

---

**PFE - PROJET DE FIN D'ETUDES - DECEMBRE 2019**

---



## **Le Big Data et l'Intelligence Artificielle dans le sport professionnel**

Conçu et rédigé par :  
Charles Fayolle, N° 20170118  
Eliott Lam-Seck N° 20170114

Sous la direction de :  
Clément Levallois – Professeur associé

## **REMERCIEMENTS**

Premièrement, nous aimerions remercier le professeur ayant supervisé ce PFE, Clément Levallois, pour le temps qu'il nous a consacré, ses recommandations, conseils et encadrement du sujet.

Deuxièmement, nous aimerions remercier David Banget, Directeur Digital de l'Olympique Lyonnais pour nous avoir accordé un peu de son précieux temps pour nous aider à mieux comprendre les enjeux de l'exploitation des données dans le marketing sportif.

Enfin, nous aimerions remercier nos amis Vincent Latimier et Valentin Styza pour nous avoir apporté de nouvelles perspectives et idées sur notre travail.

## SOMMAIRE

<b>LEXIQUE.....</b>	<b>3</b>
<b>INTRODUCTION .....</b>	<b>4</b>
<b>PARTIE 1 : Une amélioration des performances sportives .....</b>	<b>7</b>
A – Optimisation des performances physiques .....	11
B – Un outil d’aide à la prise de décisions tactiques .....	15
C – Détection des talents de demain à moindre coût : le scouting 2.0 .....	20
<b>CONCLUSION PARTIE 1 : United Managers, un aperçu du futur ? .....</b>	<b>31</b>
<b>PARTIE 2 : Un bouleversement dans et autour l’arène.....</b>	<b>34</b>
A – Enrichir l’expérience du spectateur au stade .....	35
B – Mieux cibler les supporters .....	39
C – Augmenter les revenus liés au sponsoring et autres partenariats.....	44
<b>CONCLUSION PARTIE 2 : l’approche individualisée d’un ensemble.....</b>	<b>48</b>
<b>PARTIE 3 : L’avenir de l’IA et du Big Data dans le sport remis en question.....</b>	<b>50</b>
A – Des apports mais aussi beaucoup de limites .....	50
B- L’amélioration continue de l’IA: suffisant pour contrer les limites actuelles ?.....	57
C- L’omniprésence de l’IA, vers un sport déshumanisé ? .....	61
<b>CONCLUSION PARTIE 3 : Un mariage et un équilibre à trouver entre la technologie et l’homme .....</b>	<b>64</b>
<b>CONCLUSION .....</b>	<b>67</b>
<b>SOURCES .....</b>	<b>73</b>

## LEXIQUE

### ***Big Data :***

Le « Big Data », traduit en français par les termes « mégadonnées » ou « données massives », n'a aucune définition précise ou universelle ; celle-ci variant selon les acteurs qui utilisent le terme. Toutefois, le concept se caractérise par les 3 grands principes définis par le modèle des « 3V » développé par l'entreprise de conseil Américaine Gartner en 2001.

Volume : on parle de Big Data lorsqu'il y a un volume considérable de données à traiter (seuil variable selon les utilisations et les utilisateurs et évoluant avec la forte croissance de la quantité de données stockées).

Variété : Le Big Data rassemble des données de diverses natures (brutes, semi-structurées ou non structurées), au formats variés (web, image, texte...) et de sources diverses (base de données publiques, données internes...)

Vélocité : On parle ici de la fréquence à laquelle les données sont générées, capturées, partagées et mises à jour. Le Big Data induit ainsi des flux en temps réel ou quasiment.

### ***Intelligence artificielle (abrégée en IA) :***

Ensemble de théories et de techniques mises en œuvre en vue de réaliser des machines capables de simuler l'intelligence humaine.

### ***Machine learning (traduit en français par « apprentissage automatique ») :***

Processus de fonctionnement d'un système d'intelligence artificielle par lequel le système est doté d'une capacité à apprendre à résoudre des tâches pour lesquelles il n'a pas été explicitement programmé.

### ***Deep learning (traduit en français par « apprentissage profond »)***

Ensemble de méthodes d'apprentissage automatique tentant de modéliser avec un haut niveau d'abstraction des données grâce à des architectures articulées de différentes transformations non linéaires.

## INTRODUCTION

Nous vivons dans une période où le rythme d'apparition des avancées technologiques s'accélère. Le champ des possibles pour l'humain semble aujourd'hui infini. Nous ne cessons de découvrir des nouveaux principes, inventer des nouveaux procédés, services ou produits et améliorer tout ce qui nous entoure. On effectue des progrès dans tous les domaines : la santé, l'écologie, la mécanique, l'informatique etc. Tous ces progrès ont une chose en commun : la technologie. Toutes les inventions et autres améliorations technologiques sont appliquées à ces domaines et nous permettent de ce fait de les faire grandir.

L'expression « Big Data » est apparue dans le début des années 2000 comme indiqué dans le lexique. C'est justement le progrès technologique qui a permis à ce principe de voir le jour tant nous étions en mesure croissante de récolter des données. De plus en plus de données à la fois pertinentes et exploitables. Exploitable pour l'analyse mais de manière plus importante pour effectuer des recommandations. A travers une pléthore de données, nous sommes capables de toutes les croiser, les interpréter et en tirer des conclusions sur un phénomène. On peut y voir des modèles, des répétitions pour créer des statistiques et ainsi prédire le futur grâce au passé.

Toutes ces données seraient cependant impossibles à exploiter sans une intervention technologique. En effet, elles sont étroitement liées à l'intelligence artificielle, au machine learning, au deep learning. Il est vrai qu'une telle quantité de donnée requière de l'aide, automatisée qui plus est. A ce stade, on peut observer cela dans de nombreux exemples présents dans notre quotidien. Les recommandations personnalisées de notre compte Netflix, la correction de trajectoire des nouvelles voitures, l'anticipation des crises météorologiques, un meilleur suivi de notre santé, le Big Data et l'intelligence artificielle sont partout.

Un des champs d'application ayant le vent en poupe est le sport. Véritable phénomène, prisme et reflet de notre société, le sport prend de plus en plus d'ampleur et d'importance. C'est une véritable échappatoire pour pratiquants et supporters mais c'est aussi un enjeu économique et sociétal important. Ce n'est d'ailleurs pas pour rien que les états lui accordent de plus en plus de budget et de temps. La Chine en est le parfait exemple. Elle a fait du football la priorité en termes de développement sportif et le président Xi Jinping a annoncé vouloir être en mesure de gagner la coupe du monde 2026. Alors pourquoi accorder autant d'importance à un « simple » sport ? Tout simplement parce le football est le sport le plus populaire au monde. Le président

Chinois a bien compris que le rayonnement international passe obligatoirement par ce sport. Les joueurs et clubs d'un championnat sont des ambassadeurs parfaits pour leur pays et les droits télévisions représentent une manne financière importante. Appuyés par des géants comme Alibaba ou Evergrande, les clubs investissent des sommes colossales pour attirer les plus grandes stars et ainsi rendre la Chinese Super League de plus en plus attractive. Le Football est devenu le sport numéro un à l'école et les meilleurs éducateurs sportifs étrangers sont attirés par le gouvernement pour aider à faire exploser ce sport. A travers cet exemple, on comprend que des sports comme le football ne sont pas simplement une activité physique mais incarnent des enjeux économiques, de société, politiques et même géopolitiques.

Avec une telle importance, tous les acteurs convergent vers une idée : améliorer constamment les sports et tout ce qu'ils impliquent. Améliorer les performances des athlètes, la notion de spectacle et de « show », les infrastructures, le fonctionnement des championnats etc. Tout cela dans le but, comme souvent, d'engendrer de plus en plus de revenus et de profits. Cette quête d'amélioration passe automatiquement par le progrès technologique. Avoir les meilleures machines pour s'entraîner, avoir les meilleurs stades, les meilleures tactiques, les meilleures prédictions et tout cela grâce au progrès technologiques. Progrès en termes de données et d'intelligence artificielle. Comment avoir le plus de statistiques possibles, les meilleures statistiques, la meilleure analyse, la meilleure application, comment rendre le sport plus attractif. L'utilisation de la data est exacerbée dans les sports collectifs professionnels. En effet, par rapports aux sports individuels, ils génèrent (en général) plus de revenus et ont une plus grande dimension tactique, populaire ou financière. Le champ d'application y est donc plus important et laisse plus de place aux progrès, aux essais et aux investissements.

C'est donc pour toutes ces raisons et pour notre passion commune pour le sport que nous avons décidé de nous pencher sur le sujet. En observant toutes les nouvelles applications du Big Data dans notre quotidien ainsi que dans le sport, nous avons voulu comprendre et répondre à la question suivante : « **Quel est l'impact du Big Data et de l'intelligence artificielle dans les sports collectifs professionnels ?** ». Ainsi nous avons essayé à travers de nombreuses recherches d'offrir une réponse à cette question et cela en trois parties. Le sport de haut niveau ne concerne pas seulement l'aspect athlétique. Il ne concerne pas non plus seulement l'aspect tactique ou encore technique. L'enceinte sportive ainsi que les supporters sont des acteurs tout aussi importants. Les partenaires et sponsors participent à faire vivre le sport professionnel. Le progrès technologique est impératif pour faire grandir le sport collectif dans tous les sens du

terme. Il ne faudrait cependant pas tomber dans l'excès et en abuser tant cela risque de dénaturer ce que nous affectionnons.

En premier lieu, nous avons étudié en quoi le Big Data participe à l'amélioration des performances sportives. Dans un second temps nous verrons comment ces avancées bouleversent ce qu'il se passe autour et dans les arènes sportives. Enfin, nous tenterons de modérer le sujet en exposant les limites de l'application du Big Data dans le sport.

## PARTIE 1 : Une amélioration des performances sportives

Le 13 Juillet 2014, l'équipe d'Allemagne remporte la Coupe du monde face à l'Argentine dans le mythique stade du Maracaña. Si l'unique buteur Mario Götze apparait comme le héros du jour, l'Allemagne a disposé tout au long de la compétition d'une arme secrète ayant fortement participé à la conquête du premier trophée international du pays depuis la réunification en 1990. Son nom ? Match Insights, un logiciel développé sur mesure dans le cadre d'un partenariat entre la Fédération Allemande de Football et l'éditeur SAP.

Si l'ensemble des 32 nations présentes au mondial disposaient d'analystes vidéos au sein de leurs équipes techniques, seul le staff allemand disposait d'un tel dispositif. Regroupant de vastes bases de données sur les joueurs de l'équipe d'Allemagne et ceux des sélections adverses, ce logiciel permettait à tous les joueurs et membres du staff d'accéder à ces informations de façon simple et visuelle depuis un téléphone, un ordinateur ou une tablette. Rien n'étant le fruit du hasard à ce niveau-là, ce logiciel est le résultat d'un projet qui voit le jour en 2012 à travers la compilation de données sur tous les joueurs et équipes qualifiées pour la compétition. La clé : chaque joueur est associé à un identifiant unique permettant de traquer l'ensemble de ses mouvements ainsi que divers indicateurs tels que le nombre de touches de balle, le temps de possession moyen ou la vitesse d'exécution.

L'impact de ce logiciel est mesurable via les commentaires du staff et des joueurs eux-mêmes. L'un des assistants du sélectionneur, Hansi Flick, a ainsi déclaré avant la demi-finale contre l'Allemagne : *« Les étudiants en sport de Cologne ont étudié de façon très détaillée notre adversaire, scruté chacun de leur matches, l'ensemble des articles de presse dans lesquels ils étaient mentionnés, et passé au microscope toute les informations les concernant pour nous les rendre disponibles. Nous avons désormais une énorme base de données sur laquelle se reposer et, ensemble avec nos scouts, nous sommes capables d'étudier notre adversaire et de préparer nos plans pour le match. C'est un projet sur lequel on travaille énormément depuis 2 ans. Nous avons réussi à recueillir des informations précieuses à partir des données de ces étudiants. Cela nous a beaucoup aider à nous préparer »*<sup>1</sup>.

---

<sup>1</sup> Germany's Secret World Cup Weapon: Big Data, Jack Rosenberger, CIO Insights, 18 juillet 2014



Un match que les allemands gagneront facilement sur le score de 7-1, faisant ainsi subir à la sélection brésilienne la pire humiliation de son histoire à domicile. Ce projet démontre l'énorme impact que peut avoir la collecte de données à la condition qu'elles soient exploitées de la bonne

manière. Le logiciel a par exemple permis d'améliorer la vitesse de passe de l'équipe. Le temps de possession moyen par joueur est ainsi passé de 3.4 secondes en 2010 à 1,1 en 2015, un élément clé pour développer le jeu rapide et agressif adopté par l'équipe d'Allemagne au cours de cette Coupe du monde. Autre exemple au niveau de l'aspect tactique : le logiciel a permis de faire ressortir le fait que l'équipe de France était trop concentrée au milieu du terrain et laissait des espaces sur les côtés à cause d'un mauvais positionnement des arrières latéraux. Le staff allemand a ainsi composé une équipe construite pour exploiter ces faiblesses et franchir l'obstacle français.

L'aspect visuel et « ludique » de l'application mobile a également séduit les joueurs et l'encadrement. Le manager général de l'équipe Oliver Bierhoff explique:

*« Tous les joueurs reçoivent quelques exemples de ce qu'ils ont bien fait et mal fait juste après le match. Ils peuvent le regarder quand ils le souhaitent et également regarder les différents indicateurs de performance. C'est beaucoup plus utile qu'une cassette vidéo de 90 minutes comme cela se faisait à mon époque. Les joueurs apprécient ce type de feedback. Nous avons également beaucoup de données sur l'adversaire. Jérôme Boateng a par exemple demandé à voir la façon dont Cristiano Ronaldo se déplaçait dans la surface de réparation »<sup>2</sup>*



*Capture d'écran du logiciel utilisé par la sélection allemande*

Si le cas allemand a fait date, notamment en révélant ces pratiques aux yeux d'un public plus large, le Big Data et l'intelligence artificielle sont utilisés depuis de nombreuses années dans le sport professionnel. A divers degrés et via des vêtements connectés, capteurs et divers

<sup>2</sup> Honigstein chats with Oliver Bierhoff, Raphael Honigstein, ESPN, 8 juillet 2014

algorithmes, ces outils contribuent à optimiser les performances athlétiques des sportifs, gagner un avantage tactique ainsi qu'à effectuer un recrutement de qualité à moindre coût.

## **A – Optimisation des performances physiques**

Au plus haut niveau de compétition, le moindre détail physique peut faire la différence. Les staffs techniques sont ainsi de plus en plus large, incluant divers médecins, kinésithérapeutes et préparateurs physiques. Depuis quelques années, la collecte et l'analyse de données y ont également une place de plus en plus grande, et ces derniers sont rejoints par des « data scientists », analystes et autres ingénieurs. En effet, les données sont relativement faciles à collecter, abondantes et peuvent avoir un impact positif très rapidement lorsqu'elles sont bien utilisées.

- **GPS, capteurs et textiles connectés : différents outils pour différents sports**

Les technologies utilisées pour collecter des données sont diverses et variées. Plus ou moins sophistiquées, elles sont adaptées aux besoins et aux budgets des différents sports. GPS pour mesurer la vitesse, bracelets pour le rythme cardiaque, capteurs biométriques dans la chaussure, vêtements connectés et même capsule électronique ingérable permettant le suivi continu de la température centrale : il y en a de tous types et nombre de start-ups se battent au quotidien pour conquérir ce marché.

Les sports où la performance physique pure est primordiale sont particulièrement propices à l'utilisation de ces technologies. L'aviron en est une bonne illustration car il nécessite de trouver un équilibre entre force et endurance. En effet, certains exercices destinés à augmenter l'endurance peuvent diminuer la force et inversement. Analyser une quantité suffisamment importante de données provenant des entraînements passés permet ainsi de déterminer quel est l'impact d'un entraînement précis sur les caractéristiques d'un athlète en particulier. Mark Homer, scientifique de l'équipe d'aviron du Royaume-Uni, cite par exemple la possibilité de déterminer si l'idéal est d'effectuer une ou bien deux sessions de lever de poids pour un athlète avant de passer à un entraînement d'endurance. L'équipe, qui a remporté toutes les médailles d'or aux jeux olympiques depuis 1984, fait figure de précurseur en la matière. Ils ont en effet été les premiers à constater que la majorité des mouvements des rameurs sur l'eau et à la salle de musculation étaient mesurables et qu'ils pouvaient en tirer quelque chose. Sir David Tanner, à la tête l'équipe olympique depuis le milieu des années 1990, explique qu'il a d'abord été nécessaire de choisir les données pertinentes (performances individuelles dans les bateaux et à l'extérieur, poids soulevé à l'entraînement et progrès individuels entre 18 ans et 30 ans) : celles

pouvant réellement impacter les performances en compétition. Les informations médicales, biomécaniques et performances sur l'eau de chaque athlète sont ainsi régulièrement collectées et stockées sur des logiciels et serveurs dédiés. L'équipe a notamment signé un partenariat avec l'entreprise SAP (comme l'équipe de football d'Allemagne) mais aussi avec Siemens. Parmi les bénéfices offerts par ces technologies, la possibilité de comparer le profil des jeunes athlètes à celui des plus expérimentés pour déterminer la meilleure façon d'atteindre des niveaux de performances supérieurs ou au moins similaires. Elles permettent également de diminuer les risques de blessure et d'optimiser les temps de récupération. En effet, l'analyse des données passées permet de voir quand un athlète atteint ses limites et se retrouve en situation dangereuse, celle-ci mettant en lumière les signes alarmants (données médicales, physiologiques, performances en course ou à l'entraînement)<sup>3</sup>

De la même manière, depuis l'autorisation de l'International Rugby Board (IRB) en 2010, les rugbymen portent des capteurs mesurant vitesse, distance parcourue mais aussi énergie dépensée et impact des chocs ; des données que peuvent visionner en temps réel les entraîneurs et préparateurs physiques. Ainsi lors d'un match de l'équipe de France en 2015, le joueur Morgan Parra, a été sorti immédiatement par le staff à la suite d'un violent plaquage après qu'ils étaient alertés par l'impact du coup sur leur ordinateur depuis le banc de touche<sup>4</sup>. On peut également évoquer le cas des Tigers de Leicester qui ont développé un partenariat avec IBM dès 2012 pour adapter leur logiciel d'analyse prédictive à l'anticipation de potentielle blessures de leurs joueurs<sup>5</sup>.

- **Le cas du cyclisme et de l'équipe Sky : une hégémonie mondiale permise par le Big Data**

Si l'idée de prévenir les blessures séduit de nombreuses équipes de différents sports, la recherche de la performance reste l'objectif numéro un de la plupart d'entre elles. L'équipe de cyclisme Sky (devenue depuis Team Ineos), créée en 2010, a ainsi décidé de faire de l'exploitation des données son atout numéro un dans sa quête de domination du cyclisme mondial. Avec, entre 2010 et 2018 : 322 victoires, 52 courses à étapes, 25 courses d'un jour et 9 grands tours (dont 7 des 8 derniers Tour de France), on peut dire que le pari est réussi.

---

<sup>3</sup> Driving success, SAS, 02/2015, p30-31

<sup>4</sup> Rugby : les joueurs du XV de France surveillés par GPS, Cathy Lafon, Sud Ouest, 14/03/2015

<sup>5</sup> Leicester Tigers Rugby Team Deploys Predictive Analytics from IBM to Reduce Injury Number and Severity, Newsroom IBM, 27/04/2012

Son directeur, Dave Brailsford, a cassé les traditions du cyclisme avec le concept des gains marginaux, consistant à prêter attention aux moindres détails. Il affirme ainsi que « si vous isolez tous les éléments auxquels vous pensez et que vous améliorez chacun de 1%, en les additionnant, vous obtenez un gain significatif »<sup>6</sup>. L'entraînement est notamment repensé sous l'impulsion de Tim Kerrison (préparateur physique et autre « cerveau » de l'équipe) : décrassage sur les home-trainers (vélos d'appartement) en fin d'étape, tests en soufflerie et stages en altitudes font tous l'objet d'une collecte de données. Le sommeil des coureurs fait également l'objet de mesure pour optimiser la forme physique de ceux-ci. Enfin, un élément clé de la réussite de l'équipe britannique est la gestion des efforts déployés par les coureurs pendant la course (mesurés en watts générés et visibles par les coureurs en temps réel sur leur vélo) en fonction des données collectées tout au long de l'année.

Rien n'est laissé au hasard et le cas du coureur Chris Froome est l'exemple parfait de l'importance prise par la collecte de données et par les concepts de watts et de « threshold » (seuil, palier) dans le cyclisme moderne. Pendant la course, les coureurs sont amenés par moment à effectuer des efforts plus intenses (pédaler contre le vent, remonter le peloton après être parti chercher les bidons de ravitaillement...) dépassant un certain seuil au-delà duquel il est impossible de rouler sur une longue durée (le « threshold »). Chris Froome, comme d'autres, essaye de dépasser le moins souvent ce fameux seuil, quitte à perdre momentanément sa position dans la course. Il explique ainsi imaginer avoir un sac avec un nombre de pièces déterminé au départ de la course et que chaque dépassement du seuil lui en retire une. Il sait qu'au pied de l'ascension finale d'une course, c'est le coureur à qui il reste le plus de pièces qui a le plus de chance de gagner. C'est là que l'exploitation des données fait la différence : la connaissance et le suivi des données physiques collectées tout au long des années permet une meilleure gestion de l'effort. Ainsi sur la première étape de montagne du Tour de France 2016, Froome a développé une puissance de 414 watts de moyenne avec une pointe à 556 watts sur 24 secondes, clouant sur place son principal rival Nairo Quintana. Des chiffres impressionnants, qui n'ont pas manqué de surprendre le public et les journalistes et de susciter des soupçons de dopage. Pourtant, l'entraîneur de l'équipe n'était pas surpris pour le moins du monde à la fin de la course : il a ainsi déclaré aux journalistes que son coureur avait déjà dépassé ce niveau de puissance à 15 reprises depuis 2011 en course ou à l'entraînement. L'analyse de son rythme

---

<sup>6</sup> *Dix ans de revolution – Cyclisme – Team Sky*, Chaîne YouTube L'équipe, 09/03/2019

cardiaque montrerait également que Froome est arrivé au pied de l'ascension frais et en bonne condition physique, autrement dit qu'il avait conservé la plus grande partie de ses pièces. Ce suivi permanent et ces entraînements personnalisés ont notamment permis de faire passer Chris Froome d'un VO2 Max (taux maximum auquel il peut respirer de l'oxygène) de 80.2 ml d'oxygène / kg en 2007, à 84.6 en 2015, le tout avec 5.7 kilos en moins et un threshold similaire<sup>7</sup>.

Mais cette approche très scientifique du cyclisme, ayant transformé un coureur à haut potentiel en un quadruple vainqueur du Tour de France, ne fait pas que des heureux. Au-delà des soupçons de dopage, c'est la froideur de cette implacable machine à gagner qui énerve au sein du peloton et du public. Pour Bradley Wiggins qui a pourtant profité de ce système en gagnant le Tour en 2012 avec le maillot de la Sky, l'équipe « déshumanisait le cyclisme, retirait l'émotion de la course »<sup>8</sup>. Un débat qui touche à la nature même du sport et qui n'épargne que peu de disciplines. C'est une question qui risque de se poser de plus en plus dans les années à venir, mais nous en reparlerons dans la 3<sup>ème</sup> partie de ce document.

---

<sup>7</sup> The science behind Chris Froome and Team Sky's Tour de France preparations, Joao Medeiros, Wired, 30/06/2016

<sup>8</sup> The Bradley Wiggins Show, podcast du 27/05/2019

## **B – Un outil d’aide à la prise de décisions tactiques**

Nous nous sommes jusqu’à présent concentrés sur l’aspect purement physique mais l’utilisation de ces données est bien plus vaste. Le travail de la technique se voit notamment impacté par la data, mais ces technologies s’avèrent surtout être un outil clé dans la réflexion tactique.

- **Du « long-ball game » au tiki taka : l’influence du Big Data sur les tactiques de jeu dans le football**

L’analyse des données des matchs passés permet de répondre de façon objective à des questions qui dépendent habituellement de la philosophie de jeu de l’entraîneur : a-t-on plus de chances de marquer en jouant un corner « à la rémoise » ou en centrant directement dans la surface ? Suite à un long dégagement ou une relance au pied ? En envoyant le ballon directement dans le camp adverse ou en multipliant les passes ? Les interrogations sont légions. Ces analyses peuvent influencer le style de jeu de tout un championnat, comme nous allons le voir avec l’exemple du « long-ball game » érigé en modèle par un des précurseurs des statiques dans le football avant d’être rendu obsolète par le Big Data.

Charles Reep, un comptable de la Royal Air Force, a commencé en 1950 à retranscrire tous les événements des matches de football auxquels il assistait grâce à un système de notation inventé par ses soins. L’analyse de ces données lui prenait environ 80 heures par rencontre et, au fil des années, il a analysé plus de 2 200 matches de football<sup>9</sup>.

Ces données lui ont permis de démontrer que ce sport, pourtant imprévisible dans sa globalité, comportait des constantes, des patterns prédictibles. Il fait notamment les constats suivants : un but est marqué en moyenne tous les 9 tirs, 80% des buts sont marqués à la suite d’actions de moins de 4 passes et 50% des buts marqués le sont suite à une récupération de balle dans les 30 derniers mètres. Il conclut alors que les équipes doivent passer moins de temps à se faire des passes et plus à envoyer la balle dans le camp adverse, donnant ainsi naissance au « long-ball game », une tactique de jeu qui deviendra la norme du foot anglais pendant plusieurs décennies.

---

<sup>9</sup> How data analytics killed the PL’s long ball game, Joao Medeiros, Wired, 01/2014



En 2005, Ian Franks, professeur à l'université de British Columbia, et Mike Hughes, un mathématicien pionnier de l'analyse de données sportives, ont voulu vérifier les constats de Reep à partir de données provenant du football moderne. Ils se sont notamment rendu compte que si la plupart des buts étaient marqués après des actions de moins de 4 passes, c'est uniquement qu'elles représentent la majeure partie des actions de football et non pas parce que la probabilité de marquer est plus élevée. Chris Anderson, économiste de l'université Cornell (New York) et qui étudie les données footballistiques depuis 2014 estime que Reep était un excellent comptable mais un piètre analyste, obsédé par l'idée de prouver ses idées initiales par les chiffres plutôt que de lire les données collectées avec un regard neutre. Celui-ci a par exemple sciemment ignorer que des équipes comme le Brésil, ne pratiquant pas le « long-ball game », avaient remporté la Coupe du Monde. Dans un livre, co-écrit par David Sally, Anderson écrit ainsi : « La quête de Reep d'utiliser des statistiques pour déterminer une tactique footballistique a échoué car c'était un absolutiste, déterminé à utiliser les données pour prouver ses convictions. Il aurait eu besoin d'abandonner son idée de chercher une règle générale, une recette de la victoire, et apprendre à chercher les multiples vérités et fausses vérités des chiffres en eux-mêmes »<sup>10</sup>.

Grâce à des preuves statistiques, le Big Data a ainsi contribué à la disparition et à l'émergence de différents styles de jeu plus ou moins dominateurs à des périodes données. Mais il n'y a pas une vérité universelle, et les données recueillies doivent être adaptées en fonction du style de jeu de l'équipe et du profil des joueurs qui la composent. Simon Wilson, ancien consultant pour Prozone (société spécialisée dans les statistiques sportives) et devenu depuis « Manager de la performance stratégique » à Manchester City, s'attarde par exemple sur la possession dans les trente derniers mètres ou le taux de réussite des passes vers l'avant : des statistiques positivement corrélée avec le nombre de victoires pour des équipes du calibre de Manchester. « Quand on a étudié le profil des tops clubs contre des équipes moyennes, nous avons observé que les meilleures équipes dominaient cette partie du terrain. Le taux de réussite des passes, notamment vers l'avant, était particulièrement élevé. Donc maintenant, quand nous recrutons des joueurs, nous prêtons particulièrement attention aux joueurs avec un taux de réussite des passes élevé »<sup>11</sup>. Des statistiques intéressantes pour Manchester City mais non pertinentes pour beaucoup d'autres clubs. D'où la nécessité pour chaque équipe de déterminer les statistiques les plus significatives pour elle. Pedro Marques, analyste du même club, résume ainsi : « Au

---

<sup>10</sup> The numbers game, Peguin, 2013

<sup>11</sup> How data analytics killed the PL's long ball game, Joao Medeiros, Wired, 01/2014

lieu d'étudier une liste de 50 variables, nous voulons en trouver 5 qui correspondent à notre type de jeu »<sup>11</sup>. Wilson évoque notamment une période durant laquelle l'équipe n'avait marqué aucun but sur corner en 22 matches. Après l'analyse de plus de 400 buts marqués sur corner, ils se sont aperçus avec son équipe que 75% d'entre eux étaient marqués suite à corner rentrant. Modifiant leur stratégie pour s'adapter à ce constat, ils marqueront 9 buts sur corners au cours des 12 rencontres suivantes.

Les exemples d'adaptation de jeu s'appuyant sur des constats faits à partir d'analyse Big Data qui ont porté leur fruits sont nombreux. Prenons ainsi le cas de l'Equipe de France, championne du monde en 2018 et plus précisément le cas d'un débriefing vidéo suite au premier match décevant contre l'Australie. Didier Deschamps, explique à ses joueurs que l'équipe adverse a parcouru 111km contre 102km pour l'équipe de France : « comme s'ils avaient un joueur de plus ». Il continue en leur disant que 5 joueurs australiens ont réalisés plus de 3 sprints à haute fréquence, contre seulement 1 français. Il se tourne alors vers Kylian Mbappé pour lui dire qu'il n'en a fait aucun... ce dernier marquera un but suite à un sprint à haute fréquence lors de la rencontre suivante, avec la fin de tournoi que l'on connaît<sup>12</sup>.

- **L'analyse de données en temps réel : une révolution tactique à venir**

Si les staffs techniques passent de nombreuses heures à décortiquer les données des rencontres et entraînements passés, les nouvelles technologies permettent de fournir de plus en plus de données en temps réel permettant des changements tactiques au cours même de la rencontre.

L'entraîneur peut par exemple constater l'épuisement d'un joueur, via la diminution du nombre de mètres parcourus sur le terrain et du nombre de courses à haute intensité ou de ses pulsations cardiaques (données fournies via des vêtements connectés et autres systèmes de tracking GPS). Le nombre et le pourcentage de passes, tirs ou interventions défensives réussies peuvent également confirmer la bonne ou mauvaise impression visuelle laissée par un joueur et, dans le deuxième cas, provoquer un changement de joueur.

Le positionnement des joueurs et leur zone de jeu présentée de façon graphique permet notamment de vérifier le respect des consignes tactiques et entraîner un rappel de celles-ci si elles ne sont pas appliquées correctement. En plus de cette analyse interne, propre à l'équipe,

---

<sup>12</sup> Les Bleus 2018, au cœur de l'épopée Russe, TF1, 17/07/2018

les données fournies peuvent faire apparaître des faiblesses chez l'adversaire qui n'avaient pas été identifiées au préalable. Une modification tactique ou un remplacement peut alors être décidée afin « d'appuyer là où ça fait mal » pour mener l'équipe vers la victoire. Enfin, une analyse du type de passes éliminant le plus de joueurs ou des zones de tirs présentant le taux de réussite le plus élevé peut conduire à une modification de la stratégie sur le terrain.

L'analyse de données en temps réel représente donc déjà le présent du sport professionnel mais, selon les analystes, elle n'en ait qu'à ses balbutiements et de nombreuses perspectives vont s'ouvrir à l'avenir. Premièrement, les technologies de collecte de données en temps réel n'en sont encore qu'à leur début et sont toujours perfectibles. Nécessitant d'importants investissements, elles sont aujourd'hui réservées à un nombre de sports restreints (football, football américain, rugby, baseball...) et aux ligues les plus prestigieuses. Deuxièmement, si la collecte des données est automatique et que sa retranscription en informations claires et lisibles par le staff est assurée par l'intelligence artificielle, l'analyse de celles-ci et la prise de décisions qui en découle est encore aujourd'hui assurés par des êtres humains. On pourrait imaginer qu'à l'avenir, grâce aux progrès de l'IA et du machine learning, des remplacements de joueurs ou des ajustements tactiques soient suggérés directement par la machine.

En attendant de voir ce que nous réserve l'avenir, il est un domaine où ces technologies sont déjà abouties depuis plusieurs années comme nous allons le voir dans la partie suivante : le recrutement.

## C – Détection des talents de demain à moindre coût : le scouting 2.0

- **Oakland Athletics et « sabermétrie » : précurseurs du recrutement basé sur les statistiques**

En 2003, l'auteur Michael Lewis publie le livre « Moneyball: The Art of Winning an Unfair Game » (qui fera l'objet d'une adaptation cinématographique quelques années plus tard), changeant à jamais les méthodes de recrutement en vigueur dans le baseball américain. Le livre familiarise le grand public avec le concept de « sabermétrie » à travers le récit de l'histoire des Athletics d'Oakland et de leur manager Billy Beane au cours de la saison de 2002.

Confronté à un budget limité en comparaison des autres équipes de la ligue (41 millions de dollars de masse salariale contre des sommes s'élevant jusqu'à 125 millions pour les grosses équipes de la ligue telles que les Yankees de New York), la franchise se voit contrainte de trouver des solutions pour identifier des joueurs sous-évalués sur le marché. Le staff se porte alors sur l'approche sabermétrique dont les premiers principes ont été définis par Earnshaw Cook dans son livre « Percentage Baseball » en 1964, bien avant que le terme lui-même ne soit inventé. L'essai de Cook, à travers une analyse mathématique, remet en question beaucoup de stratégies adoptées dans le monde du baseball professionnel mais l'approche scientifique et le vocabulaire jugé trop académique entraîneront un rejet global des acteurs du baseball.

Le terme de saberm<sup>13</sup>étrie ne fera son apparition qu'en 1980 à travers la plume de Bill James qui le définira comme « la recherche de la connaissance objective sur le baseball ». Elle tente de répondre à des questions objectives sur le baseball, comme "Quel joueur chez les Red Sox a le plus contribué à l'offensive de l'équipe ?" ou "Combien de circuits frappera Ken Griffey l'an prochain ?". Elle ne peut considérer les jugements, importants mais subjectifs, qui font partie du jeu, comme "Quel est ton joueur favori ?" ou "C'était un excellent match !". La sabermétrie a ainsi pour objectif de définir la valeur présente ou future d'un joueur ou d'une équipe en s'appuyant sur les statistique passées et présentes et non pas en se reposant sur des intuitions et autres jugements de valeurs.

---

<sup>13</sup> The sabermetric Manifesto, David Grabiner, 1994

Dans le cas précis des Athletics d'Oakland, le staff a jugé que la vision contemporaine des professionnels du baseball était subjective et faillible, et que les statistiques les plus couramment utilisées (buts volés, points produits, moyenne au bâton...) étaient inadaptées pour juger de la qualité et du potentiel des joueurs. Des études statistiques plus larges leur permettront de définir d'autres indicateurs plus pertinents (moyenne de présence sur les buts et de puissance entre autres), reflétant des qualités traditionnellement moins recherchées par les recruteurs malgré un impact statistiquement prouvé. Ils pourront ainsi recruter des joueurs performants ou à grand potentiel à moindre coût. Autre constat issu de ces analyses empiriques, un joueur évoluant en université a de plus grandes chances d'évoluer en MLB (Major League Baseball, plus haut niveau dans la hiérarchie du baseball) qu'un joueur repêché à l'école secondaire. Le risque serait ainsi plus élevé lors du recrutement d'un joueur de l'école secondaire que d'un joueur universitaire, peu importe les évaluations faites par les observateurs dédiés.

A la surprise générale, la franchise finira première de la conférence Ouest. Devant ce succès probant, de nombreuses équipes de la NFL prirent exemple sur ce modèle nouveau en embauchant des spécialistes de la sabermétrie, le terme Moneyball étant même rentré dans le jargon du baseball pour qualifier les équipes plébiscitant ce système pour leur recrutement. Aujourd'hui, la pratique a gagné de nombreux autres sports et les technologies modernes permettent de réaliser des économies dans le recrutement en repérant les pépites de demains partout sur la planète. Le football actuel en est la parfaite illustration.

- **Football Manager : un jeu vidéo devenu outil de travail pour les recruteurs**

Novembre 2008 : le club anglais Everton FC signe un accord avec l'éditeur Sports Interactive pour obtenir un accès privilégié à la base de données de son jeu Football Manager 2009 afin de faciliter le repérage de nouveaux talents<sup>14</sup>. Si c'est la première fois qu'un club annonce officiellement se servir de cet outil, le jeu tourne depuis quelques années déjà sur les ordinateurs de nombreux recruteurs à travers le monde.

---

<sup>14</sup> Everton FC : Football Manager 09 aide aux transferts, Guillaume Nerces Tutundjian, 17/11/2008

Le chef des scouts de Watford, Filippo Giraldi déclare ainsi « Être fort à Football Manager n'est pas nécessaire pour être un bon scout mais ça aide pas mal. Le jeu vous donne beaucoup de données utiles. C'est important de connaître les joueurs »<sup>15</sup>.

En France, le premier à avoir brisé ce tabou est un jeune recruteur de l'OGC Nice, Jonathan Beilin. En 2011, via une sortie dans le journal France Soir, il reconnaît avoir recouru à Football Manager pour recruter le défenseur serbe Nemanja Pejčinović. Il explique : « En fait, en période de mercato estival, tous les jours, on te donne dix ou quinze noms. Un jour, dans une de ces listes, mon supérieur m'a parlé de Pejčinović. Et ce joueur-là, je le connaissais déjà depuis deux ou trois ans, de Football Manager. Direct, quand on m'a proposé ce nom, j'ai dit : quand on va regarder des vidéos, on sera tous d'accord. J'ai tilté tout de suite »<sup>16</sup>.

Le jeu prenait ainsi une place énorme dans son travail quotidien : « Je m'enfermais tout un week-end ou le soir après le bureau. Je jouais entre 30 et 50 heures par semaine. T'as déjà joué à Football Manager, tu vois comme c'est prenant, ça prend un temps fou. Je connais le jeu par cœur, j'y joue depuis 1996. Il est très proche de la vie réelle. En gros, quand j'arrivais au bureau, j'avais trois écrans. Sur le premier, j'allumais Football Manager. Sur le second, une page internet avec les outils de statistiques du club. Un enfin, un écran relié à un lecteur DVD pour visionner des vidéos. Mais oui, FM était toujours allumé »<sup>16</sup>. Un discours qui a failli lui coûter son poste, les dirigeants de l'époque estimant qu'utiliser un jeu vidéo comme outil de recrutement était dénigrant pour un club professionnel.

Aujourd'hui, les choses ont bien changé et la plupart des clubs assument pleinement le recours à cette base de données. Sa légitimité sera définitivement assurée en 2014, quand Prozone signe un partenariat avec le développeur du jeu pour incorporer la base de données à leurs propres outils. Fournisseurs de statistiques pour la LFP, la Serie A, le TOP 14 ainsi que des clubs comme le PSG, Chelsea ou la Juventus, la société rend depuis disponible la base de données de Football Manager aux acteurs les plus prestigieux du monde du sport, de façon totalement officielle cette fois-ci.

---

<sup>15</sup> Comment football manger est devenu un vrai outil pour les clubs, Alix Dulac, RMC Sport, 28/11/2015

<sup>16</sup> Jonathan Beilin : « Football Manager était toujours allumé », Kévin Charnay, So Foot, 12/11/2015

Alors qu'est-ce qui rend le produit de Sports Interactive aussi attractif aux yeux de ses derniers ? C'est d'abord l'impressionnante base de données du jeu, qui comprend des informations extrêmement détaillées sur 799,643 joueurs et entraîneurs à travers le globe. Pour collecter toutes ces informations, Mark Woodger, en charge de la base de données dispose d'une équipe de 8 personnes qui supervisent le travail d'un personnel bien plus large. Une centaine de « Head Researchers » tout d'abord, qui sont chargés du suivi d'un ou deux pays/compétitions. Un cran en dessous, on retrouve des « Assisant Researchers » qui supervisent un club en particulier. Ces derniers, qui sont bénévoles, assistent aux matches du club en question, aussi bien de l'équipe une que des équipes de jeunes et équipes réserves. Ils scrutent également quotidiennement la presse à la recherche de toute information concernant les joueurs qu'ils ont vu joué. Ils vont alors remplir une note sur 20 pour chacun d'entre eux sur différents critères. Toutes les informations sont alors vérifiées et corrigées par leurs supérieurs hiérarchiques ; l'étendue du réseau et la diversité des personnes impliquées assurant la qualité des données collectées.<sup>17</sup>

Football Manager permet ainsi d'accéder à des informations sur n'importe quel joueur dans le monde facilitant le travail des recruteurs mais ce n'est là l'unique clé de son succès. En effet en plus de cette très grande quantité de données disponibles, c'est la capacité des algorithmes du jeu à évaluer de façon exacte le potentiel d'un joueur et son évolution à travers le temps qui le rend si unique. Des joueurs comme Karim Benzema, Javier Saviola, Zlatan Ibrahimovic, Kylian Mbappe, Matthijs De Ligt ou Frankie De Jong étaient tous annoncés comme de futurs cracks dans le jeu bien avant d'être révélés aux yeux du grand public. Ces dernières années ayant vu le prix des jeunes pépites s'envoler comme pour Antony Martial (70 millions à 19 ans en 2015), Neymar (89 millions à 21 ans en 2013), Ousmane Dembélé (105 millions à 20 ans) ou encore Kylian M'bappé (145 millions à 19 ans)<sup>18</sup>, il est devenu crucial pour les clubs d'être les premiers à les repérer. Si le jeu a également connu quelques échecs dans ses prédictions, le pari peut rapporter gros pour les recruteurs.

Enfin, comme pour le cas de la plateforme Match Insights développée par SAP, il ne faut pas négliger la qualité de l'interface et l'optimisation de l'expérience utilisateur comme facteur du succès de Football Manager. Imaginons que votre club est à la recherche d'un jeune ailier, rapide, qui sait dribbler et ce avec un budget limité : il suffit alors d'entrer ces critères et le jeu

---

<sup>17</sup> How Football Manager's scouting network unearths the wonderkids of real world football, JJ Bull, The Telegraph, 02/11/2018

<sup>18</sup> Sources chiffrées : Transfermarket



vous propose immédiatement une liste de joueurs correspondant au profil recherché avec une estimation de leur valeur marchande et une compilation de diverses informations les concernant. Facile d'utilisation, permettant un gain de temps important, Football Manager n'a pas remplacé les méthodes de scouting et recrutement traditionnelles mais il s'est définitivement imposé comme un outil complémentaire, y compris dans les clubs les plus prestigieux. Présentant un marché lucratif, de nombreuses sociétés se sont engouffrées dans la brèche.

- **Wyscout : numéro 1° des plateformes de scouting**

Le leader mondial en la matière se nomme « Wyscout » et a été cofondé en 2004 par trois amis, Simone Falzetti, Matteo Campodonico et Piermaria Saltamacchia. L'aventure commence avec une simple caméra avec laquelle ils vont filmer les rencontres des championnats amateurs de la région, puis les envoyer aux entraîneurs et staffs de divisions supérieures pour leur permettre de repérer des talents à même de renforcer leur équipe. Aujourd'hui, la société a bien évolué et les chiffres donnent le tournis : 599 compétitions filmées dans 124 pays à travers le monde (dont respectivement 214 et 52 pour lesquelles des statistiques détaillées sont fournies). La société compte parmi ses clients le Real Madrid, l'Olympique Lyonnais, le FC Porto, l'AS Roma, les Tottenham Hotspurs et bien d'autres clubs unanimement reconnus pour leur capacité à dénicher pépites de demain.<sup>19</sup>

Les compétitions filmées sont ensuite traduites en données (passes décisives, buts, tacles, centres, pertes de balle...) par des équipes situées en Bulgarie, au Sénégal et en Moldavie. Les données sont ensuite rentrées à la main par les équipes situées au siège en Italie après vérification. Avec 1,800 nouveaux matches analysés par semaine en moyenne, la base de données ne cesse de s'élargir.<sup>20</sup>

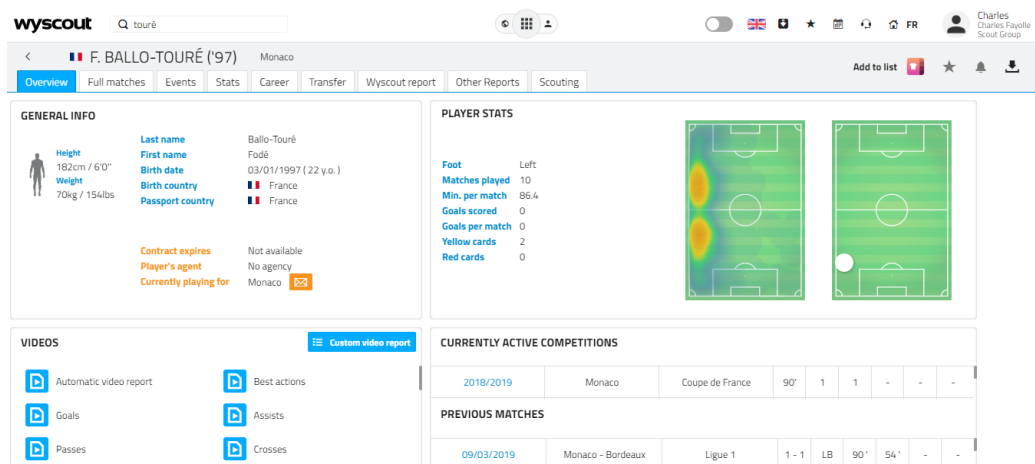
Alors comment se présente le logiciel pour les utilisateurs (recruteurs donc, mais aussi entraîneurs, agents, joueurs, arbitres, journalistes ou encore agences de paris sportifs) ?

À chaque club et sélection sont reliées les vidéos des matchs entiers (existent aussi les versions temps de jeu effectif), plus de 150 000 sont ainsi disponibles en ligne. Plus la formation concernée évolue à un niveau élevé, plus ses actions sont décortiquées : buts, occasions, contre-attaques, attaques sur touches, style défensif, penaltys concédés... Cette fonctionnalité est ainsi un outil indispensable pour les staffs techniques afin de passer au crible un prochain adversaire. L'autre aspect, probablement le plus important, concerne les individualités. Là encore, pour préparer le prochain match, mais surtout pour effectuer son recrutement. Si l'onglet carrière est assez classique, celui générique permet de comprendre en un coup d'œil quelles sont les principales qualités de l'élément sélectionné grâce à différents graphiques.

---

<sup>19</sup> On est allé chez Wyscout, leader mondial du scouting, Valentin Pauluzzi, So Foot, 28/07/2016

<sup>20</sup> On a testé Wyscout, Valentin Pauluzzi, So Foot, 28/07/2016



*Capture d'écran du logiciel Wyscout*

- **Economies de temps, d'argent et plus grande transparence**

Alors concrètement quels sont les apports de ces nouvelles technologies pour le travail des recruteurs ? Premièrement, un gain de temps et d'argent. Fini les déplacements pour aller un voir un joueur : en quelques clics, ceux-ci ont accès à des statistiques et des vidéos permettant d'éviter les trajets sur place longs et coûteux autrefois indispensables. Directeur général du FC Sion, Marco Degennaro déclare ainsi: « Nous utilisons beaucoup cet outil. Avant, se faire une première opinion sur un joueur nécessitait un déplacement. Là, en ligne, on trouve tous les matches. On économise de l'argent, et du temps. »<sup>21</sup> Un élément crucial pour les clubs ayant un budget modeste et qui n'auraient pu auparavant recruter des joueurs évoluant à l'autre bout du monde, faute de finances et de personnel pour superviser des compétitions sur d'autres continents.

Les clubs sont également moins susceptibles de faire face à des mauvaises surprises. En effet, pendant longtemps les agents de joueurs envoyaient aux clubs courtisés des vidéos de leurs joueurs avec des extraits sélectionnés pour mettre en valeur ces derniers. Il n'était ainsi pas rare de voir des joueurs d'Afrique ou d'Amérique latine apparaissant comme de futurs cracks sur les compilations mais difficilement reconnaissables une fois arrivés sur place. Avec les outils à leur disposition aujourd'hui, les recruteurs peuvent voir ce qu'ils veulent d'un joueur et non pas uniquement ce que l'agent souhaite leur montrer. Il est ainsi beaucoup plus difficile pour les agents de bluffer sur les réelles capacités de leurs joueurs.

<sup>21</sup> Sur le mercato, la révolution data, Lionel Pittet, Le Temps, 21/06/2017

- **Recruter des profils adaptés à un style de jeu spécifique**

Enfin, l'accès à des statistiques détaillées sur chaque joueur permet d'identifier ceux qui peuvent potentiellement correspondre au style de jeu pratiqué par l'équipe, un élément crucial pour les clubs à l'identité de jeu très prononcée. L'exemple de l'arrivée de James Tarkowski au Burnley Football Club en Février 2006 illustre bien l'importance de l'analyse de ces données. Jeune défenseur central évoluant au Brentford FC, le joueur présentait les caractéristiques suivantes : élégant, solide et discipliné, capable de longues ouvertures et de porter le ballon vers l'avant. Un profil intéressant sur le papier mais bien loin de correspondre au jeu défensif pratiqué par l'équipe de Burnley au premier abord. Le club est effectivement connu pour sa « nonsense approach » de la défense (tactique consistant pour la ligne arrière à relancer via de longs ballons aux centres du terrain et à inciter les équipes adverses à tirer de loin via deux lignes de quatre compactes, limitant les espaces à l'intérieur et aux alentours de la surface de réparation). Un système qui nécessite discipline, intelligence et grandes capacités dans le jeu aérien de la part des défenseurs centraux. Des critères remplis par les titulaires de l'époque Michael Keane et Ben Mee mais bien éloignés en apparence du style de jeu Tarkowski.

Pourtant, une analyse plus approfondie par le club montrera qu'il dispose du potentiel pour adopter le style de jeu de Burnley. En effet, le joueur commet peu d'erreurs défensives, reçoit peu de cartons et présente un pourcentage élevé de duels aériens gagnés. Le staff a également constaté une longueur moyenne de passes plus élevée que les deux joueurs situés plus haut, potentiellement utile pour participer au jeu d'attaque directe de l'équipe. En allant ainsi au-delà des apparences et en privilégiant une analyse statistique poussée (statistiques qui auraient été difficilement accessibles auparavant) plutôt qu'un jugement intuitif, le club s'est offert une recrue de choix qui s'est parfaitement intégrée par la suite.

Avant de conclure cette première partie sur l'amélioration des performances sportives grâce au Big Data et l'IA, attardons-nous sur le cas de la NBA qui reprend l'ensemble des aspects évoqués auparavant (performances physiques, choix tactiques et recrutement) aboutissant à une transformation en profondeur du basketball américain.

- **NBA, Big Data et IA: une transformation radicale du jeu**

Les américains étant des grands consommateurs de statistiques sportives, leurs sports favoris sont souvent précurseurs en la matière, comme nous l'avons vu avec le cas de la sabermétrie

dans le baseball. La NBA, et plus particulièrement les Houston Rockets, en sont une autre excellente illustration.

Sous la houlette de Daryl Morey, manager général de l'équipe et expert en statistique, le club est passé en quelques années du statut d'équipe de milieu de tableau à celui de sérieux prétendant au titre NBA en bonne partie grâce au Big Data.

Première étape : la création par Daryl d'une nouvelle statistique mesurant l'efficacité au tir d'un joueur, appelée « true shooting percentage ». Celle-ci est calculée ainsi :

$$TS\% = \frac{PTS}{2(FGA + (0.44 \times FTA))} \times 100$$

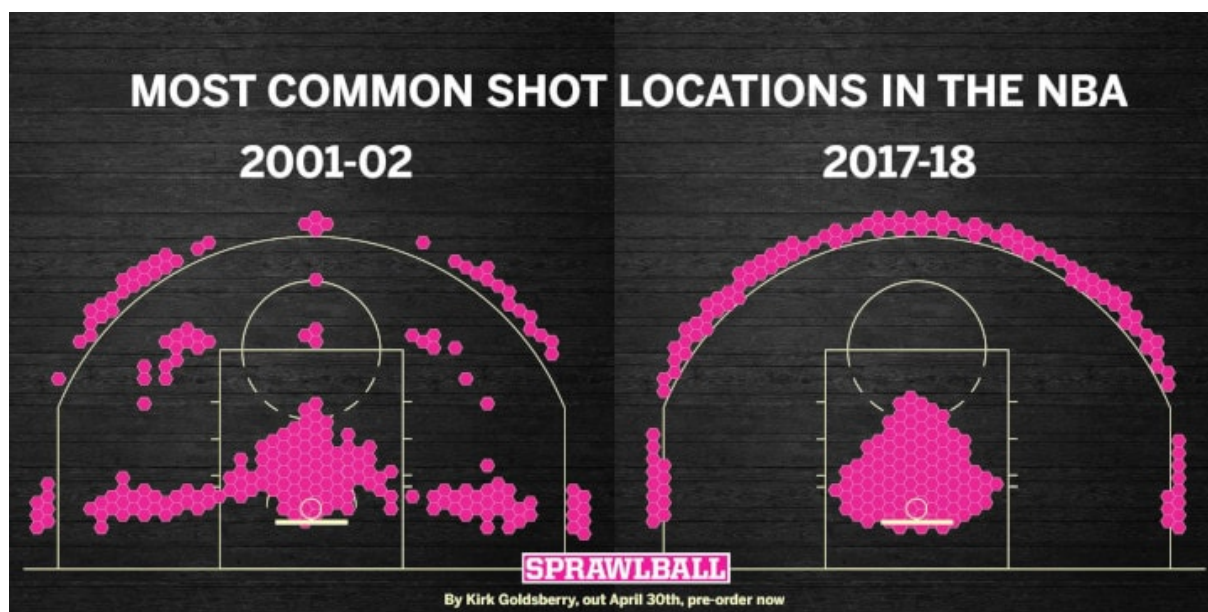
Avec:

- PTS = nombre de points marqués
- FGA = tentative de tirs en jeu
- FTA = tentative de lancer franc

Deuxième étape : l'utilisation dès 2011 d'un système de tracking 3D fourni par l'entreprise SportVU (adoptée également par 3 autres équipes de la ligue cette année-là), permettant de fournir des statistiques approfondies à l'aide des caméras placées au-dessus du terrain.

L'analyse de ces statistiques approfondies va alors révéler les situations de tirs les plus efficaces. Les dunks (88,6% de succès sur la saison 2017-2018 pour l'ensemble des équipes) et les tirs à 3 points derrière la ligne (36,7% de succès pour les Houston Rockets sur 2017-2018) sont les plus rentables, contrairement aux tirs à 2 points à longue distance, extrêmement communs dans les années 1990-2000 qui présentent le pire ratio.

Les entraînements vont ainsi être modifiés pour travailler davantage ces tirs-là, de même que la tactique en match qui va être radicalement transformée. Les rockets vont pousser cette logique à l'extrême avec lors de la saison 2017-2018, 1 184 paniers à 3 points, plus haut total de toute l'histoire de la ligue. Toutes les équipes NBA ayant progressivement adopté ces technologies, c'est le basketball américain entier qui va changer de visage comme l'illustre l'image ci-dessous.



*Source : Twitter de Kirk Goldsberry, NBA analyst et auteur de Sprawlball*

Depuis 2017, la compagnie Second Spectrum a repris le flambeau pour l'ensemble des équipes NBA. Leurs caméras sont ainsi capables de traquer et enregistrer en 3D les mouvements de la balle et de chacun des joueurs à 25 images par seconde. Mais la principale différence avec ses concurrents se situe dans l'utilisation du machine learning. Le logiciel n'était capable de reconnaître que 20 types d'actions (rebond, tir, passe...) à sa création, il en est aujourd'hui à plus de 500 grâce à l'intelligence artificielle.

Le professeur Rajiv Maheswaran, co-fondateur et CEO de la société explique ainsi les possibilités offertes : « Ce que nous pouvons faire est de demander à la machine de regarder tous les tirs similaires à celui-ci et de ressortir la probabilité de marquer soit pour un joueur en général ou pour un joueur en particulier. Et cela prend en compte tout ce qu'une personne peut voir, non seulement la position du joueur sur le terrain, mais aussi son mouvement à ce moment-là, le type de tir qu'il veut tenter et la position des défenseurs. Et ce que nous pouvons faire c'est mettre un chiffre sur quelque chose qui relevait jusque-là de l'ordre de l'intuition, du jugement qualitatif »<sup>22</sup>

Le recrutement des joueurs a également changé avec l'arrivée de ces nouvelles technologies. Les équipes privilégient désormais l'habileté au gabarit. Résultat, les joueurs NBA sont désormais en moyenne plus mince (99kg en moyenne en 2017 contre 101 en 2013) et plus agile qu'auparavant.

Et le professeur Maheswaran de conclure sur les changements à venir grâce l'exploitation des données en temps réel, une perspective que nous évoquions précédemment : « Les données vont de plus en plus vites, auparavant ça prenait des heures, on est descendus au niveau des minutes et extrêmement vite cela descendra jusqu'aux secondes. La technologie va devenir fonctionnelle en temps réel très très bientôt. Je pense que nous n'avons qu'effleuré la surface de ce qui est possible avec les données ». <sup>22</sup>

---

<sup>22</sup> How data transormed the NBA, Chaîne Youtube The Economist, 04/12/2018

## CONCLUSION PARTIE 1 : United Managers, un aperçu du futur ?

La disponibilité des données en temps réel constitue donc un outil de décision majeur pour le staff au cours de la rencontre, mais certains ont vu dans ces outils un moyen de repenser le modèle de coaching et de bouleverser la vision traditionnelle du club. Frédéric Gauquelin, un ex-professeur d'EPS a ainsi lancé en mars 2018, le projet « United Managers » en partenariat avec le club de l'avant-garde caennaise, pensionnaire de Régionale 1 (plus haut échelon du football régional et équivalent de la 6<sup>ème</sup> division). Le concept ? Les décisions tactiques ne sont plus prises par un coach mais par des internautes supporters, désignés sous le nom de « Umans ». Grâce à un partenariat avec la société de statistiques Opta, acteur majeur du secteur, des statistiques précises sur chaque joueur, des notes du staff, des vidéos des entraînements et des résumés des rencontres sont fournis aux adhérents. En se basant sur ces ressources, les « Umans » peuvent ainsi voter pour la composition de l'équipe, le schéma tactique ou encore le style de jeu, le tout directement depuis leur ordinateur, tablette ou smartphone. Les matches sont retransmis en direct sur l'application et le site internet avec tout un panel de statistiques en temps réel, permettant aux internautes de décider d'éventuels changements tactiques au cours du match (remplacement de joueur, changement de schéma tactique, ajustement du positionnement des joueurs...).



Capture d'écran de la plateforme United Managers

Le pari est réussi sur le plan sportif, le club a en effet terminé en première position de sa poule et a été promu en National 3 pour la saison 2019/2020. Pour son fondateur, ce modèle « c'est



l'avenir, c'est pour ça qu'on est là »<sup>23</sup>. Toutefois, le modèle du club est remis en question par la contestation de ses adversaires, soutenus par les instances du football français. Neuf clubs de la région se sont ainsi regroupés pour déposer une première plainte auprès de la Ligue de Normandie afin « de préserver les valeurs d'équité et d'éthique dont la Ligue se doit d'être garant ». Ces derniers estiment en effet que le projet United Managers pose problème sur 3 points principaux : la gestion du club par une société extérieure (ils estiment que les joueurs sont rémunérés par United Managers et non par le club directement), le statut de l'entraîneur (celui-ci n'ayant plus le dernier mot sur les choix tactiques) et les droits télévisés des retransmissions (la diffusion des matches en direct bénéficierait ainsi uniquement à United Managers et non aux autres clubs du championnat).

Au cours d'une assemblée fédérale le 8 décembre 2018, la Fédération Française de Football a donc tranché en défaveur du club caennais sur chacun des points soulevés. Modifiant ses règlements, elle affirme ainsi : "aucun club ne peut signer de contrat ni établir de partenariat permettant à un tiers d'acquérir, d'une quelconque manière, la capacité d'influer sur l'indépendance ou la politique d'un club ou encore sur les performances de ses équipes » ; « aucune tierce partie(...) ne peut remettre en cause la responsabilité réelle de l'équipe détenue par l'entraîneur principal, ou l'effectivité de sa fonction, en tentant d'une quelconque manière d'imposer ou d'influencer ses choix en matière de gestion sportive » et « aucune forme d'exploitation audiovisuelle des rencontres de compétitions officielles ne peut s'effectuer sans le consentement préalable et exprès de la Ligue ou du District concerné(e) ».<sup>24</sup>

Résultat, le club a fini la saison avec un modèle de management classique : les matches ne sont plus accessibles en direct mais uniquement en rediffusion et le coach Julien Le Pen est redevenu le seul et unique décideur en matière tactique, les internautes ne pouvant plus que pronostiquer sur la composition de l'équipe. Des mesures nécessaires, sous peine de s'exposer à des défaites sur tapis vert : « par ces mesures, nous contraignons nos adversaires à devoir nous battre à la loyale, sur le terrain, et non pas via de grossiers stratagèmes juridiques »<sup>25</sup> estime ainsi le club. L'espoir reste toutefois permis pour le projet United Managers, un recours juridique ayant été déposé par la structure en espérant une décision favorable dans les mois à venir.

---

<sup>23</sup> United Managers : les supporters aux commandes, CFC, Octobre 2018

<sup>24</sup> Modifications aux textes fédéraux adoptées lors de l'Assemblée Fédérale du 8 décembre 2018, FFF, 10/12/2018

L'avenir nous dira si le modèle United Managers aura gain de cause et si ce mode de fonctionnement peut se généraliser dans le futur cela souligne en tout cas un point essentiel : le sport professionnel moderne est une histoire de performances physiques mais aussi un spectacle et un business. Les spectateurs et supporters ont un poids croissant et veulent avoir d'avantage d'influence sur les prises de décisions de leur club favori. Les structures sportives se doivent de leur proposer une expérience nouvelle, plus poussée, plus personnalisée. Le tout en essayant d'en profiter pour augmenter la fidélisation de ses derniers et maximiser leurs revenus.

## **PARTIE 2 : Un bouleversement dans et autour l'arène**

L'impact des nouvelles technologies ne s'arrête pas aux performances des joueurs et à l'aspect purement physique et tactique. Au-delà de cela, d'autres acteurs voient leur expérience changer. Les supporters sont en train d'observer et de vivre un changement drastique dans leur manière d'assister à une rencontre sportive. Une expérience qui se veut plus fluide, plus riche et toujours plus intense.

Un phénomène qui profite bien évidemment aussi aux clubs tant ils parviennent à générer plus de revenus grâce à la collecte de données, aux nouveaux points de contacts et à l'optimisation de la plupart de leur processus dans et aux abords du stade. La collecte et l'exploitation de données est plus que bénéfique. La création d'un profil digital de chaque supporter à partir des données permettra de répondre à plusieurs problématiques : comment personnaliser l'approche d'un public de plus en plus divers ? Comment remplir les stades en période creuse et ainsi augmenter le taux de remplissage moyen ? Comment déterminer le bon prix d'un billet à un moment précis ? Toutes ces problématiques trouvent leur réponse dans le Big Data.

Autres acteurs majeurs, les sponsors et partenaires ne sont pas à négliger. La notion de retour sur investissement dans ce domaine est souvent floue : Comment chiffrer l'impact de la présence d'un logo d'une marque sur un maillot de foot ? Comment savoir le bénéfice rapporté par une bannière publicitaire dans le stade ? Grâce à la collecte de donnée et aux profils créés par les clubs, les sponsors et partenaires pourront savoir plus précisément quels types de supporters verront leur contenu, quand et de quelle manière. Pour aller plus loin, ils pourront modifier le contenu de leur pub en fonction de leur position dans le stade et à quelle partie du stade elles seront adressées.

## **A – Enrichir l'expérience du spectateur au stade**

Aujourd'hui la rencontre sportive en elle-même ne suffit plus aux consommateurs. A l'instar de ce qui passe aux États-Unis, précurseurs en la matière, les matchs deviennent de véritables spectacles. Le centre reste le match mais désormais, il faut aller plus loin et plus longtemps. Puisqu'assister à un match représente un symbole fort de rassemblement pour la même passion, il faut en faire un moment inoubliable. Même si la qualité de la rencontre n'est pas au rendez-vous, l'expérience se doit d'être marquante.

Les besoins des différents acteurs ont ainsi évolué parallèlement. D'un côté, les spectateurs bénéficiant des avancées technologiques réclament toujours plus. De l'autre, les clubs doivent faire face à une concurrence féroce et donc augmenter l'engagement et la fidélisation de leurs fans. Comme l'a bien dit le directeur général adjoint du PSG Frédéric Longuépée : *« Si les supporters font le choix de se rendre au stade, il faut que leur expérience soit au moins aussi agréable que s'ils regardaient le match depuis leur salon »*.<sup>25</sup>

- **Quelques contestations qui ne résisteront pas au changement**

L'installation du Wi-Fi au sein des stades a été le premier requis à la mise en place de cette nouvelle approche. Comme tout changement et innovation technologique, notamment dans le monde du sport où il y a la résistance au changement est monnaie commune, cette nouveauté a dû affronter bien des réticences. Le club du PSV Eindhoven aux Pays-Bas ou encore le Borussia Dortmund en Allemagne ont dû essuyer des contestations de leur kop (groupe de supporters le plus impliqué dans le support du club, dans les tribunes ou en dehors). Face à ces vents contraires, le club du Borussia Dortmund avait à l'époque décidé de réduire la limite du nombre de connexions simultanées. Des contestations qui se sont éteintes au fil du temps, laissant le champ libre aux nouvelles opportunités offertes par cet outil.

- **L'exemple du Stade Toulousain, club de rugby à la pointe des dernières technologies**

Un des exemples forts d'innovation est le Stade Toulousain. Club référence en France et dans le monde entier, il est devenu le premier à posséder un stade entièrement connecté. Concernant l'aspect sportif, les spectateurs ont accès aux ralentis et statistiques depuis une application smartphone pour mieux appréhender les actions clés et avoir un regard plus complet sur le déroulé du match. Ceci ayant pour objectif d'augmenter l'engagement des fans en leur permettant

---

<sup>25</sup> Stade connecté, une expérience ultime des fans, Cabestan, emarketing.fr, 27/06/2018

de vivre une expérience sportive globale sans avoir à consulter d'autres applications ou sites internet durant le match. Pour leur tranquillité, ils peuvent commander à manger et à boire depuis leur siège et se faire livrer lorsqu'ils le désirent. Cette possibilité aidera aussi à convaincre tous les supporters qui se diraient « *Bon, il y a trop de monde et je n'ai pas le temps de faire la queue pendant la mi-temps. Je vais me passer du sandwich merguez !* ». Grâce à ces nouvelles possibilités, le confort du spectateur est considérablement amélioré. Le club lui-même en bénéficie aussi. En effet, cela permet de désengorger les flux de spectateurs dans le stade et réduire les habituels goulots d'étranglement. Moins de queue aux buvettes et points de restauration, plus de consommation et un panier moyen plus élevé. En conséquence, le club bénéficie de revenus plus important et plus diversifiés.

- **Les jeux concours, un moyen « fun » pour augmenter l'engagement**

En plus du confort et de l'expérience sportive, il y a aussi une amélioration des divertissements proposés avant, pendant et après les matchs. Hyundai et l'Olympique Lyonnais par exemple propose une course de voitures sur leur application. Le gagnant se voit offrir des places confortables et proche de la pelouse<sup>26</sup>. Cela amène un côté amusant à l'application ainsi qu'un réel engagement des fans dans l'espoir de gagner les places. Du côté club, le jeu permet d'augmenter le nombre de téléchargement de l'application, la quantité et qualité de contenu sur les réseaux sociaux ainsi qu'un lien encore plus fort avec la communauté de supporter.

- **Adaptation du spectacle et des infrastructures en fonction des commentaires des spectateurs : l'exemple de Roland Garros<sup>27</sup>**

IBM, partenaire de Roland Garros depuis plus de 30 ans, a développé un logiciel basé sur l'intelligence artificielle Watson pour analyser et synthétiser le contenu concernant le tournoi posté sur les réseaux sociaux. Cela permet d'avoir des retours en temps réel sur des éléments aussi divers que les files d'attente, la qualité de la restauration ou les équipements du tournoi. Les organisateurs peuvent alors s'adapter, réagir pour améliorer la satisfaction des spectateurs. L'analyse de ces réseaux sociaux permet également de déterminer la perception des différents joueurs par le public de façon quantitative (nombre de mentions sur les réseaux sociaux) et qualitative (pourcentage de tweets positifs vs négatifs par exemple). Cela aide notamment les organisateurs à déterminer quels joueurs vont jouer sur quels cours (cours principaux ou cours

---

<sup>26</sup> Olympique Lyonnais, une activité e-commerce en très forte croissance, Anthony Alyce, EcoFoot, 06/09/2017

<sup>27</sup> L'IA et les robots envahissent le tennis, France 3, 21/05/2019

annexes) à quelle heure ou encore quels ceux qui doivent le plus mis en avant sur les supports de communication afin d'attirer le plus de monde. A noter qu'IBM et Roland Garros ont cessé leur partenariat en 2019 au profit d'Infosys, une société indienne qui a notamment choisi de développer une plateforme de statistiques réservée uniquement aux joueurs et joueuses du tournoi. Cette dernière leur permet d'avoir accès, moins d'un quart d'heure après la fin d'une rencontre, à tous les points d'un match découpés ou compilés selon de nombreux critères de recherche (volée, smash, revers, coup droit, amorties, lob) ; la reconnaissance des gestes s'effectuant sans intervention de l'homme grâce au système d'intelligence artificielle. On retrouve également une large base de données accessible par l'ensemble des spectateurs ou encore une plateforme d'analyse des réseaux sociaux.

- **La data war room<sup>28</sup> : tout surveiller pour tout contrôler et tout optimiser**

De l'autre côté de l'Atlantique, c'est sur l'aspect logistique qu'on a mis l'accent. Le Levi's Stadium en Californie est l'antre des 49ers, une équipe de football américain. Pour fluidifier et améliorer la logistique du stade, une « data war room » a été créée (traduisez « salle de guerre de données »). Une batterie de données est collectée en direct durant l'évènement pour agir sur plusieurs points de frictions possibles :

- Parking : détecter les parkings complets pour rediriger les véhicules vers d'autres zones de parking.
- File d'attente : identifier les plus vides pour alerter les équipes sur place et orienter les supporters afin d'optimiser toutes les files.
- Problèmes : observer tous les points afin d'agir rapidement en cas d'incident.

---

<sup>28</sup> Précurseurs de l'analyse de données, Laure Gruffat, FanStriker.com, 05/10/2018



*Précurseurs de l'analyse de données, Laure Gruffat, FanStriker.com, 05/10/2018*

Malgré quelques mouvements de contestations, l'avènement des données a permis aux clubs d'enrichir considérablement l'expérience de leurs fans durant une rencontre sportive. Ils ont désormais accès à plus d'informations en temps réel pour animer le commentateur sportif qui sommeille en eux, ils ont plus facilement accès à la restauration durant le match et toutes les files d'attentes ont été fluidifiées. Il y a plus d'animations autour et dans le stade ce qui permet aux clubs de ramener un public parfois plus familial mais aussi un surtout un public qui vient plus longtemps. Aujourd'hui, on vient au stade plusieurs heures avant le match afin de profiter des activités, de la buvette et des différents stands présents. Cela profite aux supporters comme aux clubs qui engrangent plus de revenus et collectent plus de données sur leurs spectateurs grâce à un nombre grandissant de points de contact.

## **B – Mieux cibler les supporters**

Dans l'optique de satisfaire davantage les spectateurs et de maximiser les revenus, le milieu sportif a progressivement adopté le concept de segmentation des supporters, à l'image de la segmentation des clients utilisée par les entreprises traditionnelles. Les supporters ne sont plus ainsi considérés comme un seul et même groupe mais comme une somme d'individualités. Les clubs se sont vite rendu compte qu'il y a différents types de fans ou de publics. Ils ont différentes attentes, différents besoins, différents budgets : un profil tout simplement différent. Les structures sportives dont donc s'adapter pour pouvoir répondre à chacun d'entre eux.

Bien analyser ou appréhender cette dimension implique plusieurs enjeux : l'image que le club renvoie à travers l'ambiance de ses tribunes et le remplissage de son enceinte. Est-ce que le stade est rempli ? Si oui, cela renvoie une belle image du club. Est-ce qu'il y a de l'ambiance ? Si oui, cela laisse croire qu'il y a une vraie ferveur autour de l'équipe. Est-ce qu'il y a des débordements provoqués par la cohabitation de plusieurs types de supporters ? Si oui, cela risque de décourager certaines personnes de venir au stade.

Il faut donc bien segmenter et cibler les supporters pour faire venir les différents groupes qui sauront cohabiter et mettre une bonne ambiance. En plus de cela, il faut savoir les fidéliser pour continuer à remplir le stade tout au long de l'année. Pour pouvoir faire tout cela, il faut savoir proposer les bons prix pour chaque match, toucher les bonnes personnes au bon moment et proposer un nombre de places adéquat.

- **Le Yield Management pour remplir les stades**

La notion de Yield Management est ici très importante. Tous les matchs ne présentent pas le même niveau d'enjeux et n'attirent pas le même nombre de spectateurs. A l'instar des prix de billets d'avion ou d'hôtels, les prix des billets vont fluctuer en fonction du match et du moment de l'achat. Il est logique qu'un match opposant deux rivaux attire plus de fans qu'une rencontre entre le premier et le dernier du championnat. Dans le premier cas, la demande va augmenter et par conséquent le prix aussi. Au contraire, le deuxième cas va attirer moins de spectateurs et donc créer un risque d'aboutir un stade au remplissage limité. Les prix vont donc baisser et des offres de dernières minutes seront proposées pour remplir au maximum les tribunes.

C'est là où l'utilisation du Big Data prend tout son sens. Toutes les données collectées de manière constante vont permettre de modifier l'offre en temps réel. Des acteurs comme Arenamatrix ont en fait leur mission via différents CRM et outils d'analyse marketing pour



aider les clubs à augmenter leur taux de remplissage. Une entreprise comme Seaters, va elle permettre aux organisateurs de combler les places vides grâce à des file d'attentes digitales qui permettent d'obtenir une place jusqu'à la dernière minute.

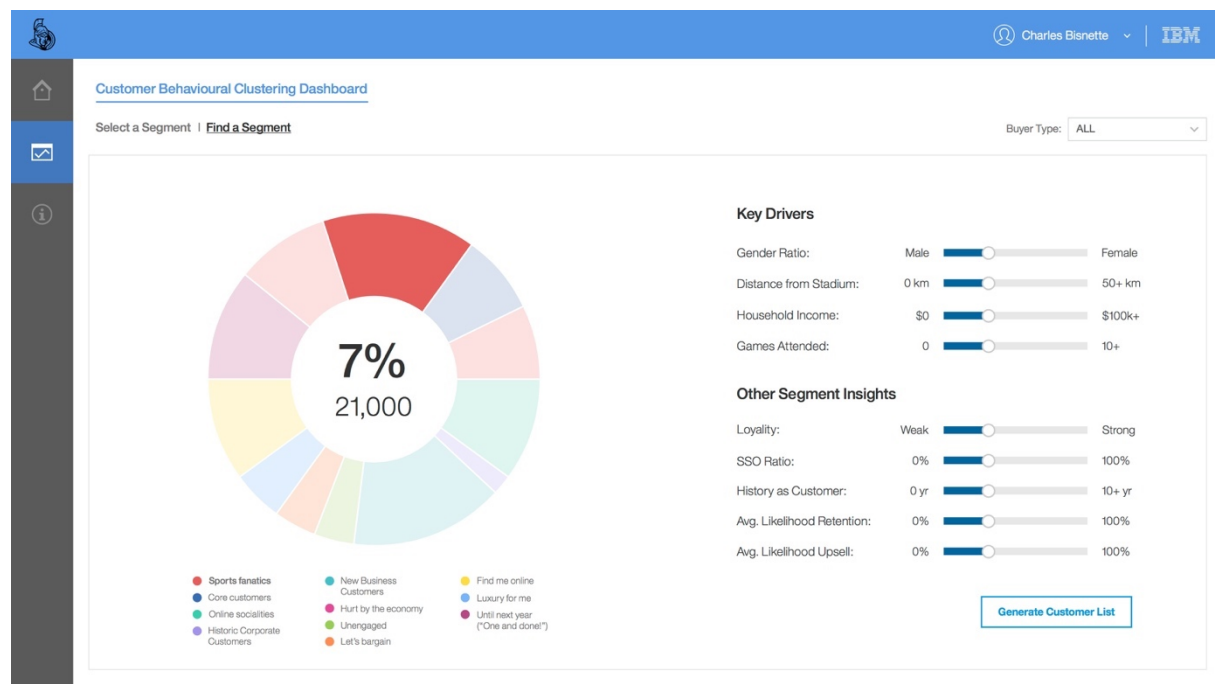
A ce niveau-là, l'Olympique Lyonnais doit faire face à une grosse problématique. S'il est indéniable que Lyon est une ville de football, c'est également une ville qui dispose de clubs de basket et de rugby de haut niveau. Le club doit ainsi s'adapter à cette concurrence pour remplir un stade qui peut accueillir jusqu'à 59 000 personnes. L'autre problème réside dans la localisation du stade. Anciennement à Gerland et donc proche du centre-ville, le stade se situe désormais à Décines en banlieue de Lyon. Du centre, il faut compter au moins 45 minutes pour s'y rendre et cela décourage beaucoup de personnes. Pour faire face à cette problématique, l'Olympique Lyonnais joue avec le Yield Management. En dehors des matchs importants, le stade est rarement plein. Cela pousse le club à proposer de nombreuses offres promotionnelles : offres étudiantes ou offres de dernière minute par exemple à des prix particulièrement attractifs. Ceux-ci peuvent en effet descendre jusqu'à 8 euros pour un match de ligue 1.

- **Le logiciel IBM Behaviour Based Fan Insight: mieux comprendre et cibler les spectateurs**

IBM Behaviour Based Fan Insight est un logiciel d'analyse de données qui a pour objectif d'augmenter le prix des billets et les revenus issus des concessions et commerces du stade via différentes actions.

Les fonctionnalités sont diverses et couvrent un grand périmètre. En amont, le logiciel va analyser un grand nombre de données brutes liés aux réseaux sociaux et autres modes de communication. Il va aussi permettre de comprendre comment le contenu contextuel (contenu basé sur les informations collectées lors des visites récurrentes) peut impacter l'engagement des spectateurs. En plus de cela, il va permettre de comprendre les préférences des fans pour proposer une communication personnalisée. Une fois la campagne lancée, il va fournir une visibilité en temps réel sur l'efficacité des promotions via différentes données clés définies au préalable. Toujours dans l'optique de vouloir toujours mieux prédire l'avenir, IBM Behavior Based Fan insight permet d'anticiper l'impact des campagnes, mesurer le pourcentage de conversion et évaluer les revenus potentiels. Une fois le profil du fan élaboré, le logiciel trouvera des offres qui lui correspondent en fonction de ce qui l'a attiré par le passé et ira même jusqu'à prédire le timing optimal pour le contacter. Enfin, pour pousser encore plus à

consommer, le logiciel saura proposer un classement quotidien de produits recommandés aux clients selon leur budget et préférences.



*IBM Fan Insight, IBM.com*

Une vraie arme d'analyse, de prédiction et d'action qui aide par exemple les Sénateurs d'Ottawa<sup>29</sup>, l'équipe professionnelle de hockey sur glace de la ville. Les deux parties collaborent sur 5 points principaux :

- Identifier et comprendre les attitudes des fans, les tendances de comportement et les caractéristiques psychologiques afin de reconnaître leurs préférences et leurs offrir des interactions personnalisées.
- Analyse des interactions entre les fans pour découvrir les opportunités d'interaction entre marques/sponsors/partenaires et fans.
- Suivre le succès et le retour sur investissement des initiatives et des campagnes en examinant les statistiques des fans et découvrir des clés et relais de croissance.
- Regroupement des données provenant de sources internes et externes et amener une vraie flexibilité dans la collecte et l'exploitation.
- Création des offres cibles efficaces pour générer des renouvellements et des ventes tout en attirant de nouveaux clients.

<sup>29</sup> Overview of behavior based fan insight, IBM

- **Les réseaux sociaux et autres jeux concours, une mine d'or pour les clubs**

L'utilisation des réseaux sociaux, les différents jeux concours ou encore les détails d'achat sont de très bons moyens pour collecter des informations sur les fans. Cela va aider à la segmentation du public et permettre au club de pousser le bon contenu au bon profil de supporter. En fonction du nombre de places achetées, du timing de l'achat, du type de match, de la tribune ou encore du type de goodies acheté, le club va pouvoir catégoriser le spectateur dans un profil bien déterminé. Une personne identifiée comme venant régulièrement en famille verra par exemple du contenu lié à la création de souvenirs avec ses enfants. Si par exemple un fan a acheté plusieurs billets séparément pour la même compétition, il se verra proposer l'année d'après un package comprenant plusieurs places. Une personne qui n'est pas venu au stade depuis longtemps recevra des offres promotionnelles pour réactiver son profil et le pousser à revenir au stade. Les personnes qui ont utilisé le Wifi du stade pendant le match sont potentiellement des fans plus engagés et donc peut-être plus enclin à racheter des billets. Les exemples sont nombreux tant les points d'analyse et de collecte de donnée sont nombreux.

- **Entretien terrain – David Banget, Directeur du Digital à l'Olympique Lyonnais**

Afin de pouvoir obtenir un exemple concret avec des informations plus pertinentes car venant du club directement, nous avons eu l'opportunité d'interviewer David Banget, Directeur Digital à l'Olympique Lyonnais. Voilà les principales informations sorties de notre entretien.

### Contexte

Le poste de David Banget a été créé il y a 5 ans de cela. Le contexte est le suivant: après des années glorieuses au stade de Gerland, tant sportivement que financièrement parlant, la direction du club voit la construction de son nouveau stade se terminer. Pensée depuis bien longtemps, cette restructuration est à la fois interne et externe. Interne car le club sait alors depuis plusieurs années qu'il se doit de réinventer son modèle pour suivre les tendances et transformations actuelles. Externe en la création de ce nouveau stade qui est parmi les plus modernes dans le monde du football. Comme évoqué précédemment, les enjeux sont désormais portés sur le digital et c'est dans cette optique que David rejoint le groupe de l'Olympique Lyonnais. Une refonte se fait et le marketing est depuis étroitement lié aux équipes techniques digitales. C'est un pôle qui se veut transversal et qui vise à optimiser la communication digitale, la relation client, le data marketing ainsi que l'expérience client.

### CRM et data marketing : 3 principaux enjeux

« Identifier 100% des clients, qu'ils viennent du monde professionnel ou du grand public. En effet cela n'est pas toujours évident car du côté B2C par exemple beaucoup de personnes prennent des places pour leurs amis ou leurs familles empêchant ainsi l'équipe de David d'identifier tous leurs clients.

La mise à disposition d'une vue en 360 degrés de ses clients. Par cela, il est entendu que l'on souhaite avoir le plus d'informations possibles sur tous. Tranche d'âge, adresse e-mail, SIRET pour une entreprise, lieu d'habitation, nombres de connexions aux différents portails de l'Olympique Lyonnais etc. Le but étant de mettre ces informations à disposition des commerciaux pour les aider dans leur travail.

Utiliser toutes ces données clients afin de personnaliser chaque approche et ainsi gagner en efficacité. Pour cela, le pôle a mis en place des arbres de direction ou parcours marketing cross-canaux. Les clients sont classés par catégories et la manière avec laquelle ils vont être atteints sera différente. Si le client ouvre souvent ces mails alors son parcours marketing sera adapté en fonction. S'il a déjà assisté à plusieurs matchs il recevra un certain type d'email plutôt qu'un autre. De cette manière-là, beaucoup de gens sont touchés et cela de la manière la plus optimale possible. La géolocalisation elle permet de gagner en portée. En effet lorsqu'un client se trouve aux abords du stade et qu'il n'a pas sa place par exemple, une notification ou email lui sera envoyé pour lui rappeler que le match est demain et qu'il y a encore des places disponibles. »

Comme nous l'a si bien expliqué David Banget, les CRM et autres bases de données sont incontournables lorsque l'on souhaite mieux connaître et donc mieux cibler ses fans. En amont, les réseaux sociaux et autres jeux concours sont des moyens à la fois amusants et efficaces pour récolter un maximum d'informations et ainsi alimenter le profil digital des supporters. En plus de cela, il existe des logiciels comme celui développé par IBM : Behavior based fan insight. Ce dernier permet de gérer l'ensemble de ces aspects : de l'analyse à la récolte de données en passant par des recommandations d'actions/campagnes et des prédictions sur les retombées. Un outil exceptionnel qui est de plus en plus utilisé dans le monde des sports collectifs professionnels. Au-delà d'une approche plus personnalisée et efficace, les clubs aspirent aussi à augmenter le taux de remplissage de leurs stades. Dans cette optique, ils optimisent les offres et prix des billets via un yield management se basant sur une batterie de données collectées dans le passé et en temps réel.

## **C – Augmenter les revenus liés au sponsoring et autres partenariats**

La data est précieuse pour un autre aspect du business sportif : le sponsoring et les partenariats. A une époque depuis longtemps révolue, les équipes de sport professionnelles payaient pour s'équiper. Maintenant c'est le contraire. Des équipementiers comme Nike ou encore Adidas débourse des centaines de millions d'euros pour s'allier aux meilleures équipes. Des entreprises comme Allianz ou même Groupama donnent leur nom à des stades (phénomène de « naming ») et, aux Etats-Unis, des entreprises comme StubHub payent des millions pour apposer leur logo sur le maillot des Philadelphia 76ers en NBA.

Avec tous ces changements, l'époque où les partenariats entre clubs et marques s'arrêtait à un simple nom sur un maillot est révolue. Désormais et comme pour de nombreuses pratiques, la tendance est au cross-canal. En effet, les points de contacts entre fans et marques/clubs sont de plus en plus nombreux. Les supporters sont de plus en plus connectés, que ce soit en ligne et paradoxalement hors ligne. A travers un « marketing 4.0 », l'idée est donc de pouvoir les atteindre tout au long de leur parcours : de l'achat du billet en ligne à la célébration de la victoire de leur équipe en passant par la mi-temps. En plus de cela, les clubs et partenaires essayent sans cesse d'augmenter la surface commercialisable de contenus publicitaires dans les stades pour accroître encore plus leurs revenus.

### **Le sponsoring**

En effet, le sponsoring est devenu en enjeu majeur. Pour les clubs comme pour les sponsors. De plus en plus de personnes assistent à des événements sportifs, que ce soit en sur place ou à travers un écran. Cette visibilité accrue vient s'ajouter au fait que le lien entre le fan et son club est un lien très fort. Ce lien émotionnel qui véhicule les valeurs nobles du sport va déborder sur la relation que le fan a avec l'équipementier et aura par conséquent un impact positif sur la marque. L'enjeu n'est donc plus à démontrer et il faut que les clubs négocient habilement leurs contrats pour en tirer un maximum de bénéfices financiers.

Avec plus de précisions et d'informations sur les profils des personnes présentes dans le stade ainsi que devant leur télé, un club est en mesure d'avancer plus précisément le nombre de personnes correspondant à la cible des équipementiers. Ce n'est donc pas au hasard qu'un géant comme Uber Eats est présent en grand sur le maillot de l'Olympique de Marseille par exemple.

- **Communiquer plus intelligemment**

Les CRM sont une vraie mine d'or pour les clubs. Ils permettent de stocker les informations personnelles sur les fans et servent de bases pour créer un profil fan digital. Ses données sont exploitées par le club mais aussi par ses partenaires. En effet, lorsqu'une communication vient de la marque directement, elle peut être prise comme envahissante ou comme un spam. En revanche, si elle vient d'un club et donc de la relation affectueuse entretenue avec la cible, la communication sera plus susceptible de toucher plus de monde. Un supporter reçoit un email de la part d'Adidas, la première chose qui lui viendra en tête serait sûrement « encore un spam d'Adidas... » alors que s'il reçoit un email de son club préféré, le Real Madrid (sponsorisé par Adidas) il sera plus enclin à l'ouvrir en se disant « tiens, peut-être des informations pour le prochain match ou des un jeu concours ! ». Le plus dur sera alors fait, la cible aura au moins ouvert l'email.

- **Des contenus publicités personnalisées en fonction de la tribune**

En termes de publicité dans le stade, là encore, les données sont cruciales pour améliorer les différents rendements liés à ce domaine. Avec différents types de données, les annonceurs pourront afficher des pubs mieux ciblés et plus personnalisées. Quel profil de supporter est majoritaire dans telle tribune, quels sont les déplacements de fans, quels achats ont-ils effectué dans le stade ou sur la e-boutique du club, sont-ils là pour un évènement organisé par leur entreprise ? La liste des données potentiellement exploitable est longue. Au lieu d'afficher le même contenu sur chaque panneau publicitaire, les annonceurs peuvent adapter le contenu en fonction du profil qui sera plus susceptible de regarder l'annonce/la démographie de la tribune concernée.

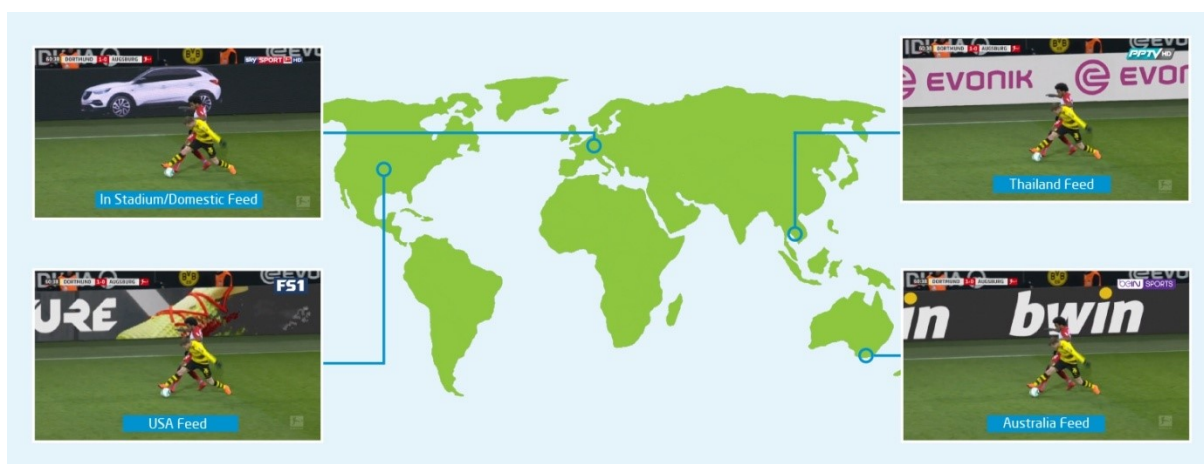
Prenons un exemple concret : une marque d'équipements sportifs est informée par un club que pour le match de dimanche, la tribune C comportera majoritairement des familles composées de parents et enfants tandis que la tribune J (suite à une offre promotionnelle) sera constituée essentiellement d'étudiants. Elle peut donc décider d'afficher du contenu d'équipements pour enfant en face de la tribune C et des équipements pour une tranche d'âge étudiante en face de la tribune J. Cela assurerait donc un plus grand retour sur investissement pour la marque et une meilleure attractivité pour les clubs qui proposeraient ce type de renseignements possibles grâce au Big Data. En effet, le nerf de la guerre des contenus publicitaires est d'avoir le meilleur

retour sur investissement en touchant les cibles choisies. Grâce à cette personnalisation, clubs et partenaires amélioreront leurs rendements et généreront donc plus de revenus.

- **La publicité virtuelle : un stade, plusieurs retransmissions, différents contenus**

La surface des espaces publicitaires dans une enceinte sportive est une manne financière importante pour les clubs. Plus les partenaires seront nombreux, plus ils pourront afficher du contenu, plus ils payeront les clubs.

L'AFL a été la première à officialiser un partenariat dans ce domaine avec la plateforme Supponor. La ligue allemande de football a approuvé en avril 2018<sup>30</sup>, l'utilisation et l'application de cette nouvelle forme de publicité révolutionnaire. En effet, selon le pays depuis lequel le supporter regarde le match, le contenu affiché sur les panneaux LED autour du terrain sera différent. Cela va indéniablement augmenter les revenus publicitaires des clubs allemands. Le cabinet spécialisé Nielsen sports estime en effet l'augmentation des revenus publicitaires annuels des clubs allemands aux alentours de 60 millions d'euros.



*Qu'est-ce que la publicité virtuelle, DBM, Sports and Business Mag, 05/04/2018*

Dans l'exemple ci-dessus, le match opposant Dortmund à Augsbourg est une bonne illustration de cette nouvelle approche. En haut à gauche, la retransmission en Allemagne affiche la pub d'un SUV Hyundai sur l'espace publicitaire en bord de terrain. En Thaïlande, c'est le partenaire Evonik qui est affiché. En bas à gauche, c'est une pub pour la nouvelle paire de crampons Puma

<sup>30</sup> Qu'est-ce que la publicité virtuelle, nouvelle technique développée en Bundesliga, SBM, 05/04/2018

qui est affiché pour les téléspectateurs des États-Unis. Enfin Bwin, spécialiste du pari sportif est le contenu poussé pour les personnes regardant le match depuis l'Australie.

Le Paris-Saint-Germain en ligue 1 a récemment adopté la technologie mais les précurseurs en la matière restent la NHL au Canada. Le hockey sur glace, sport roi au Canada a implémenté depuis maintenant plusieurs années la publicité virtuelle dans les patinoires des clubs professionnels. D'ordinaire, le plexiglass présent tout autour de la glace est une surface inexploitable pour les clubs sportifs. Si du contenu venait à être apposé dessus, les spectateurs ne verraient plus grand-chose. Grâce à la publicité virtuelle, les clubs sont en mesure d'exploiter cette surface pleinement. Grâce à la technologie d'incrustation TV, les supporters sont exposés à travers leurs écran aux logo des différentes marques partenaires des clubs. C'est le cas lorsque l'on regarde par exemple un match des Canadiens de Montréal :



*NHL un affichage publicitaire virtuelle, Alexandre Bailleul, SBB, 10/03/2016*

On y voit tour à tour des publicités pour la chaîne de fastfood McDonald ou encore Rona, détaillant et distributeur de matériaux de construction.



## CONCLUSION PARTIE 2 : l'approche individualisée d'un ensemble

Améliorer l'expérience des fans, encore et toujours. C'est le but des clubs tant l'optimisation des revenus liés au stade est cruciale. Le premier enjeu est de faire venir les supporters au stade et ici encore, les données sont d'une importance capitale. Aujourd'hui, les clubs sont capables de combiner toutes les données qu'ils collectent avec du yield management afin d'optimiser le taux de remplissage de leurs stades. Je vais notamment citer une expérience personnelle très récente avec l'Olympique Lyonnais qui illustre bien ce phénomène. Ayant déjà assisté à plusieurs matchs les années précédentes, j'ai renseigné plusieurs informations pour créer mon profil. Parmi celles-ci figure le fait que je sois étudiant. Le match entre l'Olympique Lyonnais et le FC Metz avait lieu le 26 Octobre dernier. Historiquement, les rencontres avec une équipe de bas de tableau peinent à remplir le stade. Ayant anticipé la faible affluence, le club a décidé de brader les prix à 10 euros la place et en plus de cela envoyer des offres étudiantes aux écoles partenaires dont l'emlyon business school. Une place à 6 euros nous était offerte. Opération réussie : avec 48 889 spectateurs<sup>31</sup>, (sur 59 000 possibles) l'affluence de ce match correspond à la moyenne 2018 de 49 564 spectateurs<sup>32</sup>. Quand on sait que l'année dernière a tracté un bel engouement et que le FC Metz n'est pas l'équipe qui attire le plus d'amateurs de football, l'affluence est plus que respectable. Elle valide la stratégie de l'Olympique Lyonnais.

Une fois le stade rempli, il faut parvenir à enrichir l'expérience des fans. Aujourd'hui, on ne vient plus au stade pour un match mais pour bien plus. Pour tout ce qui se passe autour aussi. Des logiciels comme IBM Behavior Based Fan Insights permettent aux clubs de connaître encore mieux leurs supporters. Connaître les tendances, les interactions qu'ils ont avec le club, leurs réactions à certains contenus, les chiffres et résultats d'une certaine compagne. Les possibilités sont nombreuses et cela permet donc la création de contenu personnalisé pour le match ou encore d'offres ciblés dans l'enceinte. Comme l'a mis en lumière notre entretien avec David Banget, les CRM et autres data marketing un temps négligé sont désormais incontournables. On veut tout connaître sur chaque supporter pour personnaliser de plus en plus les interactions qu'ils auront avec leur club préféré.

Enfin, les clubs bénéficient de toutes ces avancées technologiques et autres collectes de données dans leurs partenariats. En effet, avec toujours plus de données, toujours plus de précisions sur

---

<sup>31</sup> Stade OL HD, Twitter, 26/10/2019

<sup>32</sup> Article du site de l'OL, 06/2019

les retombées et toujours plus de prévisions, ils sont en mesure de négocier de meilleurs partenariats que ce soit pour du contenu publicitaire ou du sponsoring. On voit par exemple un phénomène de naming des stades grandissant dû à une évaluation des retours sur investissement toujours plus précise et plus grande. On est désormais en mesure de profiler toute une tribune pour déterminer le meilleur contenu à pousser en face de cette même tribune et adapter chaque panneau à chaque endroit du stade pour optimiser le ciblage. Pour aller plus loin, les publicités virtuelles permettent maintenant au club d'agrandir la surface commercialisable pour afficher des contenus publicitaires. Précurseurs en la matière, la NHL (championnat de Hockey sur glace au Canada) a initié le mouvement. Selon le pays d'où le téléspectateur regarde le match, le contenu affiché sur les panneaux publicitaires sera différent et cela est une vraie révolution tant les clubs peuvent s'associer à plus de partenaires et leur assurer une audience plus adaptée ainsi qu'un temps d'exposition plus important.

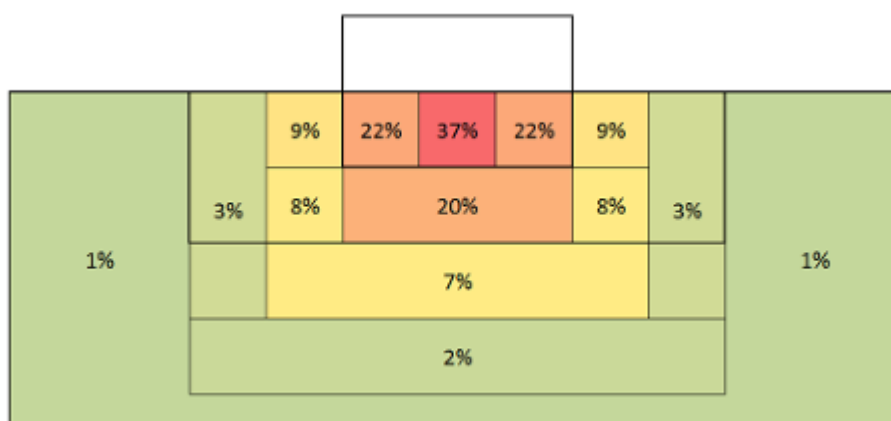
## PARTIE 3 : L'avenir de l'IA et du Big Data dans le sport remis en question

Nous avons développé au cours des deux parties précédentes tous les apports des données, de leur collecte et analyse par l'intelligence artificielle. Il est indéniable que ces technologies ont énormément fait évoluer le sport professionnel, qu'elle le transforme encore aujourd'hui et qu'elles continueront à le faire à l'avenir. Toutefois, il semble qu'il existe de nombreuses limites à ces bénéfices, des « angles morts » du Big Data et de l'IA que les améliorations technologiques cherchent à combler mais, au fond, cela est-il vraiment souhaitable ?

### A – Des apports mais aussi beaucoup de limites

- Les « expected goals », des probabilités pour des prédictions de conversion imparfaites

Comme l'illustre le concept de « sabermétrie » et les nombreux exemples évoqués dans ce document, le monde du sport est en recherche permanente de nouvelles statistiques qui se veulent toujours plus précises, toujours plus pertinentes. Toujours plus de données collectées, d'outils de mesure et d'analyse pour permettre aux staffs de gagner un avantage comparatif et d'offrir aux spectateurs une nouvelle grille de lecture. Le dernier indicateur à la mode dans le milieu du football, ce sont les « expected goals » ou « buts attendus ».



*Les expected goals, Julien Assuncao, Cahier du Football, 01/04/2015*

Cette statistique a fait beaucoup de remous dans le monde du football tant elle a semblé à la fois précise et aléatoire. Qu'est-ce que représentent les « xG » ? Elle est d'abord basée sur une donnée. La probabilité qu'un joueur a de marquer dans position donnée, avec telle partie du corps lorsqu'il enclenche une frappe. En d'autres termes, la probabilité qu'une occasion soit

convertie en but marqué. Cette donnée elle-même est basée sur des dizaines de milliers de données collectées par le passé sur des dizaines de milliers de tirs et comprend :

- Le type de tir : balle frappée avec la tête, le bon pied du joueur, le mauvais pied du joueur, le genou etc.
- Le type de passe : Est-ce que le joueur a reçu la balle dans les pieds, dans la course, en hauteur, au sol ?
- La situation d'attaque : Est-ce une contre-attaque où la défense n'a pas eu le temps de se replier ? Est-ce une attaque placée avec une défense bien en place ? Est-ce une phase arrêtée où l'aspect tactique/stratégique est encore plus amplifié ?

Une note entre 0 et 1 va alors être attribuée à l'occasion pour qualifier sa « dangerosité ». Un tir du milieu de terrain par exemple aura un  $xG = 0.01$  (probabilité de marquer de 1%) car par le passé, les tirs du rond central n'ont que très rarement aboutit à un but. Au contraire, une frappe dans les 6 derniers mètres après une passe en retrait devant un but vid aura un  $xG = 0,97^{33}$  car par le passé, la plupart des occasions similaires ont été transformés en but.

A partir de là, le but des  $xG$  est donc de déterminer plusieurs choses : la qualité d'une occasion de but et les buts attendus d'une certaine équipe ou d'un certain joueur. En se basant sur ces statistiques, il est possible de comparer les faits à « ce qui aurait dû se passer ». En évaluant la différence entre le nombre de but marqués attendus ainsi que sa variante le nombre de buts encaissés attendus et les statistiques réelles, on peut estimer qu'un joueur ou une équipe « surperforme » ou « sous-performe ». Prenons l'exemple d'un attaquant qui réalise une saison de qualité avec 15 buts marqués : si le total de ses expected goals s'élève à 10 seulement, on pourra penser que ce dernier surperforme (induisant une certaine forme de réussite) et qu'il risque de rentrer dans les rangs la saison suivante. C'est donc une donnée clé pour les recruteurs pour savoir s'il faut investir sur un joueur qui apparaît comme une perle rare ou qui flambe temporairement mais ne tiendra pas forcément sur la durée. Cette nouvelle statistique a engendré de nombreuses variantes, allant jusqu'à la possibilité de refaire le classement d'un championnat selon les « expected points ». Ce concept peut également être appliqué aux passes décisives ou aux tirs de coups francs par exemple. Ce nouvel indicateur dont la création a été permise par le Big Data et l'intelligence artificielle semble donc être en apparence le futur du

---

<sup>33</sup> Que désignent les expected goals, Wiloo, Youtube, 03/2018

football (et potentiellement du sport en général si on envisage de l'adapter à différentes disciplines).

Cette statistique comporte cependant trois principaux défauts. En premier lieu, elle ne prend pas en compte la qualité/le talent intrinsèque du joueur qui a la balle. En effet, une occasion de but a plus de chance d'être convertie par un grand attaquant que par un défenseur de niveau moyen. C'est l'exemple de l'attaquant de Tottenham, Harry Kane. Cela fait trois ans d'affilé qu'il surpasse les expected goals qu'on lui a attribués et cela est dû principalement à sa qualité unique de buteur. Le deuxième défaut est que la batterie de statistiques prise en compte pour qualifier une occasion n'est pas assez complète. Est-ce qu'il pleut ? Où est placé le défenseur au moment du tir ? Est-ce que les coéquipiers se déplacent pour brouiller les pistes et augmenter la probabilité du joueur de marquer ? Tout bêtement, est-ce qu'un supporter pointe un laser sur le gardien adverse ? Tant de variables qui ne sont pas prises en compte malgré un impact indéniable sur la qualité d'une occasion de but. Enfin, les xG ne considèrent pas la chance, la réussite. Un des meilleurs exemples est la Juventus de 2015<sup>34</sup>. Grandissime favoris pour remporter le championnat, elle pointait à la 12<sup>ème</sup> place au bout de 10 journées. Bien en dessous de ses xG, l'équipe souffrait et la direction été encouragée de toute part à licencier l'entraîneur. Après avoir repassé en revue les premiers matchs, elle considérait que c'était un simple manque de réussite tant la qualité du jeu était élevée. La suite leur a donné raison : 15 matchs, 15 victoires et un titre de champion dans la poche. Comme le montre ces exemples, les statistiques ne sont pas toujours reines dans des sports au caractère aléatoire si fort. Que ce soit sur le court ou long terme, les statistiques sont de bons indicateurs mais ne constituent pas un plan infallible. Il paraît impossible de déterminer l'ensemble des variables ayant un impact, aussi infime soit-il, sur une rencontre sportive. Une de ses variables est particulièrement compliquée à appréhender : la nature humaine.

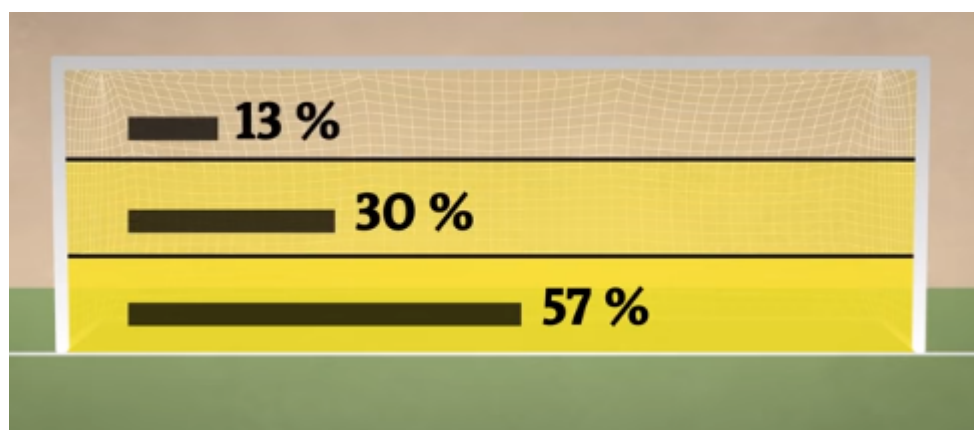
- **Quand la nature humaine s'oppose aux vérités statistiques : le cas des penaltys<sup>35</sup>**

En 2009, Michael Bar-Eli et Ofer Azar ont analysé 311 penaltys dont 94% étaient cadrés. 13% sont tirés dans la partie supérieure du but, 30% dans le tiers du milieu et 57% dans le tiers inférieur.

---

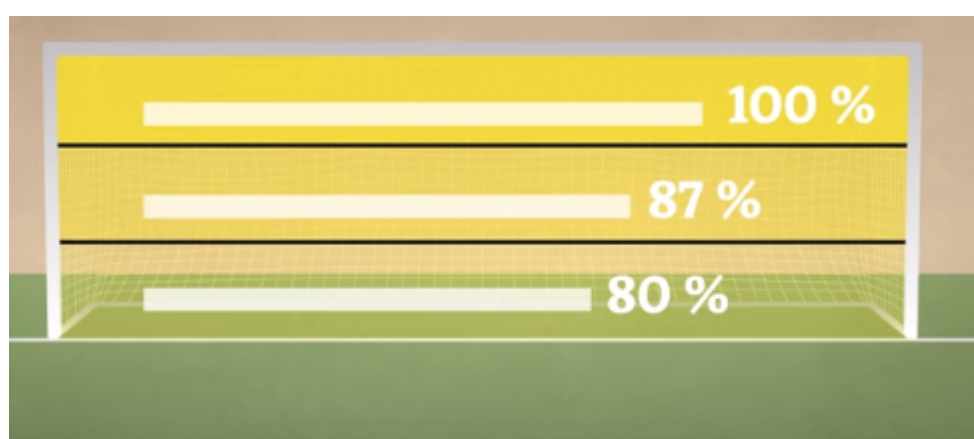
<sup>34</sup> Mais que sont les expected goals « xG » la statistique du moment, Johann Crochet, Goal, 13/11/2017

<sup>35</sup> Penalty kicks in soccer: an empirical analysis of shooting strategies and goalkeepers' preferences, Michael Bar-Eli & Ofer H. Azar, Soccer & Society, 02/2009



*Pourcentage de zones visées sur les pénaltys étudiés*

Pourtant, lorsque l'on regarde le taux de réussite par zone visée, l'ordre s'inverse totalement. Le taux de réussite du tiers supérieur est même de 100%. Alors certes, il y a plus de risques de tirer au-dessus de la cage en visant cette zone, mais même en incluant les tirs non cadrés, le taux de réussite s'élève à 90%.



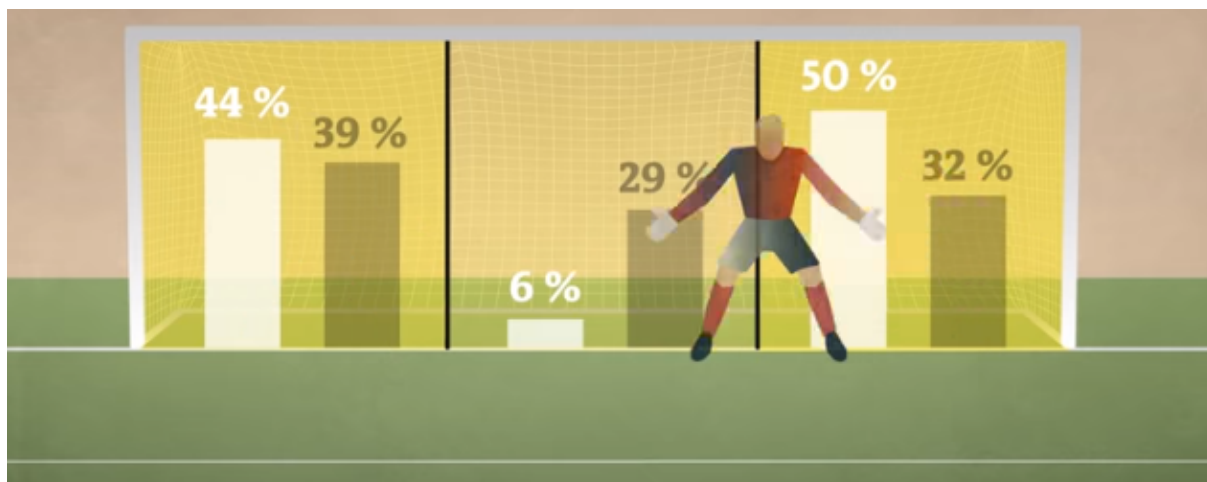
*Taux de réussite des pénaltys tirés par zone visée*

Alors pourquoi les joueurs ne visent-ils pas plus cette zone au détriment des deux tiers inférieurs qu'ils semblent plébisciter alors qu'ils auraient plus de chance de marquer ? Avec la place prépondérante des statistiques dans le sport, on peut s'étonner que cette stratégie contreproductive perdure dans le football moderne où chaque fait de jeu est décortiqué avec précaution. L'explication se situe dans la peur du ridicule et des critiques potentielles. En effet, si un joueur cadre son tir dans la zone inférieure mais que le gardien l'arrête, on mettra davantage en avant le rôle du gardien et le joueur s'exposera à moins de critiques. En revanche, s'il tire au-dessus de la barre transversale cela marquera les esprits, comment un footballeur

professionnel qui s'entraîne quotidiennement et est payé des millions peut-il envoyer un ballon au-dessus de la cage adverse en tirant à 7 mètres de celle-ci ?

Les supporters français se souviennent ainsi du pénalty envoyé sur la barre par Trezeguet en finale de la coupe du monde 2006 face à l'Italie et qui coutera aux Bleus le titre mondial. Ce penalty manqué le suivra pendant des années et écornera à jamais son image auprès du public français. Est-ce que cela aurait été le cas s'il avait tiré dans une zone inférieure et que Gianluigi Buffon avait arrêté son tir ? Rien n'est moins sûr, mais on comprend mieux pourquoi les joueurs privilégient un taux de 80% de réussite avec le risque que le gardien arrête leur tir plutôt qu'un taux de réussite de 90% avec le risque que leur tir ne soit pas cadré.

Le même phénomène s'applique de l'autre côté du ballon. En effet, une autre étude des mêmes chercheurs a démontré que si les tirs se répartissaient à peu près équitablement entre le tiers gauche, du milieu et droit du but (respectivement 39%, 29% et 32%), les gardiens choisissent un des deux côtés dans 94% des cas, ne restant au milieu que dans 6% des pénaltys étudiés. On a donc environ 1/3 des pénaltys tirés au centre du but et des gardiens qui décident de rester dans cette zone 1 fois / 17 seulement.



*Pourcentage de zones visées par les tireurs (noir) et de zones choisies par les gardiens (blanc)*

L'explication se trouve une nouvelle fois dans la peur du ridicule. Un gardien préfère rater un ballon en plongeant plutôt que rater un ballon en restant immobile, quitte à arrêter moins de pénaltys. Cela s'explique parce que les chercheurs ont dénommé le « biais de l'action ». Les spectateurs auront tendance à être moins indulgent avec un gardien qui plonge d'un côté (« ce n'est pas de sa faute », « il a choisi le mauvais côté, pas de chance ») qu'avec un gardien qui reste immobile (« il n'a même pas essayé », « c'est un mauvais gardien »).

Cet exemple permet de nous rappeler une vérité immuable, peu importe la place des statistiques, du big data et de l'importance artificielle : ce sont et ce seront toujours les acteurs sur le terrain qui auront le dernier mot ; des êtres humains avec leur personnalités leur peur, leur imprévisibilité et leurs contradictions, qui font parfois déjouer études statistiques et analyses prédictives.

- **L'échec des pronostics de l'intelligence artificielle**

Preuve que le sport reste imprévisible malgré toutes les technologies mises en place pour anticiper les résultats à venir, de nombreuses initiatives de prédiction sportive se sont cassées les dents sur la réalité du terrain.

En 2014, Google a mis à contribution sa « Google Cloud Platform » pour prédire les résultats de la coupe du monde de football et notamment la victoire de la France contre l'Allemagne en quart de finale : raté.<sup>36</sup>

En 2016, aucun des logiciels développés par Goldman Sachs, Microsoft, Yahoo, Blue Yonder ou encore l'école polytechnique fédérale de Lausanne n'ont su prédire la victoire de la sélection portugaise à l'Euro.<sup>37</sup>

Plus récemment en 2018, Goldman Sachs a à nouveau développé un logiciel reposant sur le machine learning produisant 200 000 modèles en minant des données sur les équipes et les joueurs et a fini par entrevoir 1 million de déroulés possibles pour le championnat. Calculant avec ces données en main les probabilités pour chaque équipe de battre un adversaire potentiel, le logiciel a prédit une victoire du Brésil en finale contre l'Allemagne : encore raté. Idem pour un logiciel similaire développé par la banque UBS ayant prédit la victoire finale de l'Allemagne.

Si nombre de ces algorithmes ont visé juste pour la plupart des matches, le taux d'erreur reste conséquent. Le sport est et restera imprévisible, notamment car il dépend de facteurs humains parfois irrationnels et difficiles à anticiper. Les prédictions seront sûrement de plus en plus justes dans le futur avec les progrès de l'intelligence artificielle mais nous ne serons pas à l'abri

---

<sup>36</sup> Google calcule que la France va battre l'Allemagne, Julien Lausson, Numerama, 04/07/2014

<sup>37</sup> L'euro 2016 a été compliqué pour les robots chargés des pronos, Gregor Brandy, Slate, 12/07/2016



de surprises comme le titre de Leicester en Premier League en 2015/2016 pourtant côté à 5000 contre 1 en début de saison.

De plus, certains éléments semblent difficilement mesurables, notamment tout ce qui relève de l'esthétique et de l'émotion. Un dunk au basketball comptera comme n'importe quel autre dunk dans les statistiques. Pourtant, deux dunks peuvent être plus ou moins spectaculaires et ne susciteront pas la même réaction chez les spectateurs. Il en est de même pour un dribble au football. Comment mesurer l'aura, le charisme d'un Zidane, d'un Michael Jordan ou autre Nikola Karabatic ? Ces joueurs à même de remplir des stades et soulever des foules. Il est possible d'analyser les réactions sur les réseaux sociaux mais il reste difficile de mesurer de façon objective des phénomènes qui font appel aux émotions et aux goûts de chacun.

- **La question de la sécurité des données<sup>38</sup>**

Enfin, au-delà de la pertinence des données collectées débattue précédemment, se pose la question de la sécurité de ces dernières. Une grande quantité de données collectées exigent un niveau de sécurité élevé. En effet, que pourrait-il se passer si les données d'un club tombaient entre de mauvaises mains et notamment celles d'équipes adverses ? Celles-ci pourraient avoir accès aux tactiques préparées par le club et s'y adapter par exemple. Les données physiologiques des joueurs sont également particulièrement sensibles. Des équipes adverses pourraient ainsi en profiter pour exploiter les points faibles de l'effectif. De plus, ces données permettraient d'avoir un avantage non négligeable sur le marché des transferts, ayant plus d'informations sur la forme réelle d'un joueur.

L'espionnage est monnaie courante dans le sport de haut niveau. Les cas ont ainsi été nombreux dans la formule 1, sport où la technologie est primordiale, mais aussi dans le football où Marcelo Bielsa a récemment confessé avoir envoyé un membre de son staff espionner les entraînements des adversaires de son club de Leeds. Cet « espionnage » pourrait passer au stade du numérique (si ce n'est pas déjà le cas), avec de possibles piratages des énormes bases de données existantes dans le sport de haut niveau. Si les structures sportives semblent se pencher sur cette question, les « hackers » ont souvent un temps d'avance sur les systèmes de sécurité. L'affaire récente

---

<sup>38</sup> The potential of Blockchain Technology, Christiana Henry, LawInSport, 10/05/2017

des « Football Leaks » démontre que le sport professionnel n'est pas à l'abri de fuites d'informations confidentielles.

### **B- L'amélioration continue de l'IA: suffisant pour contrer les limites actuelles ?**

Le tableau n'est toutefois pas totalement noir car les progrès de ces technologies sont exponentiels. Nous venons par exemple de citer la problématique de la sécurité des données ; sans rentrer dans les détails techniques de son fonctionnement, la technologie de la blockchain apparait comme une solution efficace à cette problématique.

On peut donc se demander quelles sont les avancées technologiques qui pourraient permettre de répondre aux failles actuelles de ces technologies ?

- **Le potentiel de la reconnaissance faciale**

La reconnaissance faciale est un outil en développement dont l'application est déjà plutôt étendue. Elle est déjà utilisée dans le domaine de la sécurité lorsque l'on souhaite délivrer une pièce d'identité à un individu ou encore lors de passages de frontières. Sommes-nous réellement qui nous prétendons être ? La reconnaissance faciale nous le dira. Le domaine de la santé y travaille aussi grâce au deep learning et à la lecture des émotions. Les professionnels de la santé sont déjà en mesure de suivre plus précisément la consommation de médicament d'un patient ou encore accompagner la prise en charge de la douleur. Dans le domaine du commerce, la Chine est bien en avance sur le reste du monde à ce sujet. Il est déjà possible dans plusieurs parties du pays de payer via un système de reconnaissance faciale. Aucune monnaie, aucune carte, aucun QR code.

#### Dans les stades

Dans le domaine du sport, son application n'est que très rare voire inexistante. Des pistes sont cependant en train de naître. Cette technologie peut être utilisée pour déterminer rapidement dans un stade l'âge moyen d'une certaine tribune par exemple. Si un club venait à découvrir que la moyenne d'âge les vendredi soir est située entre 18 et 24 ans par exemple, il pourrait utiliser cette information à son avantage. Tout d'abord en mettant l'accent sur des offres promotionnelles étudiantes pour les matchs du vendredi soir et aussi en affichant du contenu

dans le stade plus adapté à l'audience jeune. Deux manières de mieux cibler ses supporters afin d'avoir ou d'offrir un meilleur retour sur investissement.

Pour aller plus loin, les clubs peuvent utiliser la reconnaissance faciale afin de faciliter certains processus. Tout d'abord, plusieurs clubs autour du globe testent les entrées par reconnaissance faciale. Cela permettrait de fluidifier les queues à l'entrée et éviter bien des complications à l'entrée. Le système serait aussi capable de détecter les personnes définies comme VIP par exemple et donc adapter leur parcours à leur profil et donc une meilleure personnalisation de l'interaction entre le club et le supporter. Enfin, le volet sécurité est aussi une piste d'application. En effet la technologie pourrait aider à reconnaître instantanément les auteurs de troubles pendant qu'ils commettent un délit ou même après lorsqu'ils sont interdits de stade et tente tout de même d'assister à un match.

### Analyser ce que les joueurs eux-mêmes ne sauraient pas reconnaître

Le tennis est un sport où l'émotion joue un rôle essentiel. L'humeur, l'état d'âme ou encore la nervosité des joueurs ont un impact immense sur leur performance. Il n'y a qu'à voir le nombre de raquettes cassées, gestes d'humeurs et autres manifestations qui sortent un joueur de son match. On est donc en droit de se demander si par exemple, toutes les sautes d'humeur de Nick Kyrgios (immense talent australien) sont oui ou non à l'origine de ses contre-performances. Il est vrai qu'au vu des qualités tennistiques qu'il démontre sur les courts, il devrait aspirer à de meilleurs résultats. Lorsque Rafael Nadal rate un revers facile et montre des signes d'énervement, a-t-il plus ou moins de chances de remporter le point d'après ? Lorsque Roger Federer réussit un magnifique passing et harangue la foule, a-t-il plus de chances de remporter le point suivant ?

C'est là que la reconnaissance faciale intervient. En 2017, un groupe de chercheurs, professeurs, scientifiques et doctorants s'est penché sur ce sujet dans l'espoir de révolutionner un sport encore très conservateur. A travers environ 500 matchs, le groupe a compilé plus de 1700 images<sup>39</sup> de visage pour en faire leur base de données. Les images ont ensuite été traduites et classées par émotion : concentré, enflammé, exalté, abattu, anxieux, contrarié, fâché. Tout cela grâce à l'analyse de mouvements du visage comme une mâchoire qui tombe, un clignement des

---

<sup>39</sup> Facial recognition technology comes to tennis to track player emotions, John Stensholt, Financial review, 19/01/2018

yeux, des sourcils froncés et bien d'autres expressions. Ils se sont alors concentrés sur le big four à savoir Rafael Nadal, Roger Federer, Novak Djokovic et Andy Murray. Après avoir croisé les émotions avec les résultats des points observés, ils ont réussi à tirer plusieurs conclusions du jeu de ces quatre joueurs. Par exemple, lorsque Rafael Nada montre des signes d'abattement, ses chances de remporter les points suivants diminuent considérablement. Étrangement, lorsque Novak Djokovic montrait des signes d'énervement, il avait plus de chance de remporter le point suivant que s'il haranguait la foule.

Même si le tennis est un sport individuel, cette application de la reconnaissance faciale est facilement transposable à un sport collectif. Prenons une séance de tirs au but à la fin d'un match de football. Pendant que le joueur marcherait vers le ballon pour tirer, une caméra serait capable de déterminer le pourcentage qu'il a de marquer ou non selon l'expression de son visage. Cela pourrait s'ajouter à des statistiques qui pourraient être délivrées en temps réel afin d'améliorer encore plus l'expérience du fan et sa compréhension du jeu.

Aussi, cela pourrait s'appliquer dans le domaine de la santé. Lorsqu'un joueur est blessé, il est souvent difficile pour le kinésithérapeute de déterminer exactement ce que le joueur a ou quel niveau de douleur il subit avant d'avoir effectué plusieurs tests à la fois courts et longs. Avec la reconnaissance faciale, l'équipe médical du club pourrait rapidement détecter le niveau de douleur du joueur ou encore sa réaction à tel point de pression ou tel mouvement. Des douleurs que les sportifs ont souvent du mal à décrire, exprimer ou même détecter. La reconnaissance faciale pourrait donc remédier à ses problèmes et assurer une prise en charge de la blessure plus précise, rapide et efficace.

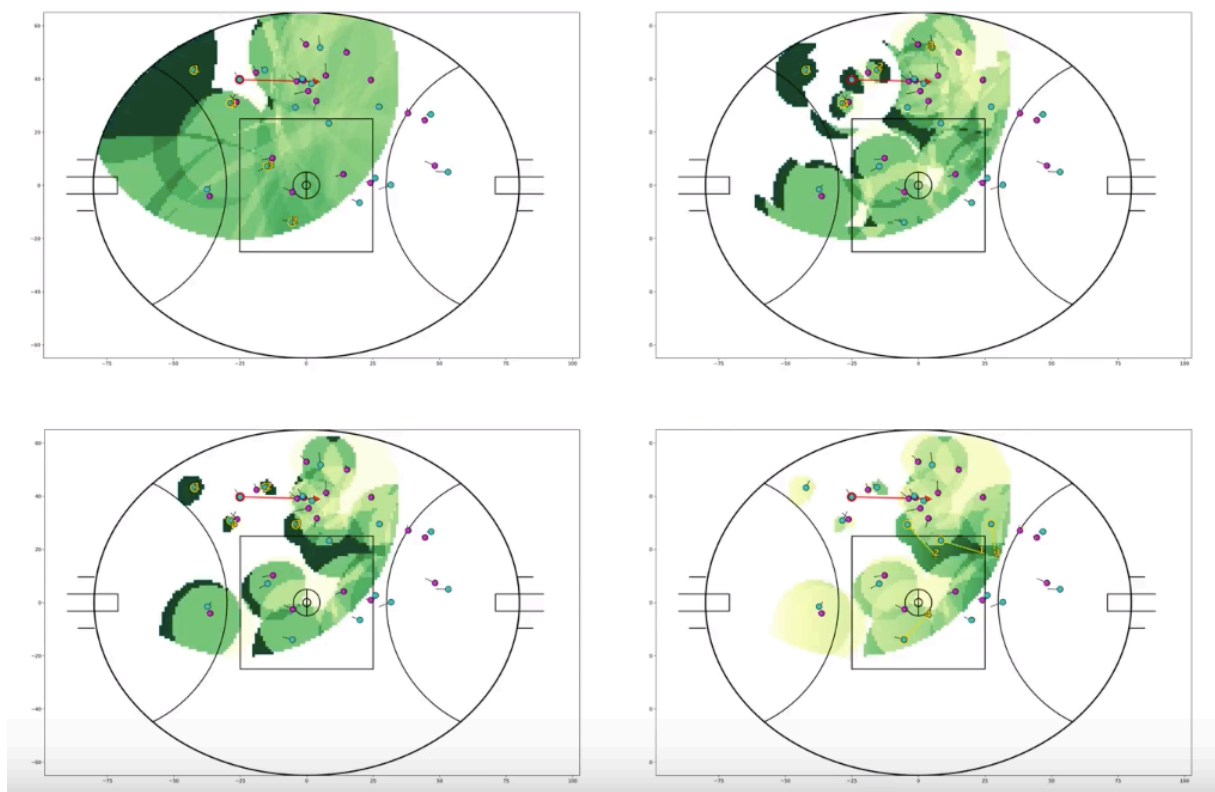
Enfin, la reconnaissance faciale pourrait aider à mesurer ce que nous disions difficilement mesurable dans la sous-partie précédente : l'émotion générée par un geste ou un joueur donné. On peut imaginer un indicateur de performance esthétique ou émotionnel basé sur l'analyse de l'expression faciale des spectateurs au moment de leur réaction face à un fait de jeu donné.

- **L'humain reste clé dans l'amélioration de l'IA : l'exemple de la meilleure passe à faire**

Que cela soit à travers l'utilisation du Big Data, de l'intelligence artificielle ou même du machine learning, l'intervention de l'humain reste nécessaire pour améliorer tout ces indicateurs. Oui, les technologies s'améliorent continuellement et nous apportent un degré de

précision toujours plus grand. Oui nos possibilités d'analyse, de traitement de données, d'observation ne cesseront de grandir. Malgré cela, une des interrogations qui subsiste est celle de l'apport humain dans ces évolutions. Un processus de machine learning est aussi performant que les données avec lesquelles on le nourrit ainsi que le modèle qu'on lui applique. Le résultat qu'on tire du processus est aussi pertinent que l'interprétation humaine qu'on en fait.

Samuel Robertson, professeur en science du sport à l'université de Victoria en Australie s'est penché sur le sujet. Le football Australien ou encore le « footy » est un sport collectif qui oppose deux équipes de 18 joueurs sur un terrain en herbe ovale. Ce sport pourrait être résumé (très synthétiquement) par un mélange de football et de rugby. L'équipe du professeur a décidé de développer un algorithme qui va déterminer quelle serait la passe la plus efficace à faire pour un joueur qui a le ballon. Le graphique en haut à gauche est le graphique numéro 1, en haut à droite, le numéro 2, en bas à gauche le numéro 3 et en bas à droite le numéro 4. Ils représentent tous un modèle différent. Le vert foncé représente la zone de passe la plus optimale, le vert clair la moins optimale. Une passe est dite optimale lorsque le risque de perdre la balle/encaisser un but est au minimum et les chances de garder la balle/se rapprocher de marquer un but sont au maximum.



Machine learning in sport : Sam Robertson, Fusion sport, Youtube, 28/10/2018

Le graphique numéro 1 prend seulement en compte la position des joueurs sur le terrain par rapport à leurs buts ainsi que la densité de la zone. Pour le deuxième graphique/modèle, ils ont rajouté la vitesse des joueurs ainsi que l'orientation de leur corps. Dans le troisième graphique, la qualité des joueurs a été ajoutée. Elle est basée sur une batterie de données concernant des données d'anciens matchs (passes réussies lors des 20 derniers matchs par exemple). Enfin, ils ont rajouté pour le dernier modèle le temps qu'allait mettre la balle (pour chaque passe) à aller du joueur A au joueur B. On observe donc que le modèle n'a pas seulement évolué, il a complètement changé selon les variables que l'équipe de scientifiques a choisi de prendre en compte dans le modèle. Lorsque le premier modèle recommande une passe courte en arrière, le dernier lui recommande une passe longue vers l'avant. Le dernier modèle, avec des entrées plus précises et plus pertinentes donne donc des recommandations plus précises et plus efficaces. Cet exemple montre que l'amélioration de l'intelligence artificielle passe obligatoirement par l'amélioration de l'humain qui l'accompagne. Avec une meilleure compréhension du jeu et une meilleure détermination des valeurs pertinentes, des progrès peuvent être effectués sans même qu'il y ait d'innovation technologique.

Les failles actuelles de ces technologies semblent ainsi pouvoir être, au moins partiellement, comblées par les innovations technologiques et l'intervention de l'homme. Mais, au fond, tout cela est-il vraiment souhaitable ?

### **C- L'omniprésence de l'IA, vers un sport déshumanisé ?**

Après tout, on aime le sport avant tout parce qu'il nous procure des émotions. Hors l'émotion est ce qui s'éloigne le plus de la technologie et des statistiques. On ne peut pas quantifier l'émotion, on ne peut pas dire « je suis heureux à 82,9% » alors que l'on peut dire que tel joueur de basket a 82,9% de chance de marquer un panier depuis telle position.

Enlever toute probabilité de perdre ou de gagner, de marquer ou de rater, c'est enlever l'essence même du sport : l'incertitude, la surprise, l'attente ou encore la déception. Si je sais que mon équipe préférée a 99% de chance de finir dans le milieu de tableau, je n'éprouverai presque rien à aller les supporter à chaque match. La notion de plaisir disparaîtrait complètement. Mais avant même d'aborder ces considérations subjectives, interrogeons-nous sur les répercussions pouvant être jugées de façon objective comme « négatives » pour le sport professionnel.

- **L'avènement de l'IA ou le renforcement des inégalités entre petits et gros clubs**

Les inégalités entre grands et petits clubs sont déjà immenses. Dans la plupart des sports collectifs, on peut facilement prédire les dix premiers d'un championnat ou encore réduire à 2 ou 3, le nombre de prétendants au titre. Or, avec l'importance croissante que prend la préparation et inversement la place décroissante que prend le hasard dans le sport, ces inégalités ne feront que croître.

En effet, les investissements financiers qu'impliquent tous ces nouvelles technologies sont conséquents. Les clubs à petit budget n'ont donc pas accès aux meilleurs outils. Les outils qui feront que votre meilleur joueur ne se blessera jamais, que votre tactique soit parfaite ou que vous recruterez le plus gros talent de sa génération avant qu'il n'éclore. Ou comment la technologie est amenée à creuser encore l'écart entre les différentes structures sportives professionnelles.

- **Des investissements lourds : l'exemple du Barça Innovation Hub<sup>40</sup>**

Prenons la toute dernière création d'un des plus grands clubs au monde, le FC Barcelone. Le Barça Innovation Hub est une sorte d'incubateur où se mêlent différentes structures. Il y a bien sûr des start-up mais aussi des universitaires, des étudiants, des représentants de marques, des athlètes, des entrepreneurs... Tout un écosystème qui va travailler ensemble pour découvrir de nouvelles choses, apporter de nouvelles connaissances et créer des nouveaux produits ou services. Tout cela dans différents domaines : le sport, la santé, la nutrition et plus généralement tous les aspects liés aux sports qui pourraient avoir un impact sur notre société.

Le développement de cet institut a engrangé un investissement financier faramineux que seul un grand club comme le FC Barcelone peut se permettre. Quel est donc le rapport avec les inégalités entre petits et grands clubs ? Sans surprise, le « Barça » sera le premier club à pouvoir tester toutes les nouveautés et pour de ce fait prendre un avantage conséquent sur ses concurrents. Le club aura grâce aux produits et services développés, des joueurs plus performants, un club mieux organisé, une expérience des fans au stade améliorée : tout simplement un meilleur club. Il est évident que des clubs dont le budget annuel ne dépasse pas 50 millions d'euros seront en incapacité d'investir dans des projets aussi importants et coûteux. Dans une période où, malgré la croissance des inégalités, l'écart entre les plus gros clubs diminue, ce sont au finals des petits détails qui feront la différence. C'est justement grâce aux

---

<sup>40</sup> Le Barça Innovation Hub, quand le football commande l'innovation, Rudy Turinay, Forbes, 13/03/2019

avancées technologiques que les clubs et joueurs gagnent en performance. Tout cela dans le but d'obtenir les informations et statistiques qui feront cette petite différence et qui les feront gagner. Le fossé entre petits et grands clubs est déjà énorme. Est-il souhaitable que l'amélioration de la performance de quelques-uns, d'une certaine élite, se fasse en creusant l'écart avec une majorité moins bien lotie ?

- **Vers un sport à deux vitesses ?**

L'autre raison pour laquelle nous aimons les sports de hauts niveaux est que nous pouvons nous identifier facilement à ce que nous voyons lorsque l'on va voir un match. En effet, que ce soit au rugby ou au basket par exemple, il y a x joueurs sur la pelouse, 2 paniers et un arbitre. C'est aussi là où nous prenons du plaisir. Pouvoir nous aussi, les week-ends entre amis ou en championnat amateur, marquer autant de buts que Sergio Aguero ou autant de trois points que Stephen Curry. Pouvoir râler après l'arbitre, avoir de la chance sur un penalty non sifflé ou une main oubliée. Nous apprécions cela en grande partie car cela ressemble exactement à ce que nous voyons à la télé ou dans les stades et nous sommes fiers de pouvoir reproduire ces choses-là à l'échelle amateur.

Ce schéma est facile à reproduire et c'est d'ailleurs ce qui se fait aujourd'hui. Du moins, ce qui se faisait. Avec la Goal Line Technology (technologie qui détermine si la balle a franchi ou non la ligne de but), la VAR (assistance vidéo pour revenir sur des décisions arbitrales importantes), les rencontres sportives professionnelles n'ont plus grand-chose à voir avec le sport collectif amateur. Les ligues ou clubs locaux n'ont clairement pas la capacité financière pour investir dans de telles technologies et nous sommes déjà, en 2019, à l'aube d'un sport à deux vitesses. La question qu'il faut alors se poser est la suivante : souhaitons-nous vraiment enlever le côté populaire et accessible des sports collectifs au profit d'un sport « parfait » et donc déconnecté de la société ?

- **Vers un sport déshumanisé même au stade ?**

Lorsque l'on parle de déshumanisation du sport, on ne peut omettre la nouvelle façon de vivre les matchs. Les stades du Real Madrid ou encore du FC Séville ont depuis quelques années des caméras partout. Tout cela pour avoir les meilleurs angles de vue, les ralentis les plus précis et encore plus d'accès à ce qui se passe sur la pelouse.



En dehors de voir des prouesses physiques ou techniques ou encore son équipe préférée, qu'est-ce que le supporter vient chercher ? Il souhaite pouvoir échanger avec les autres supporters, vivre une expérience collective. C'est l'essence même des supporters, faire et refaire les matchs. Donner son avis, ses informations et souvent, c'est le mieux informé qui « remporte » le débat. C'est cette envie que les clubs essayent de satisfaire avec encore plus de données et d'informations disponibles en temps réel sur leur application. Nous vivons déjà dans une époque où les smartphones sont omniprésents dans nos quotidiens. Les lieux de spectacles font partie des rares endroits où nous « acceptons » de poser notre téléphone. Les dernières innovations des clubs portent à croire que cela va changer. Le Wifi dans les stades, les ralentis sur leur téléphone... Quand les supporters trouveront-ils le temps de regarder directement le match, discuter avec son voisin de tribune ou chambrer les supporters adverses ?

Au-delà de cela, La Liga (la ligue de football espagnole) est en train de tester avec Samsung Gear<sup>41</sup> l'utilisation de la réalité virtuelle et augmentée pour vivre le match. Bientôt nous n'aurons plus à nous déplacer pour voir de manière unique un match. Ce sera même encore « mieux » depuis notre salon. Cela réduirait bien sûr l'affluence dans les stades mais aussi et surtout bafouerait une des valeurs les plus nobles du sport : le rassemblement. Les rencontres sportives rassemblent des individualités de tous les milieux sociaux et culturels contribuant grandement à la fameuse « beauté du sport ». Cet aspect humain, cette dimension universelle, semble aujourd'hui être mise en danger par la volonté de proposer aux spectateurs une expérience toujours plus « personnalisée », quitte à se couper des autres.

### **CONCLUSION PARTIE 3 : Un mariage et un équilibre à trouver entre la technologie et l'homme**

Nous avons vu dans cette dernière partie que dans un monde où la statistique est clé, elle n'est pas encore reine. En effet, les avancées actuelles montrent bien des lacunes et cela dans plusieurs domaines et plusieurs sports.

L'humain reste au centre du sport et même si la technologie et les statistiques sont une aide à la décision et à la préparation incroyable, rien ne remplacera l'irrationalité de l'homme. Un scientifique aura beau vous dire que si vous sautez de 10 mètres de hauteur sur un pouf géant, vous avez 99,9% de chances d'atterrir intact, vous allez y réfléchir à deux fois avant de sauter. Pour le sport c'est la même chose. Que ce soit pour les penaltys au football, la prédiction du

---

<sup>41</sup> Artificial intelligence in sport – current and future application, Kumba Senaar, Emerj, 31/01/2019

vainqueur d'un match ou d'une compétition ou encore les fameux « expected goals », on peut collecter et analyser une pléthore de données sans avoir raison. Trop de facteurs incalculables et ininterprétables (ou du moins difficile à implémenter dans un modèle) rentrent en jeu dans les sports collectifs professionnels. L'humeur du joueur, l'adrénaline qu'il ressent, la chance ou la réussite, l'intervention irrégulière de facteurs extérieurs comme les spectateurs ou une bourrasque de vent. L'aspect esthétique d'une action qui va motiver encore plus l'équipe ou un Buzzer Beater (panier à la dernière seconde au basket) qui va pousser l'équipe dans une dynamique positive sont autant de facteurs qui vont contribuer à déjouer les pronostics et qui sont pourtant immuables dans le monde du sport. Les émotions prennent beaucoup trop de place chez l'humain pour que cela ne se transcrive pas dans le sport et c'est bien là que le bât blesse pour la technologie.

Du moins c'est ce que nous avons tendance à penser avant certaines avancées. Plus la technologie et le sport évoluent ensemble plus des découvertes se font. Les données qui paraissaient pertinentes il y a quelques années ne le sont plus forcément et on arrive désormais à faire des miracles. On parvient à trouver les bonnes technologies pour les bonnes applications et amener aux intelligences artificielles la touche humaine nécessaire pour bien comprendre et analyser les données collectées. Grâce au machine learning ou encore au deep learning, les équipes de scientifiques et leurs technologies parviennent désormais à aller encore plus loin. Avec les bonnes entrées, les recommandations de passes ou de schémas de jeu sont encore plus précises et plus performantes car on y modélise maintenant les bons facteurs. On parvient aussi à décrypter les émotions des joueurs et prédire le résultat du prochain point au tennis par exemple avec une précision étonnante. Les applications sont nombreuses et s'étendent de la reconnaissance faciale pour rapidement analyser un public à la santé pour déterminer la douleur d'un joueur par exemple. Autant d'applications qui laissent penser que finalement, la technologie et l'humain ne cessent de s'améliorer pour combler encore et toujours les lacunes pointées du doigt.

Malgré tout cela, en y réfléchissant profondément, est-ce vraiment une bonne idée ? Est-ce que nous souhaitons vraiment pouvoir prédire le résultat d'un match avec une marge d'erreur de 2% ? Est-ce que nous souhaitons vraiment savoir que quand Rafael Nadal fronce les sourcils, il va obligatoirement rater le prochain coup ? Finalement, est-ce que nous voulons sincèrement supprimer le suspense, la chose qui nous excite le plus dans le sport ? La réponse est selon nous non. Le fan veut toujours pouvoir avoir la chance d'assister à une réédition de David contre

Goliath. Il veut continuer à pouvoir voir l'équipe de France de basket battre les Etats-Unis en quart de final de coupe du monde, Montpellier remporter la ligue 1 avec le onzième budget de l'hexagone ou encore le Japon battre tour à tour l'Ecosse et l'Irlande en phase finale de coupe du monde de Rugby. Au-delà de cela, un supporter vient au stade pour vibrer et sentir qu'il appartient à une communauté. L'idée d'être ensemble est centrale, l'idée de voir un ciseau retourné de ses propres yeux est exaltante, l'idée d'entendre tout un stade chanter l'hymne national de son pays est vibrante. Or, avec la direction des avancées technologiques, nous serons bientôt en mesure de vivre le match sans y être, de regarder le match au stade à travers un écran. Ce sont autant d'évolutions qui mettent en danger la beauté offerte par un spectacle sportif.

## CONCLUSION

Ayant débattu pendant de longues pages de l'impact du Big Data sur les sports collectifs professionnels, nous nous devons de conclure en donnant un avis à la fois objectif et subjectif sur la question. Après avoir épluché des centaines de documents et autres pages internet, nous avons une vue et une compréhension à la fois granulaire et globale. Les quatre prochaines prédictions sont pour nous clés dans l'évolution des sports collectifs de hauts niveaux. Elles sont basées sur nos recherches mais aussi sur notre intuition car finalement, qu'est-ce que l'on ferait sans notre instinct ?

### **Première prédiction : *L'intelligence artificielle va petit à petit prendre le pas sur les staffs techniques***

Si aujourd'hui l'intelligence artificielle est un outil, une aide à la décision pour les staffs techniques, ceux-ci pourraient peu à peu lui laisser la place.

Premièrement sur le plan tactique : à court et moyen terme, les coachs se verront suggérer directement par la machine des formations tactiques et des compositions d'équipes avant les matches et des remplacements pendant la rencontre mais garderont le dernier mot. A plus long terme, ces choix pourraient se faire automatiquement sans validation par un être humain, l'entraîneur se contentant d'être un manager/meneur d'homme, une incarnation physique du logiciel. On peut même imaginer un club qui tente le pari de ne plus avoir d'entraîneur du tout, pourquoi pas sponsorisé par un éditeur de logiciel souhaitant valoriser son savoir-faire avec potentiellement un énorme coup de communication à la clé.

Même phénomène concernant l'entraînement et le recrutement. On peut imaginer un logiciel qui compose une séance d'entraînement avec des exercices physiques individuels déterminés en fonction des données physiologiques collectées et d'exercices collectifs visant à corriger les points faibles observés au cours des rencontres précédentes. Concernant le recrutement, domaine dans lequel le Big Data est déjà très avancé, on pourrait imaginer que l'intelligence artificielle permette à un programme d'identifier les points faibles constatés au cours de la saison, d'anticiper le futur déclin ou non des joueurs grâce à l'analyse prédictive, et de déterminer les profils les plus à même de renforcer l'effectif disponibles sur le marché. Les

recruteurs deviendraient alors de simples négociateurs, de simples commerciaux au faible pouvoir décisionnaire.

Enfin, les staffs techniques risquent également de perdre du pouvoir au profit des supporters. Nous l'avons évoqué auparavant, ceux-ci réclament de plus en plus d'implication dans la vie de leur club, clubs pour qui l'enjeu économique est de plus en plus crucial. Dès lors, le modèle d'United Managers semble représenter la solution idéale, les supporters devant payer pour pouvoir avoir le droit d'influer sur la composition et la tactique de l'équipe. Nous pensons donc que, si ce modèle ne va pas forcément devenir la norme, il coexistera avec un modèle plus traditionnel, et ce partout dans le monde et dans diverses disciplines. Évidemment, cela nécessitera une modification des règlements par les instances sportives mais l'histoire récente montre que celles-ci peuvent rapidement s'adapter lorsque les enjeux économiques sont importants.

**Deuxième prédiction : *Le sport aura toujours une part d'irrationnel, d'imprévisible car il dépend in fine de facteur humain***

Les sports collectifs sont joués par des êtres humains. Le propre de l'humain est qu'il est imprévisible et irrationnel. Comme établi dans le corps de ce mémoire, c'est ce qui rend le sport attractif. On aime aller voir un match pour les émotions qu'il nous procure.

On peut prévoir l'état de fatigue d'un joueur à la 65<sup>ème</sup> minute. En revanche on ne peut pas prévoir l'adrénaline qui fera qu'il dira à son coach « Non je suis en forme aujourd'hui, j'ai encore 20 minutes dans les jambes ». On peut prédire qu'une équipe devrait écraser l'autre à l'aide de son meilleur joueur mais on ne peut pas prévoir qu'il prenne un carton rouge pour un geste d'énervement. On peut tout miser sur un joueur de 18 ans aux talent prometteur mais on ne peut pas (vraiment) contrôler ce qu'il fait en dehors du terrain et qui va participer à nuire à son développement.

La révolution technologique dans le sport et selon les avancées en est encore à différents stades de son évolution. Le VAR par exemple connaît encore de vives contestations et n'est pas adopté par tous car la décision finale appartient encore et toujours à l'humain. L'humain étant imparfait, il y a et aura toujours des erreurs.

### **Troisième prédiction : *Un inévitable encadrement financier et réglementaire des investissements technologiques ou bien une scission du milieu sportif*<sup>42</sup>**

Comme nous l'avons évoqué dans ce document, les avancées technologiques provoquent une accélération de l'écart entre les structures professionnelles disposant de moyens importants et les celles au budget plus limité, mais aussi entre le monde professionnel et le monde amateur. Une solution pour limiter ce phénomène pour les instances sportives internationales seraient d'encadrer le recours à ces technologies. La solution peut venir de l'encadrement financier des investissements technologiques notamment.

Il y a pour l'instant la règle du fair-play financier qui vise à limiter les dépenses d'un club par rapport à l'argent qu'il génère. Cette règle a été créée dans le but d'avoir des clubs aux finances saines mais aussi pour limiter le pouvoir des clubs les plus riches. Cependant, si l'UEFA est particulièrement pointilleuse pour le montant que les clubs investissent dans les joueurs, les investissements structurels du club ne sont pas pris en compte. Il en va de même pour le salary cap dans les sports américains. A l'avenir, on peut imaginer que les règlements prennent en compte cet aspect-là.

On peut prendre exemple sur la Formule 1 ou plus que dans n'importe quelle discipline, les avancées technologies sont reines. Ce sport automobile souffre depuis longtemps de nombreuses critiques liées à son côté prévisible. Les écuries les plus riches avec les meilleures outils et monoplaces gagneront toujours. Cette année encore, les plus grosses écuries que sont Mercedes, Ferrari et Red Bull finiront première, deuxième et troisième au classement constructeur. Pour y remédier, la FIA fédération internationale de l'automobile a décidé de mettre en place de nouvelles régulations pour 2021. Elles comprennent entre autres un volet financier. En effet, pour limiter le pouvoir des grosses écuries à développer de meilleures voitures, une limite annuelle de 154 millions d'euros a été fixée pour toutes les écuries.

Cela passe aussi par des règlements sur les technologies utilisées : la Formule 1 est ainsi connue pour interdire très rapidement toute innovation technologique procurant à une écurie un avantage trop conséquent sur la concurrence.

---

<sup>42</sup> FIA.com

Cet encadrement via de nouveaux règlements pourrait contribuer à limiter le fossé entre les équipes mais aussi à calmer le mécontentement croissant d'une partie du public face à ces irruptions technologiques susceptibles de dénaturer le sport. Ce phénomène pourrait passer par des règlements légers, facilitant l'implémentation progressive et contrôlée des technologies et donc de leur acceptation par le public. Cela pourrait également se faire de façon plus brutale avec des instances qui décident de « revenir en arrière » en limitant l'accès à la technologie pour revenir aux racines du sport (exemple du cyclisme où l'on débat de la possibilité de supprimer les oreillettes des coureurs et les compteurs de watt sur les vélos depuis de nombreuses années).

Si les instances ne font pas ce travail par elles-mêmes, nous sommes persuadés que nous assisterons à la création de championnats et structures sportives parallèles par les acteurs qui rejettent cette vision du sport (joueurs, clubs, spectateurs...). Cela a déjà été le cas dans le passé avec l'exemple de la création du Football Club United de Manchester créée par des fans en désaccord avec l'acquisition de Manchester United par la famille Glazer (hommes d'affaires américains) ou celle de l'AFC Wimbledon pour des motifs similaires : le refus de la prise de pouvoir de l'aspect financier sur le sportif, l'humain et les traditions. De la même manière, on peut imaginer une véritable scission entre un modèle sportif où règne l'innovation technologique et un autre plus traditionnel, rustre et idéalisé comme plus « pur » par ses fervents partisans.

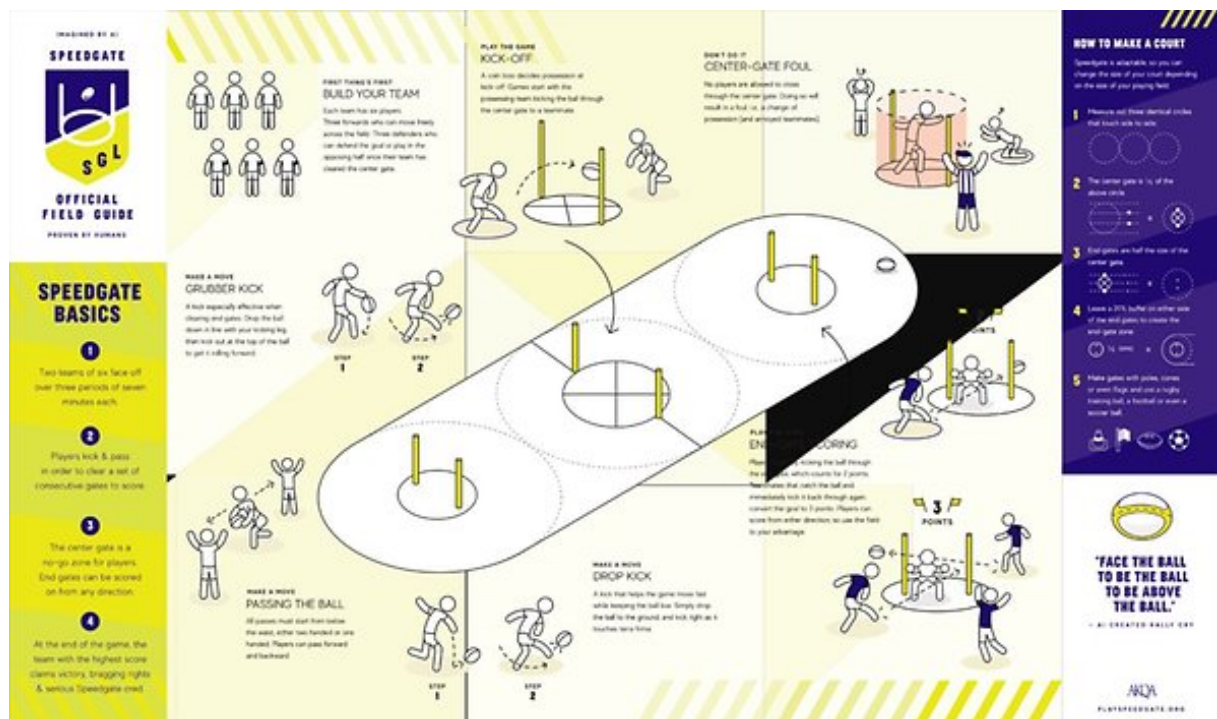
#### **Quatrième prédiction : l'IA va créer des nouveaux sports<sup>43</sup>**

Cela fait maintenant des dizaines et des dizaines d'années que nous jouons et suivons les mêmes sports. Que ce soit le basket-ball, le football ou encore le handball, tous ces sports existent depuis bien longtemps. Nous continuons à aimer ces sports car ils sont ancrés profondément dans notre société et aussi parce qu'ils présentent un mélange équilibré entre plusieurs facteurs très importants pour subsister : des règles assez complexes mais assez simple pour être comprises et appliquées, un côté ludique, amusant, compétitif avec souvent une grande marge de progression.

---

<sup>43</sup> Speedgate, le premier sport inventé par une intelligence artificielle, Fabrice Auclert, Futura Tech, 15/04/2019

L'agence de design AQKA a décidé de faire ce qu'aucun humain a pleinement réussi durant ces dernières décennies : inventer un sport. Ils ont fourni à une intelligence artificielle toutes les données relevées pour 400 sports différents. Plus précisément, un réseau de neurones artificiels l'a fait : le résultat n'est pas immédiatement au rendez-vous. Plusieurs sports aussi improbables qu'étranges ont été proposés par l'intelligence. On peut noter entre autre le frisbee explosif. Après quelques ajustements et une sélection de la part de l'équipe AQKA, le Speedgate a vu le jour. Deux équipes de 6 joueurs s'affrontent sur un terrain cylindrique (55 sur 18 mètres) avec un ballon similaire à celui utilisé dans le rugby. Il y a trois zones distinctes : la zone autour de la porte centrale où personne ne peut se trouver et 2 « portes » des deux côtés du terrain. Chaque équipe a trois défenseurs et trois attaquants répartis en attaque et en défense. L'objectif est de marquer dans la porte adverse avec une frappe du pied mais avant, il est impératif d'avoir fait une passe du pied à travers la porte centrale. La balle ou le joueur ne peut pas rester immobile pendant plus de 3 secondes, un but vaut 2 points et aucun contact n'est toléré.



Speedgate: une IA s'inspire de 400 sports pour en créer un, Bastien L, Le Big Data, 15/04/2019

Le projet a été présenté à l'Oregon Sports Authority dans l'espoir de créer une ligue dès cette année. Ce nouveau sport espère créer un réel engouement et tente d'encourager le public à être le moteur de la création de ligues autour du monde. Ce nouveau sport a tout pour plaire : une dimension physique car on ne peut rester immobile, une interdiction de contact qui attirera les



personnes craignant les blessures, une marge de progression collective et tactique importante ainsi qu'un aspect technique non négligeable.

Si l'idée peut faire rire sur le papier et que l'engouement populaire n'est pas encore au rendez-vous, on peut imaginer qu'à long terme, l'intelligence artificielle soit un réel outil pour créer de nouveaux sports en identifiant les variables qui plaisent aux spectateurs. De la même façon que Netflix écrit aujourd'hui les scénarios de ses contenus originaux en fonction de l'analyse des données collectées sur la consommation de contenus vidéo de ses abonnés.

## SOURCES

### Articles de presse (papier ou internet) :

- *Sport/économie : le sport a un poids économique « indéniable »*, Jean Damien Lesay, Banque des Territoires, 02/10/2018
- *Germany's Secret World Cup Weapon: Big Data*, Jack Rosenberger, CIO Insights, 18/07/2014
- *Honigstein chats with Oliver Bierhoff*, Raphael Honigstein, ESPN, 08/07/2014
- *Big Data: A Game Changer For Germany In The World Cup*, Mike Sandiford, JobstheWorld, 25/07/2014
- *Rio 2016: Jeux olympiques, Big Data et médailles d'or*, Bastien L, Lebigdata.fr, 11/08/2016
- *Driving success*, SAS, 02/2015, p 30-31
- *The science behind Chris Froome and Team Sky's Tour de France preparations*, João Medeiros, Wired, 30/06/2016,
- *How data analytics killed the Premier League's long ball game*, João Medeiros, Wired, 01/2014
- *Everton FC : Football Manager 09 aide aux transferts*, Guillaume Nerces Tutundjian, 17/11/2008
- *Comment football manager est devenu un vrai outil pour les clubs*, Alix Dulac, RMC Sport, 28/11/2015
- *Jonathan Beilin : « Football Manager était toujours allumé »*, Kévin Charnay, So Foot, 12/11/2015
- *How Football Manager's scouting network unearths the wonderkids of real world football*, JJ Bull, The Telegraph, 02/11/2018
- *On est allé chez Wyscout, leader mondial du scouting*, Valentin Pauluzzi, So Foot, 28/07/2016
- *On a testé Wyscout*, Valentin Pauluzzi, So Foot, 28/07/2016
- *Sur le mercato, la révolution data*, Lionel Pittet, Le Temps, 21/06/2017
- *Stade connecté, une expérience ultime des fan*, Cabestan, emarketing.fr, 27/06/2018
- *Digital, un allié puissant du sport pour une fan expérience réussie*, Fabien Chan, MCI, 14/01/2019
- *Comment vivra-t-on un match de foot en 2030 ?*, Clément Pons, Vice, 08/01/2018

- *Five high value fan segments for sports teams*, Tal Vinnik, MVP Audience, 03/05/2018
- *Olympique Lyonnais, une activité e-commerce en très forte croissance*, Anthony Alyce, EcoFoot, 06/09/2017
- *Le Yield Management adapté aux événements culturels et sportifs*, Yurplan, 10/05/2017
- *Overview of behavior based fan insight*, IBM
- *Big Data et sport: de la donnée à la performance*, William Paris, Portail de l'IE, 25/04/2017
- *7 questions à Nicolas Chanavat pour sports marketing*, Sports Marketing, 17/03/2016
- *Comment utiliser la technologie pour mieux interagir avec ses fans ?*, Victor Lamm, Microsoft, 06/2018
- *How can data help drive sports sponsorship*, Pete Giorgio, Deloitte, 2018
- *Five trends for 2017 in sports sponsorship*, Global Sports, 10/01/2017
- *The benefits of sports sponsorship in the digital age of visual data*, Logo Grab, 03/02/2017
- *Les données, chance marketing des clubs*, Florent Barraco, emarketing.fr, 05/2016
- *The transition of the sports industry and its implication for sport sponsorships*, Kinexon, 2018
- *How digital is transforming sports sponsorship*, Joe Weston, We Are Social, 15/05/2018
- *Le PSG lance la publicité virtuelle*, Alexandre Bailleul, SBB, 25/04/2018
- *Qu'est-ce que la publicité virtuelle, nouvelle technique développée en Bundesliga*, SBM, 05/04/2018
- *Business analytics, le nouveau relais de croissance des clubs ?* Charles Salvanet, ecofoot.fr, 31/01/2019
- *Les expected goals au cœur de la révolution statistique*, Julien Assuncao, Le Cahier du Football, 01/04/2015
- *Mais que sont les expected goals « xG » la statistique du moment*, Johann Crochet, Goal, 13/11/2017
- *L'euro 2016 a été compliqué pour les robots chargés des pronos*, Gregor Brandy, Slate, 12/07/2016
- *Google calcule que la France va battre l'Allemagne*, Julien Lausson, Numerama, 04/07/2014

- *Coupe du monde 2018, une IA prédit le parcours de chaque équipe et ... le vainqueur*, Julien Cadot, Numerama, 12/06/2018
- *The big risks of Big Data in sport*, Bernard Marr, Forbes, 2017
- *Le Barça Innovation Hub, quand le football commande l'innovation*, Rudy Turinay, Forbes, 13/03/2019
- *Bienvenue dans l'ère du football high-tech*, Philotée, Yellow Vision, 18/06/2018
- *How is machine learning changing sport*, William Tubbs, Innovation Enterprise, 05/06/2019
- *Big Data/Machine learning et sport*, Scient, 07/08/2019
- *How the rise of machine learning is impacting sport*, Craig Pickering, SimpliFaster, 2019
- *Artificial intelligence in sport – current and future application*, Kumba Senaar, Emerj, 31/01/2019
- *Five ways facial recognition is being used in sport*, Jessy Davis, Face First, 25/07/2019
- *Sports teams are using facial recognition to learn more about their fan base*, Eric Chemi, CNBC, 21/04/2018
- *Facial recognition technology comes to tennis to track player emotions*, John Stensholt, Financial review, 19/01/2018
- *Speedgate: une IA s'inspire de 400 sports pour en créer un*, Bastien L, Le Big Data, 15/04/2019
- *Speedgate, le premier sport inventé par une intelligence artificielle*, Fabrice Auclert, Futura Tech, 15/04/2019
- *FIA.com*
- *Artificial intelligence in sport – current and future application*, Kumba Senaar, Emerj, 31/01/2019

**Livres :**

- *The Sabermetric Manifesto*, David Grabiner, 1994
- *Moneyball: The Art of Winning an Unfair Game*, Michael Lewis, W. W. Norton & Company Inc, 2003
- *Game Changers: How a Team of Underdogs and Scientists Discovered What it Takes to Win*, João Medeiros, 2018
- *A quoi rêvent les algorithmes, nos vie à l'heure des Big Data*, Dominique Cardon, 01/10/2015
- *Penser ou cliquer*, Michel Blay, 2016

**Articles académiques:**

- *L'influence du football en Chine*, Thomas Lévêque, IESC Angers, 2017
- *Big Data and tactical analysis in elite soccer: future challenges and opportunities for sports science*, Robert Rein & Daniel Memmert, SpringerPlus, 24/08/2016
- *Match Analysis, Big Data and Tactics: Current Trends in Elite Soccer*, Robert Rein & Daniel Memmert, German Journal of Sports Medecine, 03/2018
- *Les Nouvelles techniques de billetterie pour augmenter les revenus des clubs professionnels de football en France*, Pascal Perri, Université de Bordeaux, 07/2017
- *Action bias among elite soccer goalkeepers: The case of penalty kicks*, Michael Bar-Eli & Ofer H. Azar & Ilana Ritov & Yael Keidar-Levin & Galit Schein, Journal of Economic Psychology, 10/2007
- *Penalty kicks in soccer: an empirical analysis of shooting strategies and goalkeepers' preferences*, Michael Bar-Eli & Ofer H. Azar, Soccer & Society, 02/2009
- *A machine learning framework for sport result prediction*, Rory P. Bunker, Fady Thabtah, 09/2017

### Sites Internets :

- Page Wikipedia: Sabermétrie
- Page Wikipedia: Moneyball: The Art of Winning an Unfair Game
- Stats.com
- Wyscout.com
- Transfermarkt.fr
- Secondspectrum.com
- Playspeedgate.org

### Films et documentaires vidéos

- *Dix ans de révolution – Cyclisme – Team Sky*, Chaîne YouTube L'équipe, 09/03/2019
- *Les Bleus 2018, au cœur de l'épopée russe*, TF1, 17/07/2018
- *How analytics have changed scouting in football*, Vishnu Anandraj, Chaîne Youtube Tifo Football, 13/01/2019
- *How data transformed the NBA*, Chaîne Youtube The Economist, 04/12/2018
- *United Managers : les supporters aux commandes*, Canal Football Club, Octobre 2018
- *Football : comment la peur du ridicule explique la plupart des penaltys ratés (Datafoot Ep. 1)*, Chaîne Youtube Le Monde, 04/06/2018
- *Que désignent les expected goals*, Wiloo, Youtube, 03/2018
- *Machine learning in sport : Sam Robertson*, Fusion sport, Youtube, 28/10/2018

### Podcast :

- *The Bradley Wiggins Show*, podcast du 27/05/2019

### Autres:

- *Modifications aux textes fédéraux adoptées lors de l'Assemblée Fédérale du 8 décembre 2018*, Fédération Française de Football, 10/12/2018

