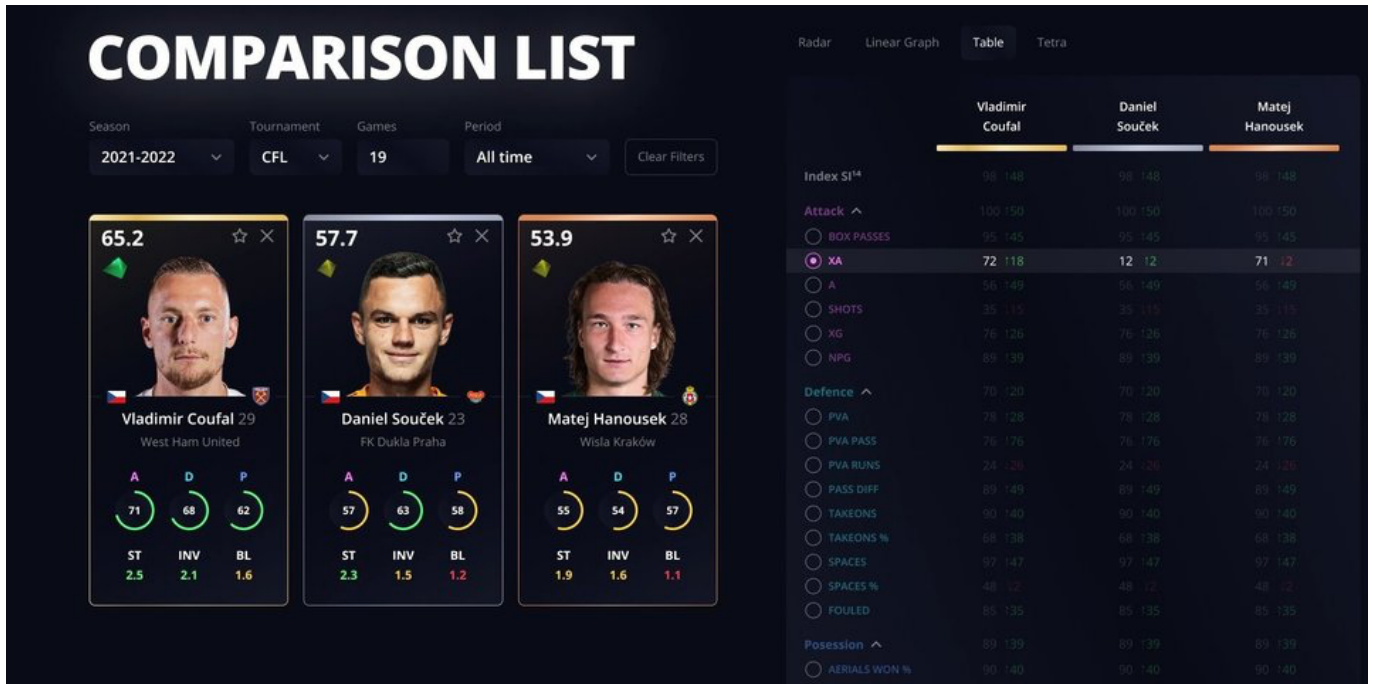# AN EX-FOOTBALL PLAYER HAS FIGURED OUT HOW TO EVALUATE THE STRENGTH OF ANY PLAYER IN THE WORLD. AND PICK THEM ON THE MARKET LIKE A TOMATO

Alexei Tryputen played for CSKA and Moskva and was a sports journalist. He now heads the analytics department of a company that offers some curious digital innovations.
In spring, Match TV already wrote about a new approach to soccer statistics, thanks to which the company represented by Tryputen won the Russian prize as Project of the Year. They began with introducing robotic trading systems for inefficient markets and came up with neurofunds, artificial intelligence, and other unexpected things for soccer.

- *Why do you need a player evaluation index?*
- *How revealing is it?*
- *What are the drawbacks of other such criteria?*
- *How is artificial intelligence different from a calculator?*
- *How do you get an idea of a player's usefulness from one championship to another?*
- *What do the insurance business and artificial vision have to do with it?*
- *Not only does the coach evaluate players in numbers, but the president of the coach as well - is this realistic?*

- Since then, we have a lot of news, - says Triputen. - We have finished working on one of the most curious products - the player evaluation index. There have been attempts to come up with something like that before. It's interesting to understand without digging into piles of numbers whether a player is good in general and in specific circumstances, in particular. **What his strengths and weaknesses are. How compatible he is with this or that tactic, league, team. How the progress or regression of players is influenced by coaches. What happens to players after they change clubs, injuries, and even weddings.**

- Spartak wouldn't mind that: they learn about the professional abilities of players and coaches after they sign the contract.

- Now they can find out before. And not just to find out, but to see and almost feel in their hands.

- What do you mean?

- The Si14 index isn't just a number. It allows you to get a visual idea of the player's class and condition. Below, I'll tell you how it works. And I'll show you.

- How is it different from the indexes developed before you?

- There aren't many of them. But all have errors, albeit of a different nature. One of the statistical sites offers a grade similar to the school grade, only given on a ten-point scale—pure taste. An American media group has come up with ratings for clubs and national teams to help predict the outcome of tournaments based on previous performances. It seems fair, but clubs with small budgets, and these are the majority, renew the roster every year. The newly assembled team has nothing to do with the previous seasons' results; its evaluation is irrelevant.

One well-known statistical company has developed a total utility index that reflects a player's efficiency. A player's efficiency is calculated based on a linear formula: did this, plus this. This approach is very primitive and close to a finger in the sky. Our story is about something else entirely.

First, a few words about the team we assembled specifically for this task. Realizing the priority of the mathematical approach, we were looking for superprofessionals, not in the least bit interested in their soccer skills. For example, the person who developed the artificial vision program was responsible for communicating with artificial intelligence. It's like in the movies about the future, when you get into the lens of a mall camera and after a few seconds, you see an advertisement addressed exclusively to you.

A guru from the insurance business worked with mathematical models. Thanks to Andrei's experience in calculating insurance risks, our methodology became coherent. Of course, we did not do without specialized soccer analysts: the artificial intelligence had to prove our mega brain Yura right. In the end, through our joint efforts, we brought this story to a beautiful outcome.

- What is it?

- So, our artificial intelligence is responsible for the calculations. The foundation is an array of data that we buy from several leaders in the statistical industry. The latest computer programs allow us to decompose a player's performance into molecules and compare him with his competitors in terms of position. Every player's movement is recorded and accounted for by artificial intelligence in the creation of statistical models. Any new information makes the models more accurate and, at the same time, allows you to see the deviation of the player from the reference values. In numerical terms, this is the Si14 index. It can be calculated for any player in the world.
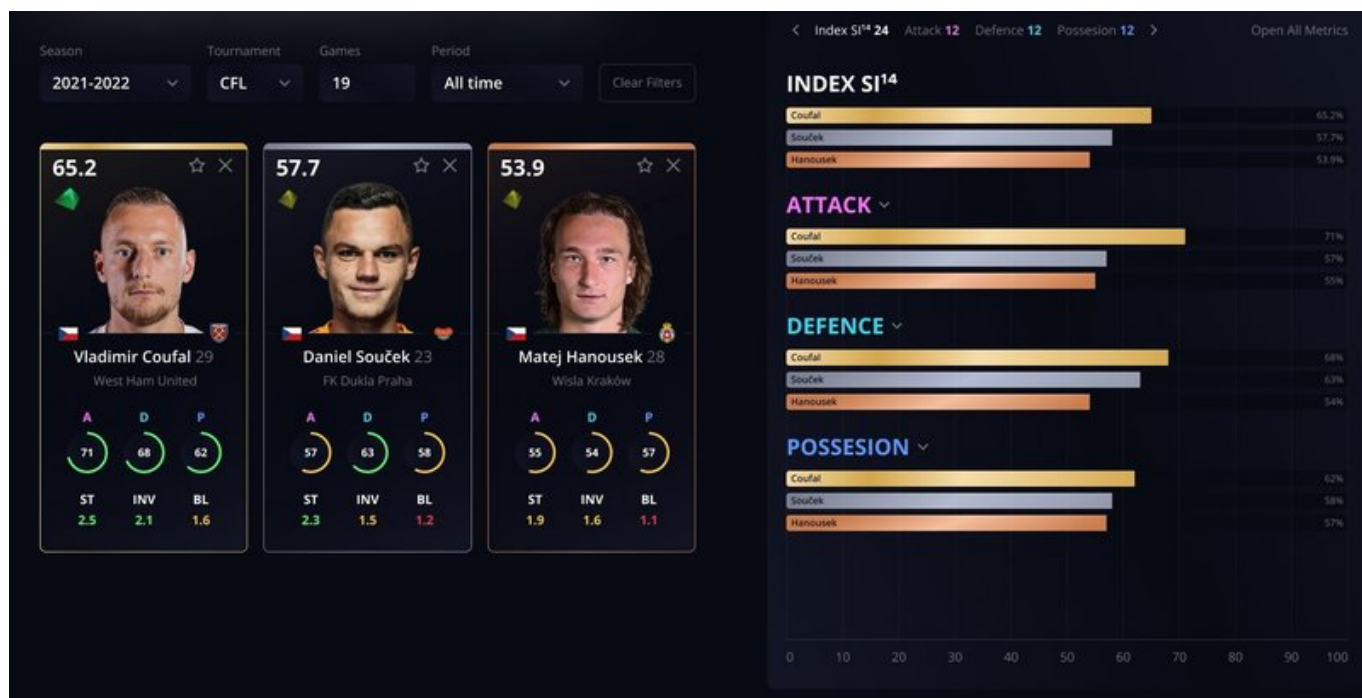
- Do your computers count faster and more than others?

- Our computers count differently than others. They do not add up the numbers - they describe the players' playing qualities by complicated non-linear curves. And then, for clarity, they translate it into digital and graphic images that are easy to use.

Artificial Intelligence knows what a model Belgian winger or a Danish league goalkeeper looks like before calculating a particular player. This is because the molds are precise - they are created based on the millions of data points that come into our daily processing. Comparing a player to a template characterizes his level. And because the Si14 index can be calculated for any period, at least for a half or a year, it also shows the dynamics of a player's level. We can see if the player is declining or gaining. And after what happens. And in what elements of the game. And what he will be able to do in the short term.

- Are you sure there is one number to describe players in different areas?

- First of all, the number is not the same. The metrics of players are divided into three groups: defensive, offensive, and on-ball actions. Each group is divided into subgroups because you can play in different parts of the field, and it will have a different effect on the result. The same applies to types of transmissions, defender connections, pressing actions, everything.

Lines are shaped not by piling everything up, but on the contrary, by full detail. It's like chess, where computers learned to play based on a) a massive number of calculated combinations and b) the previous experience of all the chess players in the world. Artificial intelligence doesn't think, but it knows how others think, and so it takes the shortest route to the goal. Our index is built on a similar idea: if you know

how it was, you will understand how it should be.

- What's the second thing?

- First of all, we identified a list of metrics that have the most significant impact on a player's efficiency. That is, not us, but the computer, of course. Imagine that somewhere, they calculate a defender's passes across the field with a factor of 0.04. They decided that the contribution of this action to the player's TTC is 4%. But why 4? What was the reasoning behind that? Most likely, it was taken out of my head, by eye, by instinct.

Let me remind you, the human factor is excluded from the calculations. The computer ran a lot of data through itself and determined those metrics of the player on which his usefulness depends. It also calculated the degree of dependence. We looked it over, cleaned something up, compared it to common sense and logic, with live examples. And we instructed artificial intelligence to use only mathematically justified parameters to calculate it.

- I wonder what you cleaned up.

- The computer found out that the quality of a player is not affected by the number of warnings he earns, his ability to shoot from the 11-meter mark, shots on the bar, or the crossbar. All of this is there, but out of connection with anything, logical context. Then why count?

- You've equipped artificial intelligence with the metrics necessary to calculate. How is it further different from an ordinary calculator?

The decision tree is a methodology used in other fields of activity, such as the insurance business. I'll repeat it: our method is not about multiplying and adding; the computer takes a different approach. To put it crudely, it is a sifting of players through a set of references that is indicative metrics. How many goals did player X score in the last game? Two. Goes into one basket. Player Y? Zero. The other. Player Z? One.

Third. Next fork: How many sharpening passes did X make? Three. Follow higher up the tree. Y none? Lower. And so on.

There are dozens of metrics ranked by importance. As a result, imaginary curves are formed for each player. And any of them will adjust the median values at least a bit. There is such a notion in physics - quantization, which allows describing a smooth line by a set of small concessions. Of course, they will never coincide completely, but the shallower the ledges, the more invisible the difference. So it is with us: the more data we have to analyze, the more reliable the pattern of player evaluation. And hence our index.
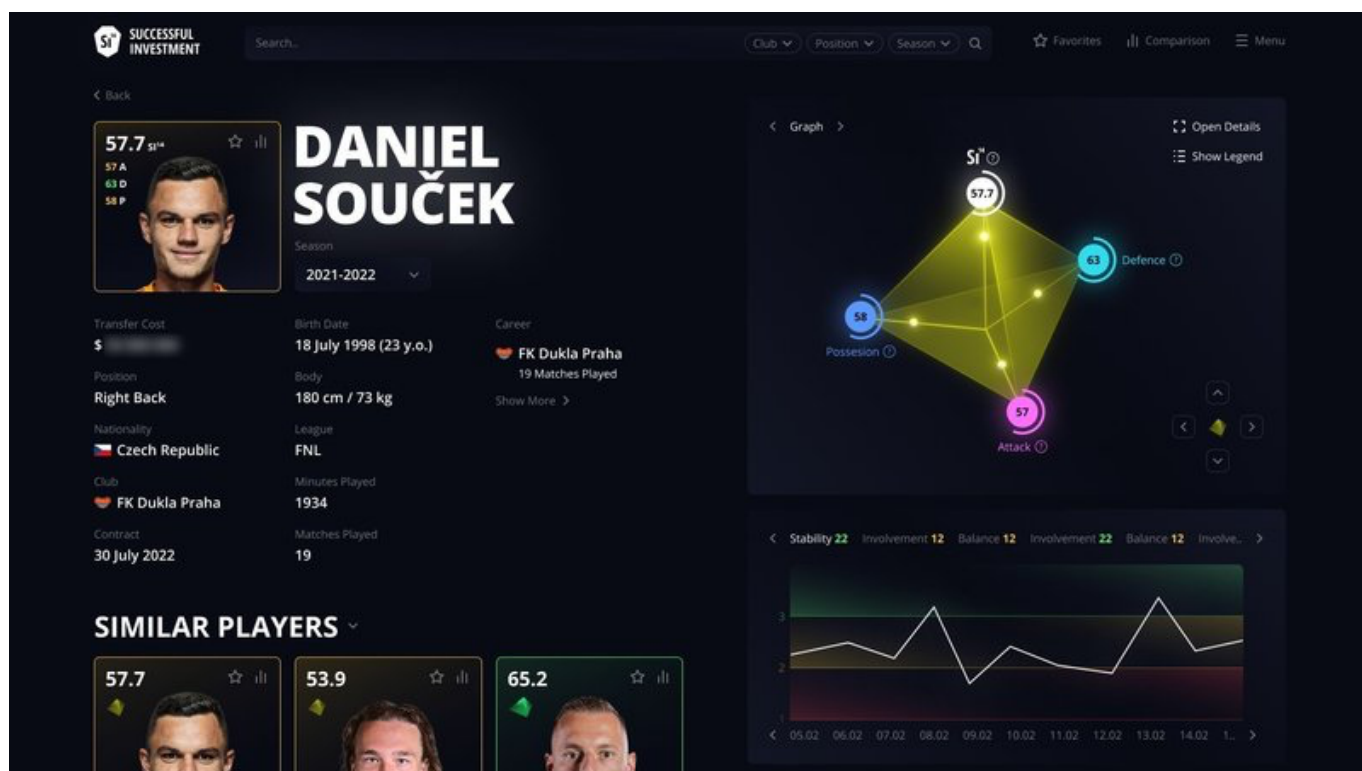
- What does it look like?

- The shape is a tetrahedron. The upper ball is the index value; the balls at the base can act on defense, offense, and on the ball. The larger the diameter of the balls and the farther they are from the axis, the higher the quality of the player.

**- Is the index the arithmetic average of the bottom three components?**

- Of course not. It depends on them, but not linearly because metrics are not part of the top three.

- If I asked you to find the right player in the Albanian championship, would you place a slide of tetrahedrons in front of me?

- Yes. It's a visual wrapper, a way of accustoming users to some perception standard. You go to the market, and you see a tomato. Your mind goes: red and big is good, green and small is bad, yellow and brown is weird or exotic. Hopefully, our index over time will begin to be perceived with their eyes, not as dried-up numbers. Especially since we also have a color gradation. Only the opposite of tomatoes: going to red - weak figures, going to green - strong.

- Have you checked the index in a real case?

- More than once. We've been doing it for many months, so we had to test the methodology on those who meet the notion of a strong player and manually control the computer. We searched through vast volumes of information. Finally, we got convinced that everything was working correctly.
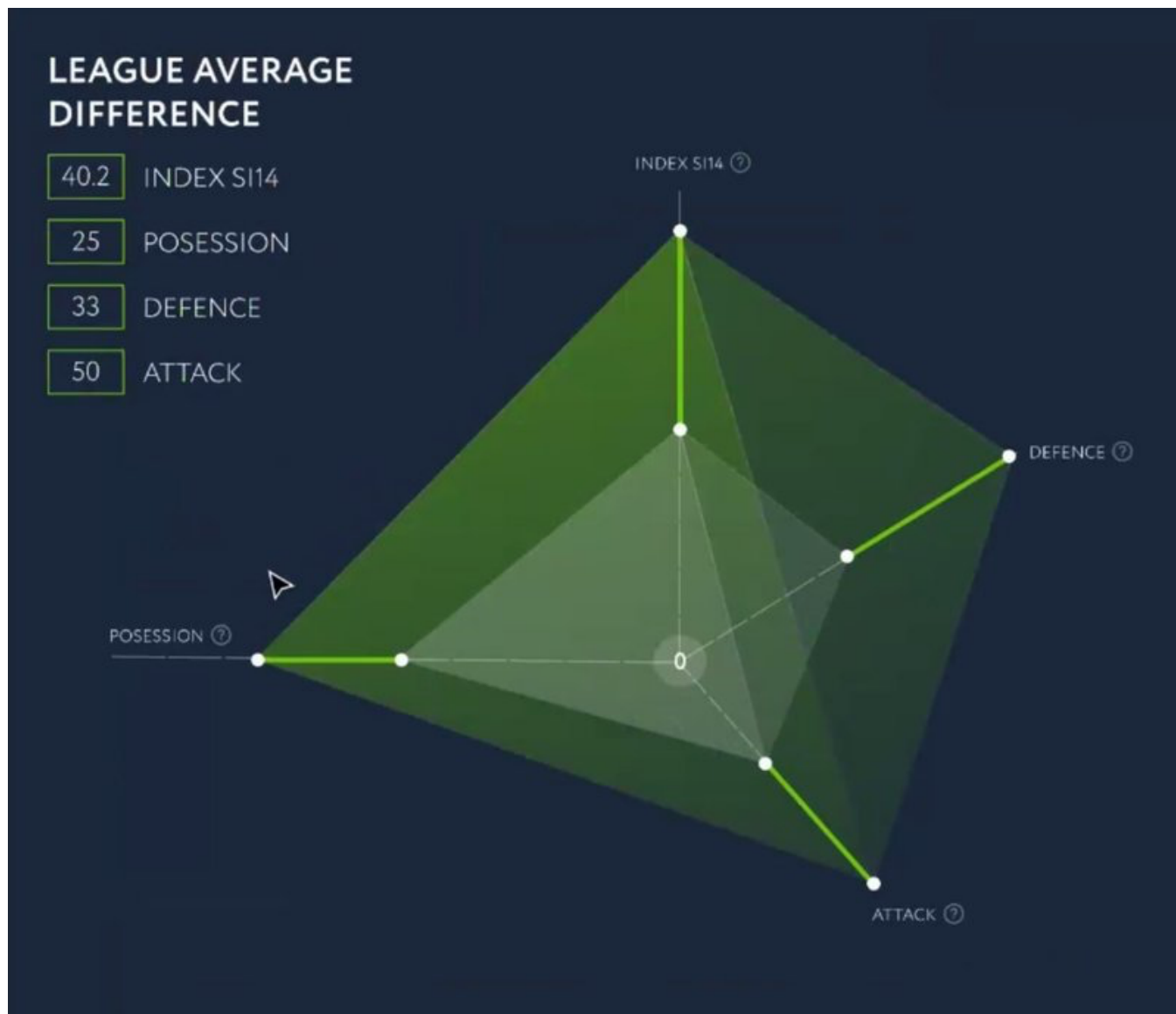
- You calculate a player, for example, from the Czech Republic. And they want to call him to England. The quality index is calculated taking into account Czech partners and opponents. What will it tell the English manager about the newcomer?

- First of all, the calculation methodology does not depend on the country. For example, we used the same algorithm in the Czech Republic and Spain championships - we were convinced of its correctness.

By default, the player is evaluated in his league and the world context. We understand that each league has its specifics, different levels of resistance, stylistic characteristics, and all this is reflected in the players' actions. Therefore, the computer has prepared profiles of all the championships of the planet; they are live matrices, changing with each round held. Thanks to them, it is easy to compare any players in the world, not quantitatively but qualitatively. In the same way, you can calculate the player's index concerning the championships of either England or Morocco. By comparing the league profiles with the model characteristics, we can quickly assess him within any tournament.

- When we put one index on top of another, Czech on top of English, the artificial intelligence will recalculate the tetrahedron, changing its components. This will allow us to guess how a player from one championship will perform in another one with a high probability.

By the way, another index we are working on is the coaching index. Many metrics characterize people of this profession. And there are a lot of parameters that allow us to derive patterns for comparisons.

- Is four numbers describing a soccer player enough for a user?

- They are far from four. For an unsophisticated consumer, our basic tetrahedron will be enough. For instance, it will show that a great forward is not defensively back, a strong defender doesn't join the attack, and a striker doesn't have a technique, scoring due to his flair and power. The bottom marbles, in this sense, are decryption of qualities, while the top one is an indicator of overall efficiency.

Well, for fans of the "sharp," we are ready to open the door to the analytical maze. By clicking on the index components, they can plunge into a sea of numbers, metrics, and stylistic characteristics of a player.

A big plus of our index is that it can take changes in the soccer world and respond to new trends. Just imagine: tomorrow the offside rules will be adjusted again. Artificial intelligence will not throw that metric out of the calculation but will change the calculation algorithm. Following the patterns, the tetrahedrons will change, but the completeness of the picture will remain. The ability to retrain is a massive advantage of artificial intelligence. And the more players it calculates, the more accurate its conclusions are. In general, expanding the mold is the future of the entire digital industry.

- "Balance," "engagement," "stability" - what are these indicators under the tetrahedron for?

- It's something that gets knocked out of complicated math but lends itself to simple math. Stability reflects a player's consistency in displaying his best qualities. Be it by season, be it by game. You can give five assists in a game where a weak opponent had two penalties. Or you can do it in five different games against the league leaders. Not the same thing.

Involvement is the number of active actions per unit of time. And the share of these actions in the soccer of the whole team. It evaluates the ability not to be distracted from the game, not disappear, and remain on the field. Mario Fernandez, for example, is far ahead of all his Russian colleagues in terms of involvement, constantly participating both in attacking and defensive actions. Others make fewer decisions and are generally less visible. As a rule, this parameter defines a player's leadership qualities, his team's importance, and his influence on team performance.

Balance characterizes the variation of a player's trump cards. Sometimes the coach needs a versatile soldier with a high stability rate, and at other times he is willing to forgive a man anything if he would score or carry the ball away. You can't call such a player balanced, but you can't call him useless either.

All three indicators are calculated using the same data set and are checked against the same metrics as the index. But at the same time, they facilitate the choice for people immersed in the subject.

- Speaking of choice. What is the practical application of your index?

- The question is not an easy one. So far, we do not see a convenient algorithm for working with the clubs. The prime cost of our services is high - expensive data, staff, equipment. Not everyone can afford it. Secondly, for effective use of any analytics, you need to be as immersed in the specifics of the club. What kind of team is the coachbuilding, his requirements for the players, what type of players he will use, and what model characteristics he does not need. Not everyone will want to open the inner kitchen to outsiders, so we do not look for clients. We work with clubs on personal contacts and requests, by an acquaintance, you could say.

- For us, it is evident that a convenient tool for everyone involved in the soccer industry has been created. The coach will visually see how the players change over any period—the club president - how the arrival of a new coach affected the team. Club analysts, selectors, agents are also among the interests. There are many variants of using the index, but I think it is better to let our child float free and offer it, for example, to fans, users of our betting exchange.

Nothing like this has ever existed in soccer before, so we just do not see all the opportunities opening up. Nevertheless, we are pleased with the successful completion of the work in which a lot of effort and innovative ideas have been invested. Many will have a chance for a deeper dive into the world of soccer. The rest remains to be seen.