

Summary

1. **Volume of investments required – \$ 230 thousand.**
2. **Production** – IT: systems of automatic text processing, information retrieval systems.
3. **Trade marks** – Dictum[®], DictaScope[®].

Company profile

Date of establishment – April 2006. The company was established to develop the project “Computer system of citation extraction from natural language texts”; first sale was in May 2006.

Signs of public recognition – winner of the START-06 competition.

Number of employees – 9 persons.

Team

Okatiev Vladimir – Director, 44 y.o. The project’s initiator, founder of the company. Responsible for search of investments and formation of funds. Candidate of physical and mathematical science (1993); engineer of Russian Federal Nuclear Center (Sarov city, 1985–1989); lecturer of computer software chair of UNN (1989–1994); director general of NizhegorodInvestService Ltd (1997–2006); director of Dictum Ltd since 2006; experienced in managing commercial projects with a budget up to \$ 12 000 ths. with a 90-persons team.

Barkalov Konstantin – Project Manager, 29 y.o. Founder of the company. Graduate from UNN CMC faculty (2000); Candidate of physical and mathematical science (1993); participant of several international projects. In Dictum Ltd he is responsible for IT-development.

Products characteristics

Dictum Ltd develops the technology of information retrieval and analysis of text information.

1) System of citation extraction will allow setting up access to text information on fundamentally new level as compared to general-purpose search systems. Responses to user inquiries will be citations rather than set of documents.

Present general-purpose retrieval systems and particularized legal system (e.g. “Garant”, “ConsultantPlus”) are intended for document search in electronic information collections. Often a user needs not a whole document but a reasonable part of information from this document.

In contrast to the mentioned systems, the developed technology of citation extraction allows to minimize necessity of reference to initial documents. Citations will contain information, which is necessary to user. The technical result from development of the proposed idea consists in reduction of volume of the information displayed on the user’s terminal by user inquiry, and reduction of intellectual expenditures on the analysis of the received information and on the decision-making.

The novelty of the proposed idea consists in following:

- citation (a text part, which has semantic completeness) is extracted as an answer;
- citation is a result of excluding unnecessary information from a text;
- imaging system allows to minimize scrolling;
- hierarchical structure of a text is actively used for presentation of search results.

Application for patenting the invention “Information retrieval system” has been sent (declarant: Dictum Ltd).

Several ways of commercialization are proposed. In particular, system will allow setting up access to legal and technical information on fundamentally new level owing to proposed navigation and retrieval engine. For subscribers of cellular communication networks access to legal information will be possible by means of cell phone. Existing systems such as “Garant”, “ConsultantPlus” and “Codex” allow searching documents, but not citations.

Other use of the system are:

- information retrieval in corporate network;
- entertainment service in the form of imaginary dialogue with public persons. Answers to user’s questions will be in the form of opinions of popular persons published in Internet, including news blocks.

Voice interface will be available in future.

2) Natural language parser DictaScope is intended for developers of linguistic software. In contrast with analogues (Dialing and Connexor software), DictaScope parser has higher speed and it may recover missing constructive fragments of sentence (ellipses).

Computer software “Syntactic analyzer for Russian DictaScope” was registered.

DictaScope analyzer is used by SIA Tilde (Riga, Latvia, www.tilde.lv) on the basis of license agreement on development of pilot version of the Baltic-EU languages electronic dictionary with automated translation features under the project of Latvian Investment and Development Agency.

Current state

Currently the project’s R&D is financed by the state Foundation for assistance to small innovative enterprises; volume of state contract is \$ 175 ths. for three years.

By now the following systems have been developed:

- 1) demo-prototype of system of citation extraction from Russian and English texts;

- 2) natural language parser DictaScope (commercial version for Russian and research version for English and German).

Development strategy

Use of funds

1. R&D	0%
2. Acquisition of fixed assets (computers and software)	20%
3. Marketing	60%
4. Acquisition of current assets	10%
5. Other	10%

Prospective outcome of investment

As a result of investments, an Internet-service with navigation and retrieval engine based on citation extraction technology will be developed. Service will be available for users of personal and pocket computers as well as for subscribers of cellular communication networks. A commercial version of text parser for English language will be developed.

Marketing & Markets

1) Volume of Russian market of legal information systems has increased from \$ 250 mln in 2005 up to \$ 500 mln in 2007. Demand for legal information will increase because of active lawmaking. Most popular systems presented on Russian market are "Garant", "ConsultantPlus" (in sum more then 90% of market) and "Codex".

The cited systems are oriented first of all on the medium and large corporations. Demand for legal information from the direction of natural persons and small companies (including private entrepreneurs) is not enough supplied at present. Dictum Ltd plans to occupy this niche by means of eLawyer-service based on citation extraction technology. At the initial stage, audience will be formed on the basis of free access to limited document collection.

Volume of the market of English-language law information at least tens as much as Russian market, as far as English is a language of international relationships, trade and transport.

Volume of the market of corporate retrieval systems is evaluated as \$ 1.2 mlrd; its growth rate is more than 30% per year.

2) At present, the market of syntactic parsers is in the forming stage. In the near future, developers of linguistic software will form the basis of solvent demand for text parsing systems, among which are DictaScope. Inquiries from developers of retrieval systems, machine translation systems and systems of speech synthesis to Dictum Ltd are evidence of demand for text parsing systems. Moreover text parser may be used in question-answering systems, autoabstracting system, systems of remote leaning and automatic knowledge checking, and other system of automatic text processing.

In particular, the volume of world marked of translation is about \$ 20 mlrd. Machine translation (e.g. "Promt", "Systran", etc.) accounts for 1% of all market. This quota increases simultaneously with the increase of quality of machine translation, which essentially depends on syntactic analysis methods.

Interaction with investor

It will be offered to investor to buy shares of Dictum Ltd. Offered share will be blocking (25.1%) at least. It will be possible to include a representative of the investor in the company's management.

Approximately after four years, venture investor can exit by sale of his share in Dictum Ltd to a strategic investor. Strategic investor may be a company – developer of retrieval or information system (e.g. "ConsultantPlus"). Forecast for the price of mentioned share after four years is \$ 5000 ths.