role of **codes of conducts** in providing a harmonised framework for the processing of personal data by federations, committees and clubs

Thanks for your attention!
Any question?

www.law.kuleuven.be/citip

