

Building a Self-Learning Search Engine: From Research to Business

Manos Tsagkias
manos@904labs.com
904Labs, Amsterdam, The Netherlands

Wouter Weerkamp
wouter@904labs.com
904Labs, Amsterdam, The Netherlands

ABSTRACT

904Labs B.V. was founded in 2014 by Wouter Weerkamp, Manos Tsagkias, and Maarten de Rijke to commercialize state-of-the-art search engine technology. 904Labs' strategic product is a self-learning search engine for online retailers, which uses some of the most recent scientific developments in machine learning and search engine evaluation. 904Labs has raised about 200K in funding and has signed pilots with large international and national companies. Since its start, 904Labs has grown with two developers and two experienced business people. In this presentation we tell how to go from research to business and the challenges it brings along?

1. FROM RESEARCH TO BUSINESS

In our talk we take you on a journey that starts at a university and ends owning a search technology company that is still in its early stages. We describe this journey in several steps, including the lessons learned after each step. We explain the current state of the company and the future plans and opportunities, and conclude by listing some of the major challenges.

1.1 Setting off

The journey of 904Labs starts at the University of Amsterdam. Witnessing a growing number of organizations knocking on the university's door to ask for help with their "scientifically interesting" problems, which turned out to be not of interest to the university, we saw an opportunity for a business entity to jump in and help these organizations solve their problems. 904Labs was born, and we thought of it much like an extension of the research group from which it originates. As one can imagine, the problems organizations face are as diverse as the research done within the research group, ranging from entity extraction and linking, via text summarization, to more regular search issues. Offering all these solutions is infeasible as a startup for three reasons: (1) there is limited manpower, you cannot do everything; (2) it is unclear to the outside world what the company actually offers; and (3) there is no scalable business model in developing custom-made software solutions.

Lessons learned: Focus. Do not try to solve everything, but pick one problem and design your solution around it. Think of it as

writing one paper. What is the core problem you are trying to solve and focus on that particular problem. Ignore the rest, as interesting as it might be, you might add it in a later paper.

1.2 Choosing a path

Working under the assumption that most businesses were by now at a reasonable level when it comes to implementing search solutions (e.g., use Elasticsearch or Solr), we figured that the world was ready to meet self-learning search. Self-learning search is a fancy name for online learning to rank, which is learning to rank tightly coupled with interleaving, which allows real-time learning and experimenting with new search algorithms.

After initial talks with a few potential clients (i.e., large e-commerce sites), it became clear that these organizations are not looking to replace their current systems with yet another solutions. Also, the search engines running at their sites are part of very complex product and customer management systems, in which we are not interested. On the other hand, these large sites could potentially benefit a lot from self-learning search, given the amount of traffic and the revenue deriving from search. Taking everything into consideration, we decided to offer self-learning search as-a-service, a middleware layer sitting between the e-commerce site and the client's existing search infrastructure. It allows for quick easy integration.

Many months later, after talks with yet other potential clients, we found out that SaaS solutions have their own problems (e.g., latency, reliability), keeping large organizations from buying. Implementing self-learning search as plugins for major search solutions (Elasticsearch, Solr) is now another viable direction we are exploring.

Lessons learned: Test assumptions. You cannot do this enough (and we actually did not do it enough). Which problems are your clients facing? What do they want? How do they want it? Ask, ask, ask. Do not sell, but learn. It is like testing the research waters with a small workshop paper before working on the full SIGIR submission. If no one cares about the workshop paper, do not bother writing the full paper.

1.3 En route

The direction has been determined, now how to actually go there? Being confident researchers, we thought implementing well-known methods into a working system should be a matter of weeks. Selling it would, of course, be easy, because who would not want to have such a cool system? So we would be making money within a few months, meaning that until then we could do without funding.

Wrong. Building a product takes time. A lot of it. Selling takes at least an equal amount of time, and probably more when you are not a natural-born salesperson. This means that you will not see any money coming in from product sales any time soon. Unless

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

SIGIR '16 July 17-21, 2016, Pisa, Italy

© 2016 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-4069-4/16/07.

DOI: <http://dx.doi.org/10.1145/2911451.2926728>

you get more people on board, for which you need ... money. With a growing need for revenue, it becomes hard to say no to consultancy projects, which take up much of your time and distract from developing your own product. We finally figured out that money is a necessity and applied for various grants, receiving 175K euro. The money gave us the opportunity to hire developers, speed up product development, and focus on business development instead of consultancy jobs.

Lessons learned: Take care of the financials. And stay on track. Try to find grants (e.g., from your university) or investors so that you do not get sidetracked by consultancy jobs. It is like working on your final technical thesis chapter, while you have already started as a postdoc somewhere else. You get paid, but cannot do the work you should be doing.

1.4 Getting lost and finding the way

Being en route does not mean you cannot change directions anymore. Or that you will not get lost. From a business perspective, finding the right client segment is a challenge, and one that continually changes. Your marketing depends on your client segment, your website depends on the marketing strategy, grants have milestones that should fit product development, which in itself is dependent on the problems faced by your potential clients. And how do you convince clients to buy your product? We might have PhDs in search technology, but why would potential clients care? They do not. Being convinced of your product is one thing, but selling it is another. After trying to get pilots and clients, with mixed success, and doing so without a clear strategy, we decided that it was time to acknowledge our weaknesses. We are not experienced business men and you do need those to find your way. We have reached out to potential investors who, instead of investing money, have jumped on board in hands on positions related to business development, marketing, and sales. Having experienced people taking care of this side of running a company helps immensely and allows us to focus on the things we are good at.

Lessons learned: Take care of the team. Make sure you have (experienced) people at the important positions in the team; you cannot do everything yourself. Sales and marketing is a very different beast. Think of writing a paper, scrapping all technical sections, and leaving only the main result and original problem ("Web search is hard. We improve MAP by 10%"). No one cares how you do it.

1.5 What the future may bring

We have a first version of our product, self-learning search as-a-service. We have pilot agreements with several companies, like Elsevier, WizeNoze, and Autostyle. We have a team that includes developers, business people, and ourselves. And we have a buffer of money. Now we need sales. Hardcore sales. We are in the money-making business, so let us try that. After money comes in, we can focus on improving the product by thinking about faceted search and online learning to rank, personalized online learning to rank, and delayed feedback.

1.6 Conclusion

Are you thinking of entering the startup arena with a brilliant idea that has proven itself in research? Do it! It is a fun ride. But do keep in mind that business is not research. People do not care about the technical and theoretical details, they want to know which problems it solves and what it brings them. So, (1) talk to potential clients, find out what their problems are; (2) Focus on one problem and create a solution for it; (3) Get funding to hire people and avoid distractions, and (4) take care of the business side of things.