

Резюме

1. **Объем запрашиваемых инвестиций: \$ 230 тыс.**
2. **Продукция** – информационные технологии: системы автоматической обработки текста, информационно-поисковые системы.
3. **Торговые марки** – Dictum[®], DictaScope[®].

Общая информация о компании

Дата создания – апрель 2006 г., компания создана для реализации проекта «Компьютерная система извлечения цитат из текстов на естественном языке», первая сделка – май 2006 г.

Знаки общественного признания – Победитель конкурса СТАРТ-2006.

Количество работающих на фирме: 9 человек.

Команда

Окатьев Владимир Васильевич – директор, 44 года, инициатор проекта, основатель компании и ответственный за поиск инвестиций и привлечение средств. К.ф.м.н. (1993); 1985–1989 гг. – инженер Российского Федерального Ядерного Центра, г. Саров; 1989–1994 гг. – преподаватель кафедры математического обеспечения ЭВМ ННГУ; генеральный директор ООО «Нижегородинвестсервис» (1997-2006); директор ООО «Диктум» с 2006 г. Имеет опыт реализации коммерческих проектов с бюджетом до \$ 12 млн. и руководства коллективом численностью 90 человек.

Баркалов Константин Александрович – ответственный исполнитель, основатель компании, 29 лет. Выпускник факультета ВМК ННГУ (2000); к.ф.м.н. (2006). Имеет опыт участия в международных проектах. В ООО «Диктум» отвечает за разработку информационных технологий.

Продукция

ООО «Диктум» разрабатывает технологии информационного поиска и анализа текстовой информации.

1) *Разрабатываемая система извлечения цитат* позволит организовать доступ к текстовой информации на принципиально новом уровне: ответом на запрос будут являться цитаты, а не набор документов.

Существующие поисковые системы общего назначения и специализированные справочно-правовые системы предназначены для поиска документов в электронных массивах. Но, часто пользователю нужен не сам документ, а лишь небольшая часть информации из этого документа. Лишняя информация усложняет жизнь.

В отличие от указанных систем предлагаемая технология извлечения цитат позволяет свести к минимуму обращения к исходным документам – необходимую информацию будут содержать выдержки из документов. Технический результат от реализации предлагаемой идеи заключается в сокращении объема информации, выводимой на дисплей пользовательского терминала в ответ на поисковый запрос, что приводит к уменьшению интеллектуальных трудозатрат на анализ полученной информации и принятие решения.

Новизна предлагаемой технологии поиска информации заключается в том, что:

- в качестве ответа извлекается цитата – часть текста, обладающая смысловой законченностью;
- цитата получается в результате исключения лишней информации из текста;
- система визуализации позволяет минимизировать скроллинг (прокрутку текста);
- иерархическая структура текста активно используется для представления результатов поиска.

Подана заявка на изобретение «Информационно-поисковая система», заявитель – ООО «Диктум».

Предполагается несколько направлений коммерциализации. В частности, система позволит организовать доступ к справочно-правовой и нормативно-технической информации на качественно новом уровне за счет предлагаемого навигационно-поискового аппарата. Для абонентов сотовой связи доступ к правовой информации будет возможен через сотовый телефон. Существующие системы - позволяют искать документы, но не цитаты.

Другие применения системы:

- поиск информации в корпоративной сети;
- развлекательно-информационный сервис – виртуальный диалог с публичными персонами: ответом на вопросы пользователя будут высказывания известных людей, опубликованные в Интернет, в т.ч. в новостных лентах.

В перспективе возможен голосовой интерфейс.

2) Анализатор естественного языка DictaScore предназначен для разработчиков лингвистического ПО. В отличие от аналогов – анализатор DictaScore обладает более высоким быстродействием и возможностью восстановления пропущенных (подразумеваемых) фрагментов предложений (эллипсисов). Зарегистрирована программа для ЭВМ «Синтаксический анализатор русского языка DictaScore».

Анализатор DictaScore используется на основе лицензионного соглашения компанией SIA Tilde (г. Рига, www.tilde.lv) для разработки электронного словаря с элементами машинного перевода для языков стран Балтии и языков Европейского Союза по проекту Латвийского агентства инвестиций и развития.

Текущее состояние

В настоящее время проект в части НИОКР финансируется государственным Фондом содействия развитию малых форм предприятий в научно-технической сфере. Сумма госконтракта – \$ 175 тыс., срок – 3 года.

К настоящему времени разработаны:

- 1) демонстрационный прототип системы извлечения цитат из текстов на русском и английском языках;
- 2) анализатор естественного языка DictaScore (коммерческая версия для русского и исследовательские версии для английского и немецкого языков).

Стратегия развития

Направление использования инвестиций:

- | | |
|--|------|
| 1. Приобретение основных средств (компьютерная техника + ПО) | 20%; |
| 2. Маркетинг | 60%; |
| 3. Обратные средства | 10%; |
| 4. Другое | 10%. |

Результат инвестиций

В результате инвестиций будет разработан Интернет-сервис с навигационно-поисковым аппаратом, основанным на технологии извлечения цитат. Сервис будет доступен для пользователей настольных и карманных компьютеров, а также для абонентов сотовой связи. Будет разработана коммерческая версия анализатора для английского языка.

Маркетинг и рынки

1) Объем российского рынка справочных правовых систем вырос с \$ 250 млн. в 2005 году до \$ 500 млн. в 2007 г. Спрос на правовую информацию будет расти в связи с активной законодательной деятельностью.

Указанные системы ориентированы, прежде всего, на средних и крупных корпоративных клиентов. Спрос на правовую информацию со стороны физических лиц и мелких компаний (в т.ч. частных предпринимателей) в настоящее время удовлетворен недостаточно. ООО «Диктум» планирует занять эту нишу путем организации услуги «Электронный адвокат» на основе технологии извлечения цитат. На начальном этапе аудитория будет формироваться на основе бесплатного доступа к ограниченной коллекции документов.

Объем рынка англоязычной правовой информации как минимум на порядок превышает российский, поскольку английский является языком международных отношений, торговли и транспорта.

Объем мирового рынка корпоративных поисковых систем оценивался в величину \$ 1,2 млрд., темпы его развития составляют более 30% в год.

2) В настоящее время рынок систем синтаксического анализа находится в стадии формирования. В ближайшем будущем компании-разработчики лингвистического ПО составят основу платежеспособного спроса на системы анализа текстов, к числу которых относится DictaScore. О существующей потребности в системах анализа текстов свидетельствуют запросы в адрес ООО «Диктум» от разработчиков поисковых систем, систем машинного перевода и систем синтеза речи. Кроме того, анализатор может использоваться в вопросно-ответных системах, системах автореферирования, системах дистанционного обучения и автоматического контроля знаний и других системах автоматической обработки текстов.

В частности, объем мирового рынка переводов между языками составляет около \$ 20 млрд. На долю машинного перевода приходится лишь 1%. Эта доля увеличивается одновременно с ростом качества машинного перевода, которое существенно зависит от методов синтаксического анализа.

Взаимодействие с инвестором

Инвестору будет предложена доля участия в компании «Диктум» - блокирующий пакет (25,1%). Возможно введение представителей инвестора в органы управления компанией.

Примерно через 4 года венчурный инвестор сможет продать свою долю в компании «Диктум» стратегическому инвестору, в роли которого может выступить компания-разработчик поисковой или справочно-правовой системы (например, «Консультант+»). Прогноз цены указанного пакета через 4 года – \$ 5 млн.