Scalekarma.

Developing products, the smart way

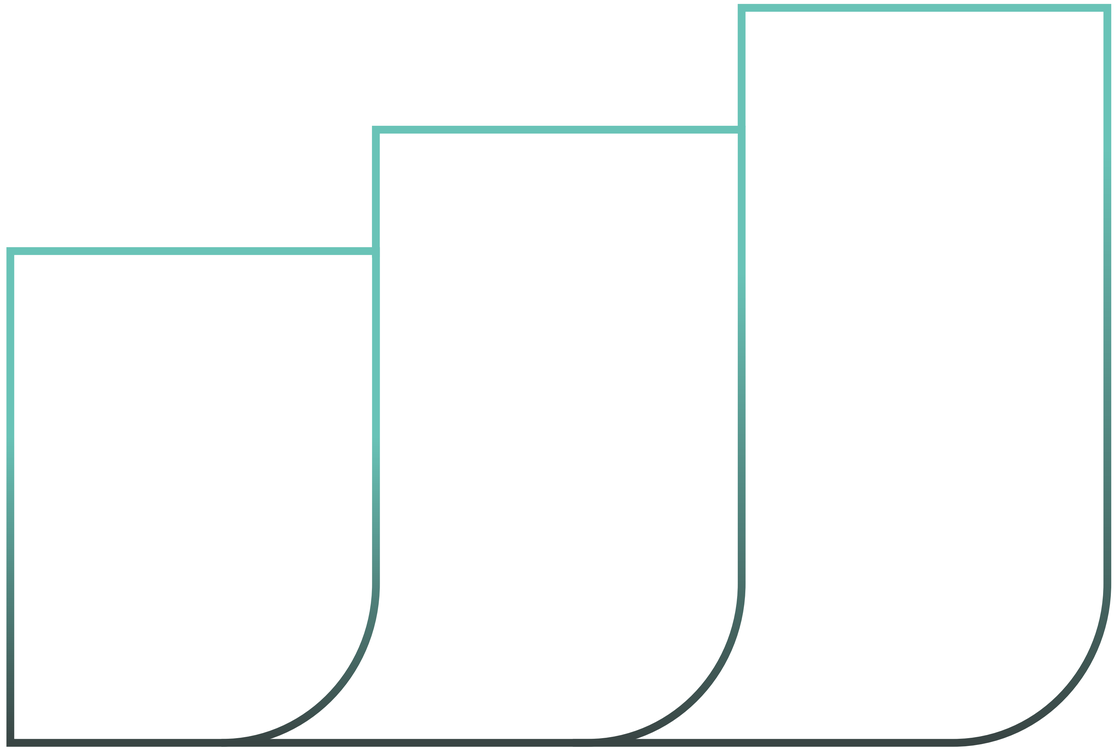
How to build better products with Scalekarma, the technology for product design, development, and marketing based on large-scale web search data analysis.

Authors

- Ilya Telegin
- Anastasiia Komarova
- Vlad Atamanenko
- Mariia Kuzmina

For Yus Teryukalov, in memoriam

You helped people discover what they were seeking for, and many found friendship, support, and inspiration.



Contents

We value your time, so this paper includes everything for the TL;DR reading experience. Each section has subtitles and a highlight. Reading through them provides a comprehensive overview of the paper contents. The TL;DR mode would take about three minutes of your time, while the full document may take over twenty. Please, have a nice read the way you prefer.

Introduction

Product-market fit challenge

> What do you need to start a business? Three simple things: know your product better than anyone, know your customer, and have a burning desire to succeed.
>
> — Dave Thomas, founder of Wendy's

Product companies are on a crusade to make the world a better place. This path is tough and has many crossroads where you need to make the right turn. Strategyzer says that seven out of ten new products and services fail to deliver on expectations. And, one of the biggest challenges is achieving the product-market fit.

Getting there involves a tough nut to crack, customer development. Traditionally, it's about a myriad of polls and in-depth interviews that take time, are expensive, and may end up providing distorted results due to miscommunication or simply data sampling issues.

Scalekarma eases those difficulties by quick and accurate reading of your customers' thoughts through their web search queries analysis. Build better products and features in tune with the true client's needs.

Living illustration

3Commas

Robotic crypto trading platform

Founded in 2017

- Raised $3M in a Series A round
- Serving 100,000+ active traders

Sergey Gustun
Product Owner at 3commas

Meet Sergey Gustun, Product Owner at 3Commas, a SaaS startup helping its users achieve a financial freedom through adopting algorithmic trading with cryptocurrencies.

3Commas has just acquired an investment round and wants to reach new clients from the broader audience: people, unaware of cryptocurrencies or setting up robotic trades.

Mr. Gustun is craving to learn more about the new audience: their values, pains, and even thoughts. He is entering the phase of a thorough and lasting customer development. How can Scalekarma help him?

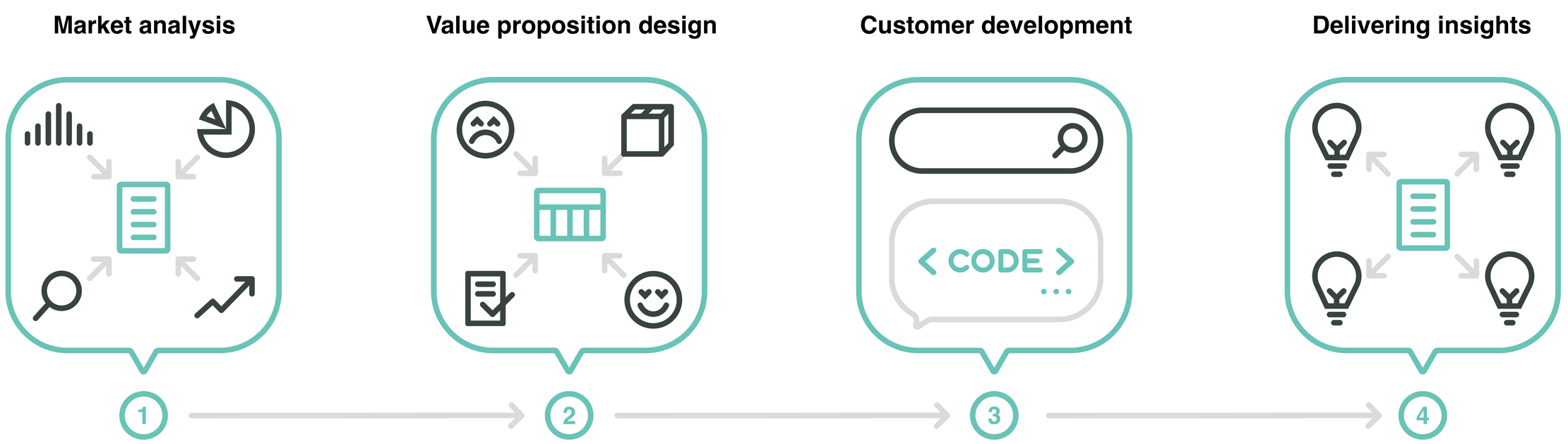
Methodology

To minimize the margin of error and optimize costs, we conduct product research as a series of interviews with the target audience. But instead of interviewing people, we use their web search queries.

Talking to search engines, we can quickly access a vast amount of data and get answers of better quality than those available through a personal live interview. The reason is, people can accidentally or on purpose give wrong information in the conversation, but they always formulate search queries accurately and honestly. Also, you can't physically conduct millions of live interviews, while you can do so with search engine data.

> At the privacy of their keyboards, people confess the strangest things, sometimes (as in dating sites or searches for professional advice) because they have real-life consequences, at other times precisely because they don't have consequences: people can unburden themselves of some wish or fear without a real person reacting in dismay or worse.
>
> — "Everybody Lies" by Seth Stephens-Davidowitz

The research methodology consists of four stages. In this whitepaper, you'll learn what happens at each stage and how it all forms a unique dataset to take your product to new heights.



Step 1: **Market Analysis**

Market data analysis: collecting secondary data

Discovering the main players on the target market

It all starts with collecting and analyzing the market context. In our example, 3Commas is a crypto trading platform. To begin with, we find the top companies that operate in the crypto industry in the US and offer similar products.

To find the top companies, we use the following criteria:

- The founders' vision of the market — they are the main holders of the company's roadmap so we need to know their goals
- Search data on queries like: 'top crypto trading companies,' 'trading bots,' etc.
- Gartner, Crunchbase, and other ratings

We study the collected top companies in detail using public materials such as reports and papers. To find this content, we use a combination of tools like Technews and Ahrefs. They help us find materials that are key to reputation and traffic generation. As a result, we understand what the leading companies in the industry are busy with and what's their vision and expertise.

Identifying the relevant industry trends

To identify current trends we select the top industry media or blogs and analyze what they write: news, articles, interviews. It helps us understand what topics are hot in the industry right now.

While choosing sources, we rely on articles mentioning the brand: the most authoritative and most visible in the search engine results. We look for specific wording in such articles to determine how the audience refers to a problem or product in their search queries. Then we use them along with brand keywords to search for publications with the greatest coverage according to Technews.

Studying data from analytical agencies (e.g. Gartner, Forrester, Nielsen, Deloitte) complements this picture, helps see industry trends and predict how events will unfold. Technology, expectations, problems and barriers, investment climate, market size — all of these fall under the details of the market context.

Revealing the top challenges and success metrics

Having analyzed the market, we find out what problems exist in the niche, identify success metrics accepted in the industry, and which of them the market leaders use to report to their shareholders. Enabling investors to assess the traction, they are likely to be crucial for each company's life. We can make a list of such key problems and metrics along with their baseline, desirable, and target levels for each industry.

Identifying decision-makers

We can determine what trends in the implementation of similar solutions exist in the target industry. Another research point covers who and how we would reach out to in order to shorten the buying cycle. For B2C products this is about the persona attributes, while for B2B products that's largely about the target roles.

Profiling the target audience

Once we know what industry leaders work on, how they live, and what they strive for, we can define the target audience segments that will benefit the most from our product.

When the collection of secondary data is complete and the general view of the industry is clear, we begin to form a value proposition template using the Strategyzer's model.

The collected data allow us to look at existing products and value propositions in the market of interest and find useful insights for the new proposal.

We deepen the expertise and better understand what results are expected from a specific class of products in the target niche.

Step 2: **Value Proposition**

Creating value proposition canvas (VPC) outline

What Gains do audiences in the industry strive for?

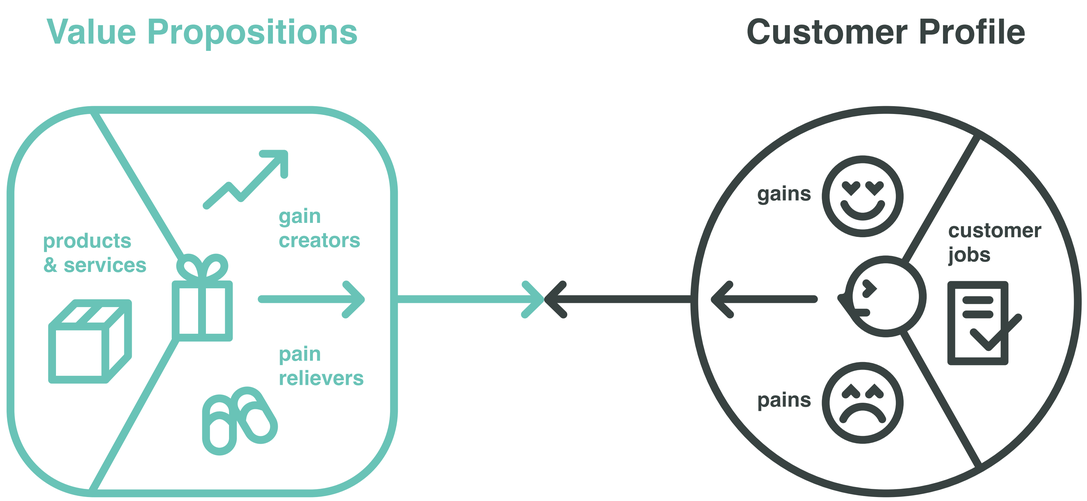
Let's start by defining the goals and benefits that our potential customers want to reach on this market. For cryptocurrency algotrading, this will mean simple access, building trust and loyalty, providing high profitability, helping people achieve financial security and freedom.

What are the problems (or Pains) in the market?

Finally, we take a look at the obstacles along the way and identify the client's problems or pains. As for the crypto, it might be the fear of fraud, software reliability, capital protection, token price volatility, and unobvious crypto taxes across different legislations.

What kind of work (or Jobs) can our product do?

Jobs describe the things your customers are trying to get done in their work or in their life. A customer job could be the tasks they are trying to perform and complete, the problems they are trying to solve, or the needs they are trying to satisfy. This term is widely known due to the proven and widely adopted Jobs To Be Done framework. Determining the jobs hierarchy helps understand who and why will use our product, and adjust our communication accordingly.



Visual representation of Strategyzer's Value Proposition Canvas.

Crafting a Value Proposition Canvas

After collecting and studying the data, business analysts create the first iteration of the value proposition template. It's a kind of sketch that we later enrich with details from in-house company vision, expertise, and with data from public sources.

For example, based on what we have learned about algotrading, the 3Commas platform will help its adopters overcome inflation, stay ahead of the natural market movement, reduce trading risks, and get rid of negative psychological factors while committing trades. All this contributes to strengthening the innovation and enhancing trust.

The new knowledge about the audience is then applied to company and product messaging to convey values from the points of customer jobs, gains or pains to increase chances of mastering the market.

> **Based on the data collected, we design the value proposition canvas. This deliverable is useful to create or improve products, as well as to adjust marketing communications. It's also a basis for further analyses.**

Customer Job	Domain
Rapidly assess crypto market	general
Overcome inflation	expert
Automate cryptocurrency trading	expert
Reduce crypto trading risks	expert

Customer Gain	Domain
Simple registration	general
Passive income	expert
Quick trading bot set up	expert
Leverage advanced algorithms	expert

Customer Pain	Domain
Unobvious crypto taxes	general
Fear of fraud	general
High transaction fees	expert
Algotrading software failures	expert

View full sample

Step 3: **Customer Development**

Value proposition canvas validation and expansion

In-person interviews problem

When the secondary data is collected and the value proposition basis is described, we proceed to validate it with holders of the product expertise from a client company.

The in-house experts provide the necessary deep immersion in a market niche and help confirm which of the VPC features are on their roadmap already, and in which directions they want to move the product further.

However, it is difficult to find and contact any expert respondents “outside” the new industry. Geography and time zones might become the first problem, while questionnaire design and interviewers training are almost sure to add up to the obstacle list.

After coding the first cycle of interviews, more iterations might be necessary for obtaining rich and accurate expert information. And interview cycles require a plenty of time and effort.

To get it, we would run the first interview cycle, derive expert keywords with frequency analysis, and iterate further until the list of such keywords stops growing.

Search data analysis is the key

To solve this problem, we found a way to address our questions to a unique mind — search engine. To ask the questions we use the search bar. To get the answers coded as keywords, we adhere to marketing services such as Ahrefs.

People often refer to search engines to solve their problems, and the search results thus contain the entire array of solutions. The methods of manipulating such data are much wider than those for classic live interviews.

To prepare for the data extraction, we design sequences of search queries in a way similar to the user's path of product recognition. Such sets include all the steps of the classic AIDA model from minimal awareness to formalized tasks and decisions about “hiring” a product to perform them. To build such a sequence, we use the concept of the **awareness ladder.**

Awareness ladder, the concept

Awareness ladder is a tool that helps you understand at which stage of the customer journey your target audience segments are and what questions they ask at each stage.

The ladder has five main steps that serve as the basis for our interview questions to Google.

Question for each stage	Answers examples
What?	**financial freedom**
How?	**overcome financial insecurity**
How exactly?	**financial freedom checklist**
In what ways?	**investment instruments**
In what way?	**algotrading service**

First-order ladder

After designing the VPC and validating it with the product specialists, we expand the canvas by formulating user search queries corresponding to different steps of the awareness.

We call the dataset a **first-order awareness ladder**. It's an intermediary data needed to further fetch relevant user searches.

The process of designing the first-order ladder is where we adhere to the professional interviewing experience. To do the job, we use the skills of experienced customer development specialists. Thus we minimize the margin of error and enhance the quality of all the output datasets.

→ View a sample of 3commas First-order Awareness Ladder

> **After designing the VPC and validating it with the product specialists, we expand the canvas by formulating user search queries corresponding to different steps of the awareness. We call the dataset a first-order awareness ladder. It's an intermediary data needed to further fetch relevant user searches.**

Fetching user searches

This is where we get to fethcing the data. Search engines are smart enough to deliver relevant search results for many types of complex queries, like those originating from voice search.

Scalekarma leverages the search engine intelligence to fetch relevant user searches, eliminating possible inconsistency with the target audience language in the VPC.

We form search queries from the extended VPC and run each element through Google Search, following the behavioral footprint of the target audience. Then, we treat URLs from search results as answers to specific questions.

Then we adhere to the relevant search CTR distribution curves and outline the results that fit best. For each top result, we determine organic keywords with Ahrefs. This helps us understand which keywords exist in the search engine database and how they answer the questions of the target audience.

The whole process of fetching relevant searches from Google goes as follows:

1. Long-tail search queries are sent to the search engine with a specific search context, like geography, etc.
2. Google's algorithms reduce long-tail to a specific corpus of keywords, identify seep phrases, and show SERP. The SERP position of the resulting URLs highly correlate with their accuracy.
3. Knowing the CTR distribution in SERP, we determine the best URLs containing reasonably accurate and complete answers to the user search queries.
4. We use these URLs to build reports on organic keywords in Ahrefs. This way we answer the question: “what exactly would representatives of the target audience have entered into the search bar instead of a long-tail query to get their question answered.”
5. Thus we get coded answers to the questions of our "interview" from the representatives of our target audience.

Step 4: **Product Insights**

1 2 3 4

Fetching and organizing product semantics

Product semantic kernel

As a result of fetching user searches, we get a collection of keywords resembling **a semantic kernel**. Attributed with the ladder elements, it describes how the target audience perceives the product (or similar products) at different stages of acquaintance with it.

1. Each keyword is assigned to a query from the first-order ladder so we could trace every keyword down to its parent product feature.
2. It allows matching the keywords with custom attributes (e.g. audience segment, level of expertise).

The resulting kernel contains coded responses to a customer development interview conducted with tens of millions of respondents — this is the approximate sum of keyword search volumes in the kernel. As for the keywords, we end up collecting five- to six-digit amounts.

The product kernel is then used for deep market research, product positioning, building SEM/SEO, marketing, and communication strategies. Using it, we can assess market volumes and translate the value proposition into the language of the right audience.

Since the product semantic kernel is large, it can be tricky to use. To simplify its usage, we build a dashboard with the detailed stats for each element of the extended VPC dataset. The dashboard is capable of quickly building keyword corpuses for each VPC element.

The stats are based on the derived underlying organic keywords for each VPC item: Count, Total Volume, Cost, Difficulty, etc.

solution to get financial security	37,490 Organic demand volume	3.49 Average keyword cost per click	61 Keywords count	22.69 Average keyword difficulty
digital personal financial advisor	502,860 Organic demand volume	11.57 Average keyword cost per click	546 Keywords count	41.19 Average keyword difficulty

View full sample

Having both the product kernel and a dashboard for it, we proceed to segmenting the derived datasheet.

Segmenting the derived datasheet

We use natural language processing algorithms to segment the resulting set of keywords. We group them according to the extent to which the audience is ready to purchase our product.

The output of such a segmentation is also a dataset. We refer to it as the **second-order awareness ladder**. Its steps contain a collection of all words and phrases related to the product in the minds of the target audience segmented by different search intents. The dataset can be used to correctly position the product in the target market across the variety of media communications and serves as the basis for creating or improving content marketing strategies.

For example, for the 3Commas cloud platform we picked up 79,600 keywords. Such a kernel is much cleaner from the perspective of parasitic search queries: quantitatively, about 8–17 times compared to analogs built with other methods. Since we accurately formulate “product interview” questions and choose best answers to them from the search data, we naturally filter out garbage keywords with no relevant meaning. We save time on manual filtering or developing sophisticated algorithms for that.

keyword	volume	cpc	difficulty	ladder-step
cryptocurrency	**197,000**	**1.5**	**88**	**awareness**
how to invest money	**20,000**	**6.0**	**71**	**how-to, hl**
buying stock 101	**200**	**35.0**	**67**	**how-to, templ.**
personal financial advisors	**250**	**25.0**	**84**	**tool discovery**
financial portfolio management software	**10**	**55.0**	**28**	**comm. intent**

Part of the second-order awareness ladder demonstrating the keywords and their parameters.

View full example in Google Sheets.

This segmented semantic context is enough to assess the organic demand, design a new product or improve an existing one, convey its features and values in a way consonant with the needs of the audience within your target market.

Conclusion & What’s Next

The resulting dataset provides a context for product design, development, and marketing that ultimately supports the company at every stage of its organizational life cycle. However, this isn’t the only deliverable produced by Scalekarma. The full set of deliverables is described below and supports the decision-making and helps uncover an optimal way for planned growth.

With all the data at hand, Sergey now has the product superpowers he needed to commit a tectonic shift in product vision, generate hypotheses, test them, and acquire new audiences.

Thus, with no additional manipulations with the kernel, you can understand how the audience refers to products that have Gains, Pains, and Jobs similar to yours. For example, Sergey from 3Commas found out that he shall experiment with the ‘online broking service’ and ‘digital financial advisor’ product positioning if he wants to reach audiences unaware of the crypto industry.

Search query analyses provide audience insights way beyond the initial VPC contents. We reveal the exact wording the target audience uses when looking for solutions to their problems and identify more meaningful keywords and contexts driving the product thought further. This means more hypotheses and deeper insights.

> **Having the semantic kernel and knowing the market context, you can shape the strategy to get targeted product traffic. As a result, you can spend a minimum of time, money, and human resources to make all marketing efforts at product promotion more effective.**

Deliverables

value proposition canvas (VPC)
built together with client's product specialists and decision-makers, based on the market analysis, in-house expertise, and product development strategy.

View full sample

plain semantic kernel
of real-world search queries attributed with the extended VPC elements: words and phrases used by the target audience to discover existing or planned products and features.

View full sample

detailed report
built by Scalekarma product analysts and including the market and competitor analyses, proposed product development directions, and the outline of the product vision and assessments of needed content investments to convey it to the target audience.

extended VPC
transformed into an awareness ladder holding the potential search queries of the target audience.

View full sample

product semantic kernel
a form of the plain kernel with keywords segmented by user intent: an extent to which the target audience is ready to make the purchase decision.

View full sample

product development dashboard
a list of extended VPC items attributed with in-depth statistics for the related underlying keywords: volume, cost, complexity.

View full sample

Scalekarma.

Revolutionize your product marketing with Scalekarma.

Ilya Telegin

Co-Founder & CEO

email
ilya@scalekarma.com

telegram
@reddishradish

Schedule a demo

Yandex · ElectroNeek · CRYPTOPAY · 3Commas · SEMRUSH · chatfuel · Uploadcare · SEXLIKEREAL · HEARST · TSU · Bookform · Tweek · scada · Сеть городских порталов SHKULEV MEDIA HOLDING

Developing products, the smart way