

LA DATA DANS LE SPORT PROFESSIONNEL : APPORTS ET LIMITES

PTUT

2024 – 2025

PRÉSENTÉ PAR :

**Dorian Verney, Louis De-cazes,
Paul Dupont--Fauvel et
Pierre Lamouroux**

Introduction

Aujourd'hui, il n'est pas nécessaire d'être un athlète de haut niveau pour disposer d'informations précises sur ses performances sportives. Grâce à une montre connectée et parfois à une application sur son smartphone on peut obtenir facilement la fréquence cardiaque, la vitesse moyenne, la distance que nous avons parcourue et le temps de récupération. Autant d'informations qui étaient, il n'y a encore pas si longtemps, réservées à certains domaines notamment dans la préparation physique professionnelle mais aussi dans le secteur médical.

De nos jours, toute cette technologie est en accès direct pour le grand public ; preuve que la culture de la donnée a envahi le sport à tous les échelons et pour tous les pratiquants. Dans ce cadre général, la donnée et donc la « data » dans le sport désigne l'ensemble des informations émanant des capteurs biométriques, du GPS de suivi d'effort, de l'analyse vidéo, des algorithmes permettant de tirer des indicateurs précis concernant les performances sportives.

Dans le cadre du sport professionnel, cette révolution est encore plus marquée. Le football, longtemps pensé comme un sport d'intuitions, de passions, de gestes, s'est lui aussi laissé transformer par cet enjeu technologique. Nous avons choisi de nous concentrer uniquement sur le football professionnel, car il est le sport plus suivi à l'échelle mondiale et en parallèle la data est omniprésente dans les grands championnats. De plus le football actuel représente le marché économique le plus important dans le monde du sport avec des clubs, des championnats et des investissements présents sur tous les continents. Si les statistiques comptaient, jusqu'alors le nombre de buts marqués, de cartons distribués, de passes décisives réalisées, l'arrivée des capteurs corporels, des logiciels d'analyse vidéo et des bases de données comportementales ouvertes en font un champ d'analyses à part entière : Optimisation des performances, affinement du recrutement par le recours à des indicateurs factuels, préparation des matchs par une connaissance approfondie des adversaires et même aide à l'arbitrage, sont les nouveaux domaines exploités par la Data.

Mais de cette transformation, des interrogations naissent cependant. La donnée est-elle réellement apte à rendre compte de toute la richesse et de la complexité d'un sport collectif comme le football ? Les clubs ont-ils tous les mêmes moyens d'analyse ? La surinterprétation des chiffres ne risque-t-elle pas d'éradiquer le ressenti, l'intuition, voire l'humain du jeu ? Et quelle est finalement la place de l'entraîneur, de l'intuition, de la cohésion d'équipe face à la rationalité froide des algorithmes ?

Dans quelle mesure l'usage des données dans le football professionnel transforme-t-il positivement le sport, et quelles en sont les limites ? Pour répondre à cette problématique, nous étudierons la question générale de la data dans le football (I), avant d'analyser les apports concrets pour les clubs et les joueurs (II), puis de discuter des risques, des inégalités et des dérives éthiques ou humaines qu'elle peut engendrer (III).

Table des matières

INTRODUCTION	2
I La data dans le football professionnel : de quoi parle-t-on ?	4
I.1 Qu'est-ce que la data dans le football ?	4
I.2 L'évolution historique de la data dans le football	5
I.3 L'écosystème actuel de la data football : acteurs, économie et réglementation	6
I.3.1 Les acteurs économiques de la data football	6
I.3.2 Les usages commerciaux et dérivés	7
I.3.3 Réglementation et encadrement	7
II Les apports de la data dans le football professionnel	8
II.1 Optimisation de la performance sportive des joueurs	8
II.2 Analyse des adversaires	10
II.3 Analyse tactique et aide à la décision pendant les matchs	10
II.4 Suivi de la santé des joueurs et des entraînements	11
II.5 Préparation physique et prise en compte de l'environnement d'une compétition	12
II.6 Améliorations du recrutements et valorisation des joueurs	12
II.7 Autres utilisations de la data dans le football : aide à l'arbitrage	13
III Les limites et les risques associés à la data	14
III.1 Déshumanisation et pression sur les acteurs du football	14
III.1.1 Les joueurs réduits à des chiffres	14
III.1.2 Jugements « simplistes » par les médias et supporters	14
III.1.3 Le risque d'une perte d'instinct	15
III.2 Des données peu fiables ou incomplètes	15
III.3 Des données falsifiables ?	16
III.4 Problèmes éthiques et dépendance technologique	16
III.4.1 La vie privée des joueurs en question	16
III.4.2 Une dépendance croissante à la technologie	17
III.4.3 Le cas de la VAR : la data ne résout pas tout	17
Conclusion	18
Bibliographie	19

I La data dans le football professionnel : de quoi parle-t-on ?

I.1 Qu'est-ce que la data dans le football ?

Dans le football moderne, le terme 'data' désigne un ensemble structuré d'informations collectées à différentes étapes du jeu, allant des performances physiques aux décisions tactiques, en passant par l'analyse comportementale des joueurs. Il est donc important de distinguer 2 types de data :

- **La data brute** : par exemple cette data peut-être récupérée par des capteurs GPS, ces données servent essentiellement pour des analyses primaires du jeu, telles que le nombre de tirs, le nombre de kilomètres parcourus... qu'un joueur serait susceptible de réaliser. Ces data ne nécessitent pas d'analyse approfondies. On peut retrouver ces données sur des applications de sport pour le grand public, comme OneFootball. Cette application nous propose un grand nombre de données, telles que les ballons aériens gagnés, ou le nombre de récupérations de balles.
- **La data avancée** : est comme son nom l'indique une forme de donnée plus évoluée. Cette data doit être travaillée avant de la transmettre aux entraîneurs d'un club. Aujourd'hui, lorsque l'on parle de data dans un club de haut niveau, on évoque beaucoup plus souvent le terme de « data avancée ». Cette forme de data nécessite une expertise, c'est pourquoi de nombreux clubs de football possèdent des data analystes. Cette data avancée donne beaucoup plus d'informations sur le travail d'un joueur, sa fatigue, son influence dans le jeu que simplement son nombre de passes ou de tirs. La data avancée peut également servir à aider un autre type d'acteur du football, à savoir les arbitres.

Après avoir abordé les deux types de data, nous allons évoquer les différences concernant les types de données. Ces données sont divisées en quatre groupes :

- **Les données quantitatives** : ces données sont les plus connues, elles correspondent à des chiffres objectifs comme le nombre de tirs cadrés, le pourcentage de passes réussies, etc.
- **Les données biométriques** : ce sont les données recueillies par les capteurs portés par les joueurs, on peut par exemple avoir leur rythme cardiaque, la fatigue musculaire ou encore la fréquence respiratoire. Les préparateurs physiques s'aident de ces données pour adapter les entraînements pour éviter les blessures.
- **Les données vidéo** : les données vidéo servent à analyser les déplacements et positions des joueurs sur le terrain, cela permet d'avoir le schéma tactique en temps réel. Ces données sont récupérées grâce à des caméras disposées tout autour du terrain ou des logiciels de tracking.
- **Les données psychologiques** : ces données servent à garder un suivi mental concernant chaque joueur de l'équipe, on peut donc déterminer sa capacité à gérer la pression sa motivation ou encore son potentiel de progression. Ces données sont souvent collectées par des questionnaires d'auto-évaluation après chaque entraînement où chaque match.

I.2 L'évolution historique de la data dans le football

L'utilisation de données dans le football remonte à plusieurs décennies, mais elle se limitait autrefois à des statistiques basiques comme le nombre de buts, de passes ou de fautes.

La première fois que la data avancée a été utilisée fut en 2001. Manchester United avait 2 ans auparavant était sacré champion d'Europe, avec comme entraîneur Sir Alex Ferguson. Il a pris une décision surprenante, vendre Jaap Stam, acheté 3 ans avant. Jaap Stam a été vendu à Lazio Rome. Cette décision fut non-comprise par les supporters du club car le joueur était considéré comme l'un des meilleurs défenseurs. La raison, bien que jamais avoué publiquement par Alex Ferguson, venait du fait, qu'après certaines analyses, Sir Alex Ferguson a remarqué que son défenseur taclait moins qu'avant. L'entraîneur ayant peur d'un déclin de son défenseur, l'aurait vendu dans la foulée.

Cette anecdote est un tournant dans le monde de la data, Jaap Stam fut donc le premier joueur de football vendu en raison de l'utilisation de la data.

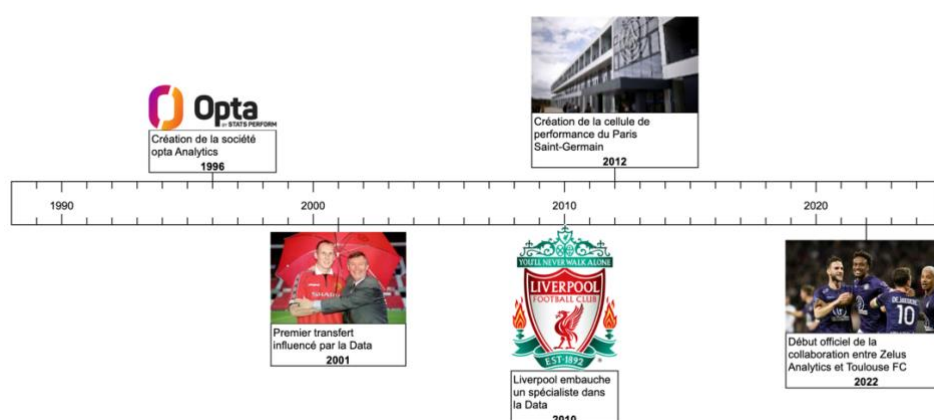


Figure 1 : Synthèse des grandes étapes de l'arrivée de la data dans le football

La data dans le football a donc émergé dans les années 2000. Les premières entreprises spécialisées dans ces services commencent à apparaître : on peut citer l'entreprise **Opta Sports** qui fut créée en 1996. L'entreprise avait pour but de collecter et de distribuer de la donnée sportive brute. C'est dans les années 2010 qu'Opta Analytics fut développé, cette branche d'Opta se concentrant sur l'analyse et l'interprétation de données fournissant des insights modélisés aux clubs professionnels à destination des recruteurs et des analystes.

C'est durant cette même année qu'un professionnel fut embauché par Liverpool pour gérer l'ensemble des datas. En effet le club de première division anglaise, recrute **Ian Graham**, un chercheur en physique théorique, afin d'analyser les performances et le potentiel des joueurs via des méthodes statistiques avancées. Cependant, à ce moment-là, Ian Graham travaille d'abord comme consultant via la société Decision Technology, avant d'être officiellement intégré au club en 2012, où il dirige le département de recherche de Liverpool.

Ce recrutement marque un tournant décisif, pour la première fois, un club mise sur les données pour orienter ses décisions stratégiques, notamment en matière de recrutement. Sous sa direction, Liverpool commence à développer des modèles prédictifs complexes, capables d'identifier des talents sous-évalués sur le marché, et d'optimiser les performances individuelles et collectives.

Après Liverpool, les plus gros clubs commencent à recruter dans le domaine. Le PSG après son rachat par QSI¹ lance officiellement sa cellule performance et possède donc des experts en

¹ Qatar Sport Investment

données. L'objectif était de se doter des meilleurs experts pour pouvoir gagner la Ligue des Champions. Avec le tout nouveaux campus du PSG inauguré en 2024, cette cellule spécialisée est maintenant réservée à l'ensemble des acteurs allant des équipes jeunes du PSG, afin d'améliorer leur performance, aux équipes féminines et l'équipe première masculine. Ce nouvel environnement porte ses fruits, avec les victoires en LDC de l'équipe première ou encore celle des jeunes en finale de la coupe de France des U19.

En France, à partir de 2020, la data s'est démocratisée en devenant l'axe stratégique du recrutement de club plus modeste comme le Toulouse Football Club (TFC).

1.3 L'écosystème actuel de la data football : acteurs, économie et réglementation

Aujourd'hui la data est loin d'être un simple outil au service de la performance sportive. Ce secteur s'organise autour d'acteurs spécialisés, de nouveaux usages commerciaux et d'une réglementation encore en construction. La data est désormais une ressource stratégique, à la fois convoitée et encadrée.

1.3.1 Les acteurs économiques de la data football

Le traitement, l'analyse et la commercialisation des données footballistiques mobilisent de nombreux acteurs spécialisés. Les entreprises, les plus influentes du secteur sont **Opta**, **StatsBomb**, **Wyscout** ou encore **Catapult**, **Hudl** et **Zelus Analytics**. Ces entreprises proposent par exemple des solutions basées sur la biométrie ou la vidéo. Les sociétés spécialisées dans la data football offrent des **services de** collecte, d'analyse, de visualisation et d'exploitation des données **liées aux performances sportives**.

Ces sociétés vendent leurs services à prix élevé, via des **abonnements** annuels ou des licences. Ce marché pèse plusieurs centaines de millions d'euros. Néanmoins il semble difficile d'obtenir des informations précises et publiées par ces nouveaux acteurs du sport. L'essor de l'analyse de la performance favorise ces entreprises.

De nos jours, beaucoup de clubs créent des cellules d'analyses, composées de data scientifiques ou analystes vidéo. Par exemple, le TFC a été racheté par le fonds d'investissement RedBird Capital et collabore avec Zelus Analytics pour l'intégration de la data pour le recrutement ou encore les stratégies du club.



*Figure 2 : Association
Adidas+Kinexon pour le ballon
connecté de la coupe du monde
2022 au Qatar*

Pour finir, afin d'exploiter la donnée, il faut l'acquérir. Le domaine de **l'IoT²** dans le sport et le football se développe toujours plus d'année en année. L'objectif est de développer et positionner des capteurs sur tous les acteurs et objets du foot : cela passe par le joueur, le ballon, les chaussures, le ballon, des pilules (cf. § II.5), etc. On retrouve comme acteur des entreprises comme Capapult, Kinexon ou encore STATSports.

² IoT : Internet Of Thing (Internet des objets connectés)

I.3.2 Les usages commerciaux et dérivés

Au-delà des terrains, les données collectées trouvent de nombreux débouchés dans de nombreux secteurs, avec une forte valeur économique :

- **Dans le domaine de l' « Entertainment » :** La société EA Sports FC utilise des bases de données extrêmement détaillées pour créer et rendre réaliste son jeu FIFA. Ces données sont souvent issues d'un partenariat avec Opta. Dans les jeux récents chaque note, chaque caractéristique (vitesse, dribbles, tir) s'appuie sur une collecte rigoureuse de performances réelles de chaque joueur. La réalité augmentée constitue également un nouveau vecteur d'interactivité : ainsi, demain pourquoi ne pas être dans son salon, comme si nous étions dans un stade ? Cela permettrait aux clubs de vendre des nouveaux produits et ainsi augmenter le nombre de places vendues.
- **Dans le domaine des paris sportif :** Ce secteur génère plus de 1,8 milliard d'euros en France. Ces applications sont liées à la data avec des algorithmes de cotation s'appuient sur des milliers de données en temps réel pour estimer les probabilités, gérer les risques et adapter les cotes proposées aux parieurs.
- **Dans le domaine de l'audio-visuel :** Enfin, les chaînes comme Canal + ont largement démocratisé la data en l'intégrant dans leurs émissions d'avant et après match. Grâce à ses nombreux graphiques, Canal + valorise l'expérience spectateur en transformant la data en argument de vente pour fidéliser les abonnés. Cette valorisation de la donnée crée un cercle économique en rendant le football plus technique pour le grand public. Cela permet également de produire plus d'émissions ou d'augmenter la durée de celles-ci pour capter et garder le téléspectateur devant sa télévision et donc générer du revenu publicitaire.
- **Dans le domaine de la valorisation des stades :** en effet, la data est devenue un outil essentiel dans la gestion commerciale des stades. 2 axes sont les plus développées :
 1. **La sécurité :** les clubs analysent les flux de spectateurs pour optimiser l'accueil les points de vente ou encore la sécurité, cela évite les mouvements de foule trop important.
 2. **La gestion de la vente de produits dérivés :** grâce aux données collectées, les Stadium managers peuvent savoir quelle zone du stade sont les plus fréquentés et quels sont les habitudes de consommations des spectateurs en analysant les chiffres des buvettes, des boutiques.

La data permet donc de transformer un stade en un véritable lieu de consommation optimisé et donc de renforcer la stratégie marketing du club.

I.3.3 Règlementation et encadrement

Face à la croissance de l'usage des données dans le football, la FIFA³ et l'UEFA⁴ ont mis en place des initiatives comme le **FIFA Performance Programme** ou le **FIFA Innovation Hub**. Cela encadre et prône une utilisation éthique des technologies de la data. Le FIFA innovation Hub permet d'explorer, tester et intégrer les innovations technologiques dans le football.

³ Fédération Internationale Football Association

⁴ Union Européenne Football Association

Durant les matchs, la transmission aux entraîneurs de **données avancées** dépend des règlements des compétitions et des fédérations. En Premier League, l'utilisation des données est encadrée par des règles strictes pour garantir l'intégrité du jeu et l'équité sportive. Les clubs et analystes peuvent utiliser ces données pour préparer les matchs, mais l'accès aux données en temps réel pendant les matchs est limité. Ces données sont réservées aux analystes ou aux diffuseurs comme l'UEFA via son MatchCentre. Ces utilisations servent à enrichir l'expérience des spectateurs et téléspectateurs. Cependant les entraîneurs ont quand même la possibilité d'avoir accès en temps réel aux **données quantitatives** (évoqué en 2.1).

Pour finir les données biométriques sont considérées comme des données sensibles selon le **RGPD**⁵. Il faut donc avoir le consentement des joueurs et une sécurité renforcée pour pouvoir utiliser ces données. La propriété des données et la domination des entreprises privées soulèvent des questions sur ces exigences.

Le tableau ci-dessous met en évidence les domaines de protection et les articles du RGPD :

Article RGPD	Domaine protégé
Article 6, alinéa 1a	Consentement
Article 6, alinéa 1b	Contrat
Article 6, alinéa 1c	Conformité avec les obligations juridiques
Article 6, alinéa 1e	Intérêt public
Article 6, alinéa 1f	Intérêts légitimes

II Les apports de la data dans le football professionnel

II.1 Optimisation de la performance sportive des joueurs



Figure 3 : Fiche technique d'un joueur

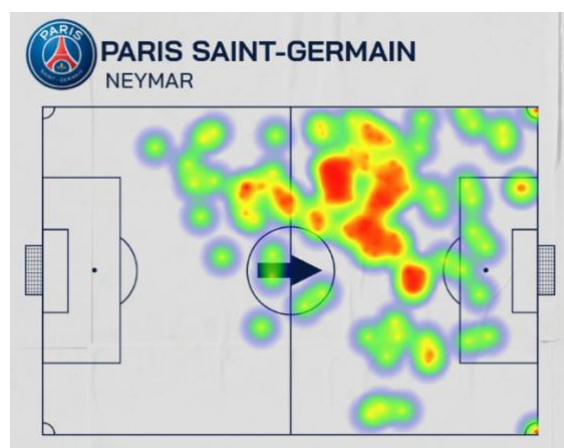


Figure 4 : Heatmap d'un joueur : carte de chaleur de Neymar durant la victoire 2-1 du PSG contre le RCL en 2021

⁵ Règlement Général de Protection des Données

Dans le football moderne, la recherche de la performance dépasse largement l'intuition et l'œil humain. L'intégralité des clubs professionnels a décidé depuis les années 2000-2010 d'intégrer dans leur quotidien de nombreux outils technologiques pour améliorer et optimiser leurs performances sportives. L'usage de capteurs GPS, d'outils de mesure biométriques et de plateforme d'analyse de données permettent aujourd'hui un suivi précis de chaque joueur à l'entraînement comme en match. Chaque joueur est alors analysé et noté sur différents critères :

- La figure 3 représente la **fiche technique** d'un joueur sur plusieurs statiques comme le nombre de but, le nombre de fautes, etc.
- La figure 4 met en avant les **positions préférentielles** de Neymar durant 1 match : cela reflète typiquement ce que peuvent amener les capteurs GPS portés par les joueurs.

Une autre donnée comptabilisée durant les matchs est **l'Expected Goals (xG)**. Cet indicateur mesure la probabilité qu'un tir se transforme en but selon sa position, son angle et bien d'autres critères.

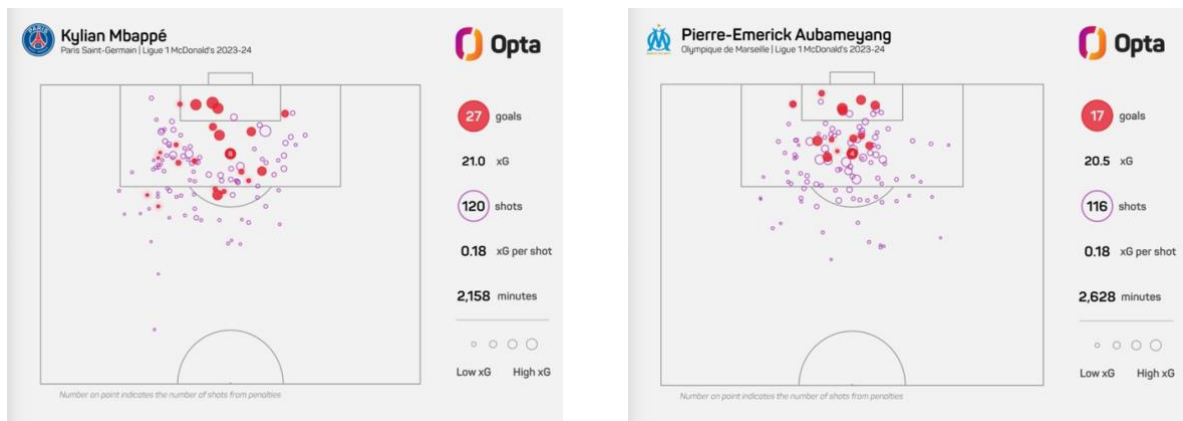


Figure 5 : Expected Goals de Mbappé vs Aubameyang durant la saison 2023-2024 :

Les xG prennent surtout leur sens au moment de cumuler les données d'un joueur ou d'une équipe sur un match ou une saison afin de comparer avec le total de buts inscrits sur la même période. :

- Dans le cas de Mbappé, lors sa saison de Ligue 1 2023-2024 on distingue qu'il a plus de but que d'Expected Goals, dans ce cas on peut dire que le joueur **surperforme** par rapport à la qualité des tirs qu'il se procure.
- Au contraire si on regarde le cas de Pierre-Aymerick Aubameyang, on observe qu'il a accumulé presque autant d'expected Goals (20,5 xG) avec un nombre comparable de tirs que Mbappé. Pourtant sur l'ensemble de la saison il a marqué 10 buts de moins, on peut donc dire que le joueur est moins réaliste devant le but.

Ces données jouent un rôle important pour les entraîneurs qui peuvent prendre du recul pour évaluer la capacité de leur équipe à se créer des occasions. Des changements sont donc étudiés à l'entraînement pour améliorer les performances de chacun.

Cette approche individualisée favorise un suivi plus fin et une personnalisation du travail pour chaque footballeur.

De nos jours, chaque entraîneur est accompagné de plusieurs data analystes pour préparer les matchs et entraînements. Ces data analystes récupèrent l'intégralité des données de tous les

joueurs, pour les traduire en indicateurs concrets pour les entraîneurs et préparateurs physiques grâce à des logiciels comme **Catapult** ou **STATSports**. Ces analyses de l'ombre sont de plus en plus importantes à l'heure actuelle où les plus grandes équipes disputent des matchs tous les 3 jours. L'analyse de la charge de travail permet par exemple d'ajuster les séances d'entraînement pour éviter des blessures dues à des sursollicitations.

II.2 Analyse des adversaires

La préparation d'un match passe aussi par une analyse fine des adversaires grâce aux données disponibles. Les analystes et l'entraîneur étudient les tendances tactiques de l'équipe adverse, que ce soient ses systèmes préférés, ses points faibles et les joueurs clés d'un match, à différents moments du match. Ensuite des **modèles** sont alors créés pour **prédire** les zones à attaquer ou les comportements sur coup de pied arrêtés.

En effet grâce aux données collectées et aux extraits vidéo, les datas analystes ont la capacité de repérer les déplacements typiques ou les combinaisons préférées des joueurs adverses, pour cela ils peuvent par exemple utiliser les heatmaps.

En amont de sa finale de la Ligue des champions en 2025, Le PSG a analysé les pieds forts du gardien et des défenseurs adverses, pour ajuster les pressings des attaquants. En effet si le défenseur est droitier l'attaquant va presser vers son pied droit pour qu'il dégage avec son pied gauche. Ces détails peuvent jouer un rôle important durant les matchs, pour la récupération du ballon.

Ces analyses tactiques sont alors cruciales pour bâtir une stratégie adaptée pour gérer les temps forts et faibles d'un match. Brentford choisit son plan de jeu selon les faiblesses statistiques de leurs adversaires. Brentford a été racheté par Matthew Benham, un ancien analyste financier et parieur professionnel. Il a racheté Brentford en 2012 et a appliquées les méthodes du moneyball (Baseball) et de la data avancée. Cette approche lui a permis de faire monter le club en première League, le 29 mai 2021.

II.3 Analyse tactique et aide à la décision pendant les matchs

Dans le football, l'accès aux données avancées est limité dans certaines situations pour les entraîneurs. En effet pendant les matchs officiels (FIFA, UEFA, LFP⁶, ...), l'accès aux données en temps réel diffère selon les fédérations, les championnats ou les compétitions. Cependant il est autorisé d'enregistrer des données GPS ou biométriques pour les analyses post match. Lorsque nous voyons les entraîneurs sur leur tablette, les données affichées sont souvent très limitées, on peut par exemple voir les replays des actions ou des statistiques basiques.

En comparaison, dans le rugby professionnel notamment au niveau international et en Top 14, les analystes et entraîneurs ont la possibilité de consulter les données avancées comme les GPS la biométrie ou d'autres types de statistiques. Ces données visibles en plein match, sont issues de capteurs GPS, gyroscopes, accéléromètres ou encore vidéo tracking. Dans le rugby les analystes en tribune travaillent en temps réel et peuvent communiquer directement avec le staff, on les voit d'ailleurs régulièrement avec des casques audios. Les analystes peuvent par exemple signaler un

⁶ LFP : Ligue Football Professionnel

problème de fatigue pour suggérer un remplacement. Ne pouvant pas tout analyser durant les matchs beaucoup d'autres indicateurs sont exploités en post match.

Il est possible d'expliquer cette différence d'utilisation en raison d'une culture du rugby historiquement ouverte à la science de la performance et à la dimension tactique. En revanche le football est plus conservateur et les instances veulent éviter que les entraîneurs « coachent par les algorithmes en live ». Pour autant, il n'est pas impossible que cette vision de l'utilisation de la data pendant un match évolue : en effet, l'aide à l'arbitrage par la vidéo, utilisée initialement par le rugby, est maintenant également utilisée dans le football.

II.4 Suivi de la santé des joueurs et des entraînements

La collecte continue de données permet également de suivre l'évolution de la forme physique des joueurs à court et moyen terme. En comparant les performances physiques sur plusieurs semaines, les data analystes ont donc la capacité d'identifier les signes de fatigue ou de baisse de forme des joueurs. Cette analyse permet aux entraîneurs de **prévenir les blessures**. Ces informations permettent d'adapter le programme d'entraînement de manière proactive, notamment en allégeant les charges pendant les entraînements pour les joueurs. Certaines équipes, comme le club de Monaco, vont encore plus loin en intégrant des algorithmes prédictifs pour anticiper les risques de blessures : Le documentaire, « La révolution DATA » de Canal+, montre que les joueurs reçoivent un retour régulier sur leurs efforts, les progrès et les points à améliorer.

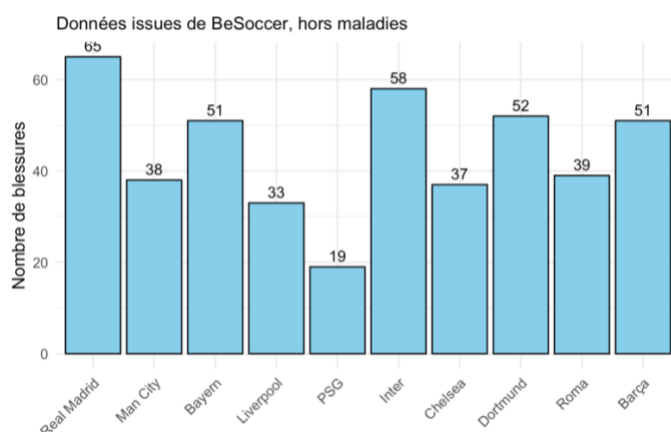


Figure 6 : Nombre de blessures physiques pour équipes du Top 10 Européen – Saison 2024/2025

On peut retrouver ces méthodes au PSG où le nombre de blessures recensées en 2025 est bien inférieur aux autres équipes du top 10 européens (Figure n°6). Ce résultat s'explique notamment par une meilleure utilisation de la data pour individualiser les charges d'entraînement ou d'optimiser la récupération. La rotation de l'effectif (4,4 changements par match) et le fait d'avoir les meilleurs analystes dans le domaine permettent le faible risque de

blessures. De plus le PSG possède l'une des équipes la plus jeune ce qui favorise la diminution de blessures malgré le rythme important de match durant la saison 2024-2025.

La ligue 1 étant considéré comme le 5^{ème} plus grand championnat, certains évoquent le fait que Luis Enrique a une plus grande marge de roulement dû à un championnat moins « concurrentiel », où dès le milieu de la saison le PSG dispose de plusieurs points d'avance sur le deuxième. L'ensemble de ces points peuvent expliquer ces chiffres exposés.

➔ La data permet donc d'améliorer le suivi de la santé des joueurs, mais elle ne peut pas être considéré comme l'unique facteur.

II.5 Préparation physique et prise en compte de l'environnement d'une compétition

En prévision des fortes chaleurs attendues lors de la Coupe du monde 2026 aux États-Unis, Au Canada et au Mexique, Thomas Tuchel, sélectionneur de l'Angleterre, a décidé de mettre en place une préparation spécifique et novatrice. Son objectif est de préparer l'équipe d'Angleterre aux fortes chaleurs attendues lors de la Coupe du monde. Pour cela, lors d'un stage à Gérone, des tentes chauffantes ont été installés pour simuler une température de 36°C et des exercices sous haute chaleur ont été réalisés. En parallèle, les joueurs ingèrent des **pilules électroniques** capable de mesurer leur température corporelle interne et d'autres **paramètres physiologiques**. Ces méthodes déjà testées en athlétisme ont pour but d'adapter l'hydratation et l'intensité des efforts pour chaque joueur. Cette initiative permettra peut-être de limiter les risques de fatigue et coup de chaleur.

II.6 Améliorations du recrutements et valorisation des joueurs

Dans le football, le recrutement s'est largement transformé avec la data. Désormais, les clubs ne se basent plus uniquement sur les observations des recruteurs, mais aussi sur des milliers de données objectives. Ces énormes bases de données permettent d'analyser non seulement les ligues supérieures mais également les ligues secondaires et championnats exotiques. Ces méthodes ont été aperçues par la première fois en Angleterre mais également en France à Toulouse. Le TFC⁷ collabore depuis plusieurs années avec la start-up Zélus Analytics qui développe des algorithmes pour repérer des profils précis. Ces algorithmes permettent de choisir un entraîneur ainsi que des joueurs qui vont avoir la capacité de développer un bon jeu d'équipe en peu de temps car chaque profil de l'effectif est choisi pour s'intégrer à l'équipe et au style de l'entraîneur. De plus les entraîneurs sont aussi modélisés, notamment à travers l'analyse de leur style de jeu, leurs décisions tactiques et leur impact statistique sur la performance de l'équipe.

Concernant Toulouse, ce choix a été payant car l'équipe a gagné le 29 avril 2023 La Coupe de France contre le FC Nantes quelques mois après son titre de champion de France de Ligue 2.

➔ L'utilisation de la donnée a donc été un **phénomène accélérateur du temps**. Pour des clubs de « milieu de tableau », comme le TFC, la notion de style de jeu et de joueurs dédiés pour un style étaient antérieurement basés sur les centres de formation. Le club de Nantes a longtemps mis en avant pour la qualité de ses joueurs issus de son centre de formation et collant à son style de jeu « à la Nantaise ». Mais ces pratiques peuvent prendre plus de temps.

La data permet également de **valoriser un joueur** en anticipant son potentiel de progression ou en mettant en avant des qualités invisibles à l'œil nu, comme le pressing réussi ou la création d'espaces. Ces points importants dans le jeu d'une équipe n'étaient auparavant jamais mis en avant. L'objectif est donc double pour ces clubs :

1. cela permet d'améliorer l'équipe sportivement à moindre coût,
2. tout en réalisant dans certains cas de nombreuses plus-values lors des transferts.

⁷ TFC : Toulouse Football Club

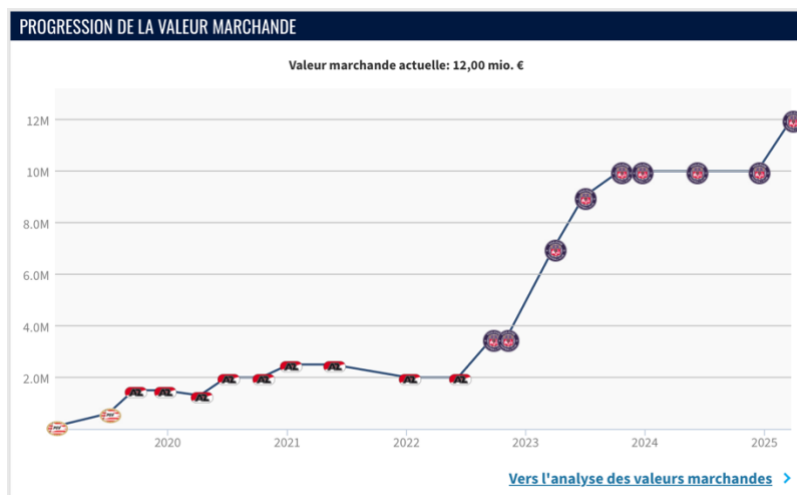


Figure 7 : Evolution de la valeur marchande de Zakaria Aboukhal

des déplacements sans ballon et une polyvalence offensive. Durant la saison 2022-2023 il a su marquer les esprits en marquant 6 buts et 5 passes décisives avec notamment un but en finale de la Coupe de France. Il a performé en Ligue 1 mais également durant la Coupe du monde 2022 où son pays, le Maroc, arrive en demi final. Pour conclure ce joueur représente un recrutement intelligent de la part du TFC car il a été acheté à un prix raisonnable, et a un impact direct sur le terrain avec une forte valorisation sportive et financière aujourd'hui.

Avant d'arriver pour 2 millions d'euros en juillet 2022 au TFC⁸, Aboukhal a évolué au PSV Eindhoven et à l'AZ Alkmar. Les analystes du TFC ont su remarquer que le joueur souvent remplaçant montrait grâce aux données une efficacité prometteuse lorsqu'il jouait. Ils ont détecté chez lui de nombreux pressings hauts, des créations d'occasions depuis les ailes,

II.7 Autres utilisations de la data dans le football : aide à l'arbitrage

La data n'est pas seulement utilisée par les clubs professionnels, du côté de l'arbitrage, on peut par exemple parler de la VAR et la détection semi-automatisée du hors-jeu, utilisée pour la première fois lors de la Coupe du monde 2022 au Qatar (Figure 9). Ainsi, chaque arbitre s'aide des données pour préparer le match et savoir où se positionner durant le match. Prenons par exemple la finale de la Ligue des champions 2024-2025, l'arbitre du match aurait pu se positionner plus souvent du côté droit de l'attaque parisienne en formant un triangle avec son arbitre de ligne et son 4e arbitre pour pouvoir avoir une meilleure vision sur l'endroit où se déroulait la plupart des fautes concédées par les Parisiens (Figure 8).



Figure 8 : La data pour aider l'arbitre

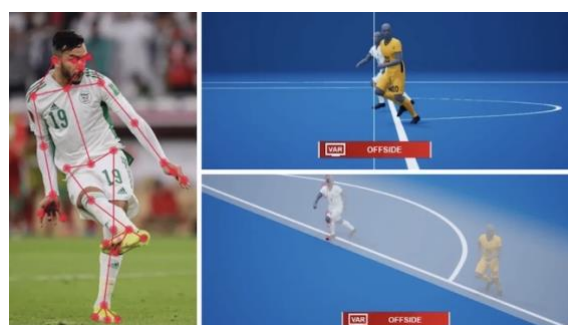


Figure 9 : Image du hors-jeu semi-automatique

Lors du tout nouveau championnat du monde des clubs qui va débiter le 14 juin, l'IA⁹ va être utilisée pour la première fois pour la détection du hors-jeu.

⁸ TFC : Toulouse Football Club

⁹ IA : Intelligence Artificielle

III Les limites et les risques associés à la data

L'utilisation des données dans le football professionnel représente une avancée majeure. Néanmoins, si la data est mal exploitée ou utilisée sans cadre, elle peut créer de nombreux problèmes ou dérives comme :

- Accentuer une **pression sur les joueurs** qui seraient focaliser sur leurs statistiques et non sur la performance d'équipe.
- Accentuer d'avantage, la **différence entre le football amateur et professionnel**. Ces datas ne sont également pas disponibles pour tous les clubs. Les grands clubs, dotés de moyens importants, peuvent se permettre d'embaucher des data analysts, d'acheter des logiciels spécialisés, ou encore d'équiper leurs joueurs de capteurs biométriques dernière génération. À l'inverse, les clubs de divisions inférieures, ou issus de pays à plus faibles ressources, n'ont pas accès à ces outils. Les moyens autour de l'arbitrage d'un match ne seront pas également les mêmes.
- Créer une **fracture technologique**. Dans un sport déjà marqué par les écarts financiers, la data peut devenir un nouveau facteur d'exclusion, renforçant l'écart entre l'élite et clubs modestes. Si Brentford ou Midtjylland sont souvent cités comme exemples de "petits" clubs qui ont réussi grâce aux statistiques, ces exemples pourraient devenir les derniers, avec les investissements majeurs réalisés par les grands clubs.

III.1 Déshumanisation et pression sur les acteurs du football

III.1.1 Les joueurs réduits à des chiffres

Avec la montée en puissance de la data, les joueurs sont de plus en plus évalués selon des indicateurs chiffrés : nombre de kilomètres parcourus, duels gagnés, expected goals, pressing effectué. Cela peut donner une vision froide et désincarnée du football. Le joueur n'est plus vu comme un être humain, mais comme une série de statistiques.

Cette logique peut avoir des effets psychologiques importants. Certains joueurs ressentent une pression constante sur leur niveau de performance au détriment du plaisir de jouer. Darwin Núñez en est l'exemple : l'attaquant de Liverpool a été fortement critiqué à ses débuts car ses expected goals étaient élevés, mais il ne marquait pas assez. Les fans et médias l'ont jugé uniquement sur ce chiffre, oubliant ses appels, son pressing ou ses déplacements. Le risque est de transformer les footballeurs en produits standardisés, constamment comparés, notés, suivis.

III.1.2 Jugements « simplistes » par les médias et supporters

Le football est aussi un spectacle, et les statistiques spectaculaires comme les buts ou les passes décisives sont souvent mises en avant dans les médias. Mais cela occulte d'autres aspects essentiels du jeu, moins visibles : le repli défensif, la récupération de ballons, les appels sans ballon.

On peut citer Mbappé, malgré son début de saison en demi-teinte, décevant sur le plan individuel, Il est resté décisif dans les moments clés au Réal Madrid. Cependant la presse et les supporters

Espagnol ont eu tendance à résumer ses performances à ses statistiques immédiates, sans considérer le contexte global. Les jugements simplistes ont parfois ignoré l'impact invisible qu'il a sur le jeu, il a d'ailleurs reçu le trophée du soulier d'or¹⁰ pour sa saison. Le contraste entre ses apports réels et la perception extérieure illustre bien une tendance à juger les joueurs sur des impressions rapides, plutôt que sur l'analyse de leur rôle tactique et de leur influence collective.

Certains joueurs sont donc injustement jugés, voire critiqués, simplement parce qu'ils ne brillent pas dans les statistiques "vendeuses". À l'inverse, la data peut aussi valoriser des rôles longtemps sous-estimés, comme des milieux défensifs. N'Golo Kanté n'a jamais été spectaculaire en termes de buts ou passes décisives (64 matchs = 1 but et 2 passe décisives), cependant grâce aux données de récupération ou de kilomètres parcourus on peut voir que sa présence sur le terrain a été bénéfique. Par exemple pendant le match Autriche-France (1-0) à l'Euro 2024, N'Golo Kanté a récupéré 9 ballons et parcouru 11,8 km sur le terrain ces statistiques sont impressionnantes et pas forcément mises en valeur mais décisives durant un match.

➔ Si aujourd'hui encore, les attaquants sont plus mis en avant que les autres positions, il conviendra aux observateurs du jeu (supporters ou journalistes) de mettre en avant toutes les données du jeu pour valoriser chaque acteur d'une équipe.

III.1.3 Le risque d'une perte d'instinct

Le football est un sport important dans le cœur de beaucoup de personnes, il est alors beaucoup jugé et analysé, ce qui peut provoquer un risque de brider la créativité des joueurs. Certains entraîneurs ou observateurs parlent d'un football trop programmé, où les joueurs n'osent plus prendre de risques s'ils ne sont pas "validés" par les données.

Cette tendance à l'**hyper-analyse** peut conduire à une paralysie de l'action : trop réfléchir, trop anticiper, au lieu d'agir instinctivement comme en témoigne l'Italie ou certains analystes dénonçaient un jeu trop rigide, où les joueurs appliquent des schémas dictés par les datas au détriment de la spontanéité. Lors du match Italie-Macédoine du Nord (éliminatoires CDM 2022), l'Italie a eu **32 tirs cadrés pour 0 but**. La domination était "statistiquement parfaite", mais ils ont manqué de créativité dans les 20 derniers mètres.

Enfin, on oublie régulièrement que le football est un **sport d'équipe** et pas une somme de talents individuels. On peut par exemple citer le PSG qui avait sur le papier les meilleurs joueurs du monde (Messi, Mbappé et Neymar) en attaque, malgré cet effectif, le PSG n'a pas remporté la Ligue des champions durant cette période.

III.2 Des données peu fiables ou incomplètes

La qualité des données n'est pas homogène partout. Dans certains championnats, notamment dans les pays en développement, les données sont peu précises, voire inexistantes. Cela pose un problème d'équité, notamment lors du recrutement, un joueur talentueux peut être sous-estimé parce qu'il évolue dans un championnat où les données ne sont pas suivies correctement.

¹⁰ Soulier d'or : Le Soulier d'or européen est une récompense attribuée au meilleur buteur de tous les championnats de football de première division d'Europe confondus

Cela peut expliquer par le : “**garbage in, garbage out**”. Si les données de départ sont biaisées ou incomplètes, les analyses et décisions qui en découlent, risquent d’être erronées. On peut citer Rasmus Højlund qui a été recruté par Manchester United en 2023 pour environ 70 millions d’euros. Ce recrutement s’est appuyé sur des données prometteuses : vitesse, puissance, profil de buteur moderne, fort potentiel statistique (xG, frappes, appels en profondeur). Dans les faits, il sortait d’une seule bonne demi-saison à l’Atalanta de Bergame, avec peu de buts en Série A. Son profil “data” était flatteur, mais trop limité dans le temps, et centré sur des données telles que son style de jeu, sa technicité. Finalement sa première saison a été très compliquée et cet exemple est la preuve que si les données de base sont partielles ou mal interprétées, une erreur de casting est possible.

- ➔ La data doit ainsi être considérée comme une aide supplémentaire : les analyses data doivent toujours être complétés de rapports visuels réalisés par des recruteurs ou des analystes vidéo.

III.3 Des données falsifiables ?

Les enjeux économiques du football étant très forts, certaines dérives pourraient naître, comme la tentation pour un club ou un agent de manipuler des statistiques afin de maximiser la valeur marchande d’un joueur. Le marché du football est actuellement influencé par les statistiques notamment dans le domaine du recrutement (cf. § II.6). Il existe donc un risque ou certain acteur tenterait de biaiser les données diffusées aux recruteurs en transmettant uniquement les plus flatteuses. C’est dans ce contexte que certains acteurs s’inquiètent d’une potentielle **dépendance excessive à des données** issues de prestataires privés comme Opta ou StatsBomb. Ces prestataires pourraient sous pression d’un club ou intérêt commercial fournir des statistiques erronées aux concurrents.

En parallèle, les enjeux économiques étant toujours plus importants, ces prestataires devront protéger les données du risque de **cybercriminalité**. Il leur faudra justifier de la bonne qualité de l’information et qu’elle n’a pas été modifiée par des tiers mal intentionnés.

- ➔ La cybersécurité autour de la donnée devient donc un enjeu majeur pour les clubs et les acteurs du football, qui doivent protéger leurs données sensibles concernant les joueurs, les entraîneurs et les équipes, compte tenu des potentiels risques de piratage ou de fuite d’information.

III.4 Problèmes éthiques et dépendance technologique

III.4.1 La vie privée des joueurs en question

Les données récoltées ne concernent pas seulement les performances sportives, mais aussi l’intimité des joueurs : rythme cardiaque, sommeil, alimentation, géolocalisation. Cela pose de vraies questions éthiques. Le club de Valence CF a par exemple testé des capteurs durant le sommeil de certains joueurs. Cela a suscité une controverse, certains estimaient que cela envahissait leur intimité, en particulier si les données sont utilisées pour juger la forme mentale ou l’hygiène de vie. De plus Cristiano Ronaldo, très soucieux de sa vie privée, a souvent refusé

l'utilisation des capteurs afin de collecter des données biométriques en dehors d'un match pensant que la data dans le sport doit rester dans le sport et non en dehors d'un terrain de foot.

- ➔ Cette surveillance permanente peut ainsi être perçue comme une intrusion et une remise en cause de la notion de « **vie personnelle / vie professionnelle** ». Si un joueur est constamment observé, même hors du terrain, cela remet en cause sa **liberté individuelle**. À long terme, cela peut générer de la méfiance, du stress, voire des conflits entre joueurs et encadrement. Le joueur devra ainsi être attentif aux clauses contractuelles et à ses obligations envers le club, qu'il aura signé dans son contrat.

III.4.2 Une dépendance croissante à la technologie

Plus les clubs s'appuient sur les données, plus ils deviennent dépendants de la technologie. Est-ce qu'un coach peut encore prendre une décision sans la confirmation des données ?

Laurent Batlles, ancien entraîneur de l'AS Saint-Étienne avait exprimé ses réserves sur l'importance accordée aux statistiques et aux analyses de données en insistant sur la phrase « la prise d'information humaine est primordiale ».

Cette dépendance fragilise les équipes qui ne développent plus assez leurs compétences humaines : sens du jeu, intuition, expérience. L'équilibre entre données et intelligence humaine devient alors crucial.

- ➔ Même s'il y a une aide de la technologie et que celle-ci prend toujours plus de place dans le haut niveau, à la fin, si toutes les équipes disposent de ces mêmes ressources, la force d'une équipe sera toujours :
- Dans sa capacité à respecter un plan de jeu,
 - Surprendre ses adversaires en sachant sortir de ses habitudes.

III.4.3 Le cas de la VAR : la data ne résout pas tout

La VAR (assistance vidéo à l'arbitrage) est l'exemple parfait d'un outil technologique censé apporter plus de justice mais qui reste fortement critiqué. Malgré les ralentis, les lignes de hors-jeu automatisées, et les analyses millimétrées, les décisions continuent de faire débat.

En effet, certaines actions reposent toujours sur une **interprétation humaine**, même avec les données : Une main ou une faute peuvent être analysées de manière différente selon les arbitres comme pendant le match Tottenham - Liverpool (sept. 2023), un but valable de Luis Díaz a été refusé à cause d'une erreur humaine dans le traitement de la VAR. L'image était claire, mais le message transmis à l'arbitre était erroné.

Avant son arrivée, la VAR était mise en avant comme l'outil qui allait révolutionner l'arbitrage, en rendant les décisions plus justes. De plus, l'ensemble des technologies nécessaires à son fonctionnement peut constituer une inégalité entre les différents niveaux d'une ligue, le niveau professionnel et amateurs et entre les différentes compétitions. De nombreux débats existent entre les différentes instances représentatives du football pour décider au mieux l'intérêt de l'apport de la donnée.

Conclusion

L'utilisation croissante des données dans le monde du football professionnel a bouleversé les pratiques, les entraînements et les perceptions de ce jeu. En effet, elle permet d'optimiser les performances par une meilleure compréhension de l'effort physique, une prévention fine des blessures, une analyse tactique plus affinée.

En aval elle aide à recruter des talents sous-évalués offrant aux clubs à plus petits budgets des moyens de rivaliser via l'innovation. Enfin, cela permet d'analyser les adversaires, d'apporter des aides à l'arbitrage comme avec la VAR, d'offrir aux supporters des expériences personnalisées, etc. La data s'invite de fait dans tous les aspects du jeu.

Malheureusement ces apports ne peuvent pas être toujours bénéfiques. Si la data rend les décisions plus rigoureuses, plus rationnelles, elle pourrait accentuer des inégalités structurelles entre les clubs, réduire des joueurs à des projets statistiques, alimenter des jugements rapides des supporters et médias. Les risques de déshumanisation sont bien réels, tout comme ceux liés à la vie privée ou à une dépendance excessive à la technologie

Le constat est donc nuancé : la donnée est un formidable instrument d'aide à la performance et à la décision, elle ne peut être considérée comme miraculeuse ou comme substitut à l'intuition, l'émotion et l'expérience humaine ; son usage doit être éthique, encadré et réfléchi, respectueux des acteurs du jeu autant que de son essence.

Dans cette dynamique, de nouveaux métiers ont émergé, comme celui de data analystes, désormais présent dans de nombreux clubs professionnels. Demain, l'irruption de l'intelligence artificielle posera de nouveaux défis : quelles seront les limites de l'automatisation dans les choix tactiques ? Quel rôle jouera encore l'entraîneur, ou le directeur sportif ? À quel moment jugera-t-on qu'un joueur n'ait pas assez "performé" selon les algorithmes, et avec quelles conséquences sur sa carrière ?

Le football reste un sport collectif, où le geste juste, parfois invisible dans les statistiques, prime souvent sur la performance individuelle. L'émotion du jeu, l'imprévu, le rôle du vestiaire ou du mental ne se capturent pas encore entièrement en données. C'est là tout l'enjeu : continuer à faire progresser le football sans lui faire perdre son âme.

Bibliographie

Reportages :

- **La révolution DATA** : Canal + ; Documentaire du 05/08/2024 :
https://www.canalplus.com/sport/canal-foot-manager-saison-2/h/24864742_50001
- **Mercato : la data, "un formidable moyen de gagner du temps" pour les clubs de foot** : TF1 ; Reportage du 4 janvier 2024 par Maxence Gevin
<https://www.tf1info.fr/sport/interview-mercato-un-formidable-moyen-de-gagner-du-temps-pourquoi-la-data-est-de-plus-en-plus-utilisee-par-les-clubs-de-foot-2281472.html>

Webographie :

- Blessures équipes TOP 10 :
<https://fr.besoccer.com/equipe/blessures-sanctions/paris-saint-germain-fc>
- Valeur marchande Aboukhal :
<https://www.transfermarkt.fr/zakaria-aboukhla/marktwertverlauf/spieler/393238>
- Rôle de la DATA :
<https://insights.morganphilips.com/fr/le-role-predominant-de-la-data-dans-le-sport>
- Les secrets d'Opta :
https://decouverte.challenges.fr/entreprise/sports/les-secrets-d-opta-societe-reine-des-data-du-foot-mondial_602577
- Opta Analytics :
<https://www.statsperform.com/opta-analytics/>
- Brentford Football Club :
https://fr.wikipedia.org/wiki/Brentford_Football_Club
- Impact de la data :
<https://larevuedestransitions.fr/2024/03/30/quel-est-limpact-de-la-data-dans-le-sport/>
- Hors-Jeu semi-automatique :
<https://real-france.fr/le-real-madrid-va-inaugurer-le-hors-jeu-semi-automatique/>
- Tout comprendre des Expected Goals :
https://ligue1.fr/fr/articles/l1_article_497-le-meilleur-tutoriel-pour-comprendre-les-expected-goals
- Classement UEFA :
<https://fr.uefa.com/nationalassociations/uefarankings/club/?year=2025>
- Clés de la « Data Foot » :
<https://www.formationfootball.fr/blog/data-foot#:~:text=Les%20data%20foot%20quantitatives%20repr%C3%A9sentent,impact%20tactique%20d'une%20action>
- Data Liverpool :
<https://www.liverpoolfc.com/news/behind-the-badge/398645-ian-graham-liverpool-fc-behind-the-badge>

- TFC : ZELUS, l'arme d'analyse de données :
<https://www.ladepeche.fr/2022/06/09/tfc-la-societe-zelus-larme-exclusive-danalyse-de-donnees-derriere-le-recrutement-toulousain-10348051.php>
- Football Manager :
<https://www.telegraph.co.uk/football/2018/11/02/football-managers-scouting-network-unearths-wonderkids-real/>
- Préparation Angleterre 2026 :
https://rmcsport.bfmtv.com/football/coupe-du-monde/coupe-du-monde-2026-tentes-chauffantes-pilules-les-etonnantes-methodes-de-l-angleterre-pour-anticiper-les-fortes-chaleurs_AV-202506060626.html
- Arbitrage CDM Clubs :
<https://www.lequipe.fr/Football/Actualites/Des-arbitres-munis-de-cameras-un-hors-jeu-semi-automatique-queelles-seront-les-nouveautes-experimentees-lors-de-la-coupe-du-monde-des-clubs/1568448>
- Apports de la DATA :
<https://balises.bpi.fr/data-et-football/>
- DATA dans le football Français :
<https://www.simonleon.fr/post/ou-en-est-la-data-dans-le-football-francais>
- Réglementation :
<https://www.entsportslawjournal.com/article/1082/galley/4956/view/>
- Réglementation Premier League :
<https://www.premierleague.com/player-privacy-policy>
- La data révolution :
<https://www.google.com/url?sa=i&url=https%3A%2F%2Fwww.lequipe.fr%2Fexplore%2F09-la-data-revolution%2F&psig=AOvVaw2FuzlrQcK3iDQlvsEa2n7&ust=1750075066407000&source=images&cd=vfe&opi=89978449&ved=0CBgQ3YkBahcKEwjAzaXVsfONAxUAAAAAHQAA AAAQBA>