

**Программный комплекс автоматизированного синтеза  
экспертного заключения по инновационным проектам (на  
примере проектов программы «СТАРТ»).**

**С.А. Митрофанов,**

первый проректор РГУИТП, д.т.н., профессор, зав. кафедрой «Вычислительная техника» МГИЭМ, зав. кафедрой «Прикладная информатика» РГУИТП.

email: [.Mitrofanov@itbu.ru](mailto:Mitrofanov@itbu.ru)

**А.В. Рождественский,** к.п.н., заместитель руководителя Федерального агентства по образованию

**С.Л. Макаров,**

ст. преп. кафедры «Вычислительная техника» МГИЭМ, соискатель степени к.т.н.

email: [msslmsl@mail.ru](mailto:msslmsl@mail.ru)

В статье рассматривается возможность автоматизации получения экспертного заключения по инновационным проектам, заявки с описанием которых подаются для участия в программе «СТАРТ». Предлагается алгоритм обработки заявок, рассматривается математическая модель, лежащая в основе формализации синтеза экспертного заключения, а также описываются некоторые характеристики и перспективы создаваемого программного комплекса, построенного на основе упомянутого алгоритма.

Ключевые слова: синтез, экспертное заключение, решение

The main idea of the article is a possibility of creating a software which could generate an expert estimation of START program innovation project description stored in a certain structure file. The software algorithm and some features of the software as long as the future of this kind of approach to the problem are also discussed.

Keywords: synthesis, expert opinion, decision

В настоящее время при экспертизе инновационных проектов, рассчитывающих на инвестиции, используется только человеческий труд экспертов. Число заявок по различным программам фондов поддержки инновационных проектов постоянно растёт. В то же время эксперт не всегда может дать стопроцентно качественную оценку того или иного проекта, может быть необъективным, может что-то пропустить или не увидеть, поэтому оценку одного и того же проекта поручают нескольким независимым экспертам, что повышает объективность экспертизы, но и с другой стороны повышает временные и материальные затраты на экспертизу. В связи с этим достаточно актуальной становится задача помощи эксперту и даже его возможной замены на некоторых достаточно формальных этапах рассмотрения проекта. Очевидно, что окончательное решение должен принимать только эксперт и речь идёт о возможности автоматизированного синтеза некоторого предварительного заключения по проекту и об информационной поддержке решений эксперта.

Рассмотрим возможность автоматизации работы эксперта на примере анализа заявок по программе «СТАРТ» государственного Фонда содействия развитию малых форм предприятий в научно-технической сфере. Суть каждого инновационного проекта, претендующего на инвестиции Фонда, должна быть изложена в так называемом файле заявки. Этот файл, в соответствии с требованиями Фонда, имеет чётко определённую структуру. Он состоит из следующих разделов: данные о проекте, интеллектуальная собственность, информация о создаваемом предприятии, участники проекта, научно-техническая часть проекта, коммерциализуемость научно-технических результатов, план действия по реализации проекта, план развития предприятия, характеристика команды (заявителей), аннотация проекта, календарный план выполнения НИОКР (1-й этап проекта), смета затрат на проведение НИОКР 1 этапа проекта, укрупнённый календарный план выполнения 2-го этапа проекта, укрупнённый календарный план выполнения 3-го этапа проекта, календарный план по проекту. Экспертное заключение, формируемое экспертом по этой

заявке, также жестко структурировано и состоит из конкретных вопросов по каждому разделу заявки, ответ на который оценивается экспертом количеством баллов из заданного интервала баллов для данного вопроса, например:

1. Научно-техническая часть проекта

1.1. Теоретические основы, идеи, суть проекта для коммерциализации. Оценка от 0 до 5 баллов.

1.2. Научно-технические риски проекта. Оценка от 0 до 3 баллов.

1.3. Устранение научно-технических рисков проекта. Оценка от 0 до 3 баллов.

1.4. Научно-технические публикации по проблеме. Оценка от 0 до 2 баллов.

1.5. Новизна и аналоги. Оценка от 0 до 3 баллов.

1.6. Интеллектуальная собственность – права и защита. Оценка от 0 до 4 баллов.

Кстати, такая формальная структуризация заявки и экспертного заключения, разработанная Фондом, в значительной степени позволила поставить и решить задачу автоматизированного синтеза экспертного заключения. По каждому из этих разделов эксперт выносит своё мнение – насколько подробно изложена информация, отвечает ли информация раздела на тот вопрос, на которых пытается ответить эксперт, выражая своё мнение в виде текстового комментария и соответствующей количественной оценки этого мнения в баллах. Можно предположить, что эксперт формирует свое заключение на основе некоторой мысленной модели-эталона, существующей в виде набора требований, которым должна соответствовать рассматриваемая им заявка, и чем ближе то, что он видит в заявке, к эталону, тем выше оценка, которую он назначает согласно своим знаниям по каждому разделу.

По такому же принципу может работать и программное обеспечение, занимающееся анализом и проставлением оценок по каждому из разделов файла заявки. Таким образом, чтобы сформировать экспертное заключение, нужно знать эталоны и правила, по которым выстраивается заключение эксперта, причём правила для каждого раздела должны быть свои, а для некоторых может не существовать эталонов – только логические правила. Нами был разработан алгоритм (рис.1), на основе которого возможно построение программного

комплекса, решающего поставленную задачу автоматизации синтеза экспертного заключения.

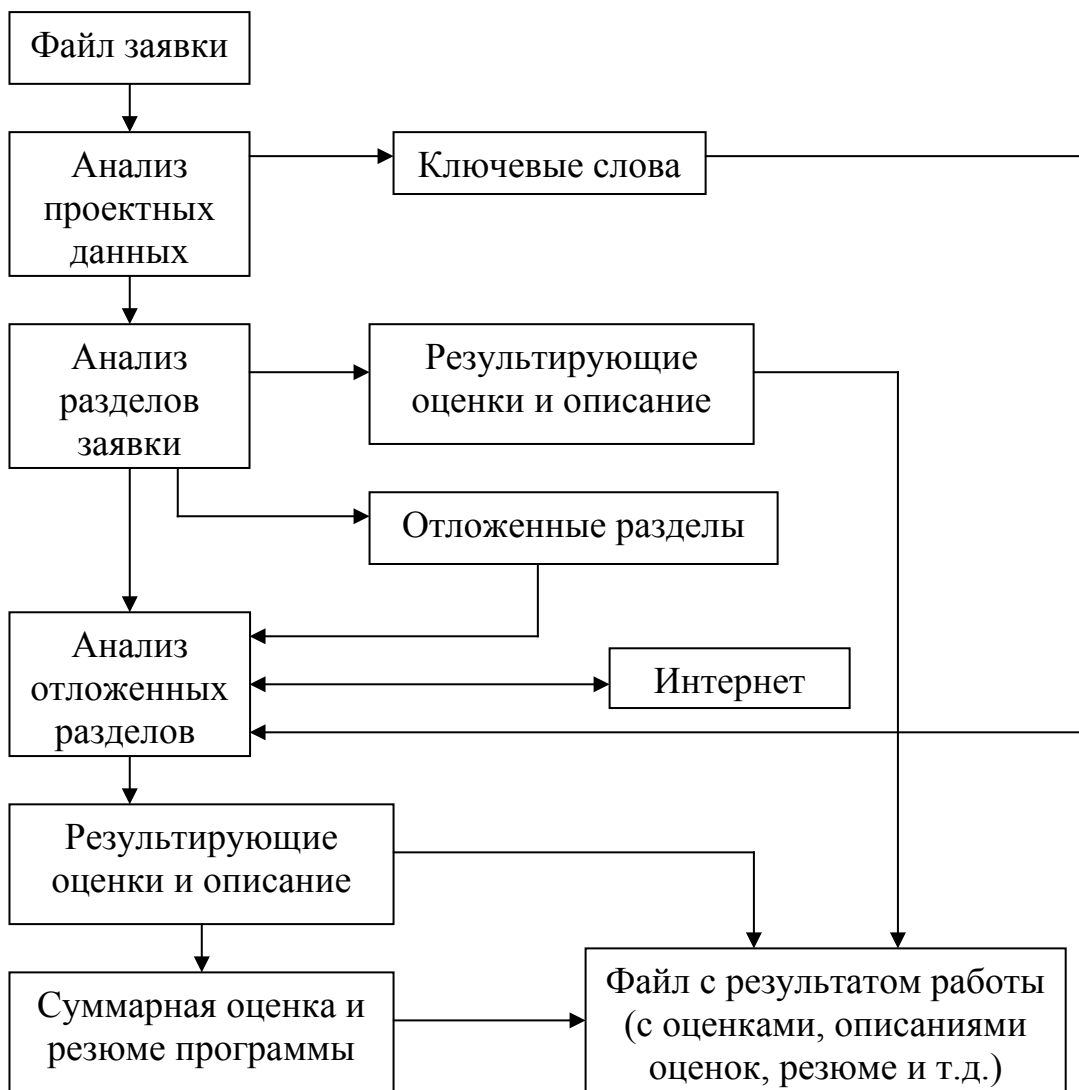


Рисунок 1. Укрупнённая блок-схема алгоритма программы.

Алгоритм состоит из нескольких больших этапов. На первом из них проводится анализ данных о проекте – выделяется полезная информация, которая пригодится в будущем, например – название проекта, ключевые слова, которые, по мнению авторов, однозначно характеризуют именно их проект, сведения об авторах и т.д. Второй этап предполагает анализ каждого раздела, которых в заявке всего пять. На этом шаге алгоритма происходит работа с разделами, не требующими большого времени для обработки программой, а

разделы, не отвечающие этому требованию, запоминаются и их обработка откладывается на будущее. На третьем этапе происходит анализ отложенных разделов. Одним из них является научно-техническое описание идеи проекта и описание аналогов. Схема работы по этим описаниям приведена на рис. 2. Информация об аналогах проекта берётся из соответствующего раздела, а научно-техническая часть анализируется с целью составления терминологического портрета документа (ТПД) и моделей терминологических портретов запросов (ТПЗ) для поискового механизма. Затем по каждой из построенных четырёх моделей ТПЗ, одну из которых формируют выделенные на первом этапе рассматриваемого алгоритма ключевые слова, а три других – составленные определённым образом термины различной длины строится запрос для поискового механизма и происходит поиск конкурентов анализируемого проекта в сети Интернет. Каждый найденный документ запоминается системой, а затем, после окончания поиска, происходит анализ всех найденных документов. Для каждого документа, в свою очередь, строится ТПД, которые сравниваются с ТПД научно-технического раздела. Выделенные в результате такого сравнения наиболее близкие документы обрабатываются не предмет названий продуктов или фирм, описание которых заложено в содержании этих документов, и затем происходит завершающий подэтап этого большого этапа рассматриваемого нами алгоритма – сравнение названий конкурентов и/или фирм, найденных в сети с информацией по конкурентам/аналогам, указанной в соответствующем разделе файла заявки инновационного проекта, в результате которого и формируется оценка. Этот этап алгоритма – самый трудо- и времяёмкий, однако достаточно эффективный по своим результатам, т.к. позволяет оценить новизну инновационного проекта на сегодняшний день, что практически нереально требовать от эксперта.

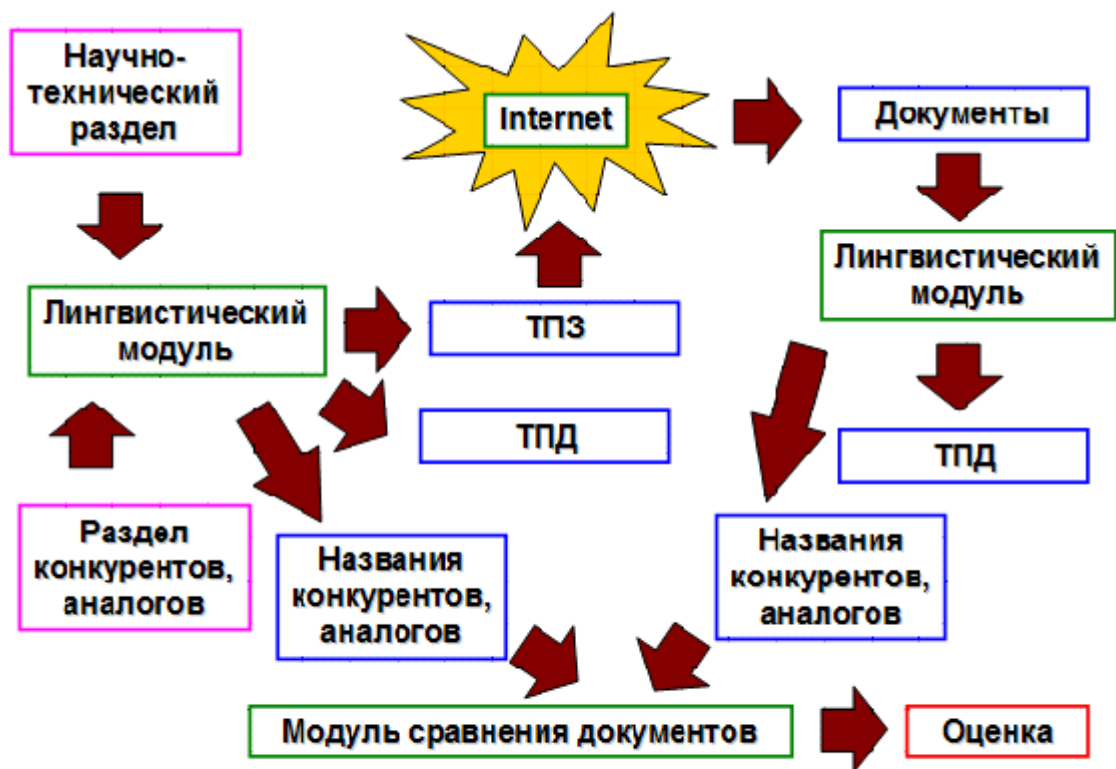


Рисунок 2. Анализ научно-технической новизны проекта.

Математическая модель синтеза экспертного заключения выглядит следующим образом:

$$F(x) = \bigvee_{i=1}^N x_i \oplus t_i,$$

где  $N$  – количество всех оценок по всем разделам,  $x_i$  –  $i$ -я оценка,  $t_i$  – резюме по  $i$ -й оценке (текстовое описание оценки), причём значения  $x_i$  и  $t_i$  могут получаться двумя способами в зависимости от рассматриваемого раздела:

1)

$$t_i = a_i \cap b_i,$$

где  $a_i$  – исходный текст  $i$ -го раздела,  $b_i$  – эталонный текст  $i$ -го раздела, и

$$a_i = \bigvee_{k=1}^M w_k, \quad b_i = \bigvee_{l=1}^O w_l,$$

где  $M$  – количество всех имён существительных  $w_k$  исходном тексте  $i$ -го раздела,  $O$  – количество всех имён существительных  $w_l$  в эталонном тексте  $i$ -го раздела, а

$$x_i \approx \text{round}(|t_i| \left\lfloor \frac{x_{\max}}{|b_i|} \right\rfloor),$$

где  $x_{\max}$  – максимальная оценка по разделу.

Например, возьмём реально существующую заявку, раздел "План действия по реализации проекта". Исходный текст этого раздела следующий:

"Компания «Оптимизирующие Технологии» (создана в 2007 г.) является экспертом в области создания оптимизирующих компиляторов и систем автоматического распараллеливания вычислений;

Нами созданы бета-версии продуктов: Универсальная Библиотека Трансляции (UTL), Анализатор, Авто-Распараллеливатель и Оптимизатор. Демонстрационные дистрибутивы и документация к ним доступны с сайта компании [www.optimitech.com](http://www.optimitech.com);

в настоящее время поддерживаются платформы x86, Power Cell, Itanium. Перечень поддерживаемых платформ постоянно расширяется.

релизы продуктов запланированы на:

UTL полная версия:	январь 2009 г.
UTL Analyser:	январь 2009 г.
UTL Auto-Parallelizer:	январь 2009 г.
UTL Optimizer:	январь 2009 г.

Актуальная задача текущего периода - завершение разработки и продвижение продуктов на рынок."

Эталонный текст для этого раздела следующий:

"Инвестиционные риски"

Рассмотрим формирование оценки и описания. В бланке эксперта эта оценка называется следующим образом:

"3.5 Имеются ли неопределённые моменты (риск в плане действий), намерения по их уменьшению? Оценка (от 0 до 3):"

$t_i$  в данном случае будет формироваться следующим образом: так как существительное "риск" не упоминается в исходном тексте раздела, множества  $a_i$  и  $b_i$  не пересекаются, то  $t_i = ""$ , однако для пользователя пустая строка не информативна, поэтому  $t_i$  присваивается значение

$t_i = \text{"Риски в плане действий по реализации проекта не приведены."}$

$$x_i \approx \text{round}(|t_i| \left\lfloor \frac{x_{\max}}{|b_i|} \right\rfloor) \approx \text{round}\left(0 \left\lfloor \frac{3}{1} \right\rfloor\right) \approx 0$$

2)

$$x_i \models r_i,$$

где  $r_i$  – некоторое логическое правило, которому должно удовлетворять определённое соотношение данных (обычно численных) из разных частей  $i$ -го раздела, а

$$t_i \models f(x_i).$$

Например, из той же заявки рассмотрим получение следующей оценки:

"4.5 обоснован ли уровень заработной платы сотрудников малого предприятия? Оценка (от 0 до 3):"

Правило  $r_i$  для этой оценки (неполное, так как их много) звучит следующим образом:

"1-й год проекта – может быть и 0, второй год – больше, чем первый, 3-й – больше, чем второй, в итоге – больше, чем в каждом году (таким образом зарплата повышается из года в год), итоговая зарплата > \$20 000, зарплата второго года, третьего года и итоговая не может быть равна 0"

В качестве входных данных имеем следующую таблицу:

Выработка на сотрудника

1-й год проекта (руб./чел)	2-й год проекта (руб./чел)	3-й год проекта (руб./чел)	Итого (руб./чел)
2 974 583	0	0	1 115 469

После применения правила получаем:

$$x_i \models 0$$

$t_i =$  "Некорректное развитие фирмы (зарплата за второй год, третий год или итоговая зарплата не может быть равна 0)."

В ходе анализа программа выставляет не только оценки, но и даёт заключения и рекомендации, чего именно не хватает в отдельных элементах заявки, сравниваясь с эталоном, или указывает на ошибки, допущенные при заполнении, например, таблиц. В итоге все замечания и предложения отражаются в резюме, рядом с итоговой оценкой.

В настоящее время создаётся программный комплекс, работающий по рассмотренному алгоритму. В настоящее время программный комплекс



проходит опытное тестирование на заявках программы «СТАРТ», можно привести пример экспертного заключения, синтезированного разрабатываемым программным комплексом:

Научно-техническая часть проекта

1.1. Теоретические основы, идеи, суть проекта для коммерциализации

Оценка (от 0 до 5): 3

Идея проекта не нова, однако показаны преимущества перед основными конкурентами.

1.2. Научно-технические риски проекта

Оценка (от 0 до 3): 3

Научно-технические риски определены.

1.3. Устранение научно-технических рисков проекта

Оценка (от 0 до 3): 2

Перечень НИОКР, приведённый в заявке, недостаточно соответствует решению научно-технических проблем (отсутствует информация по следующим проблемам: прибыль, возрастание объёма продаж).

1.4. Научно-технические публикации по проблеме

Оценка (от 0 до 2): 1

Степень компетентности заявителей отражена не полностью (есть только публикация).

1.5. Новизна и аналоги.

Оценка (от 0 до 3): 3

Новизна и ближайшие аналоги определены в заявке.

1.6. Интеллектуальная собственность – права и защита.

Оценка (от 0 до 4): 3

Интеллектуальная собственность не защищена, но заявители ясно представляют защищаемые объекты интеллектуальной собственности. Права будут зарегистрированы на фирму.

Суммарная оценка по разделу: 14

2. Коммерциализуемость научно-технических результатов

## 2.1. Перспективность коммерциализации результатов

Оценка (от 0 до 4): 3

В заявке приведена недостаточно весомая аргументация, почему предлагаемый результат может стать основой бизнеса (отсутствует информация по следующим проблемам: прибыль, возрастание объёма продаж).

## 2.2. Обоснованность стратегии коммерциализации

Оценка (от 0 до 4): 4

Стратегия коммерциализации обоснована.

## 2.3. Анализ потребностей рынка

Оценка (от 0 до 4): 3

Уровень анализа потребностей рынка недостаточно высокий (отсутствует информация по следующим проблемам: потребность).

## 2.4. Конкуренты на рынке

Оценка (от 0 до 4): 4

Конкурентные (альтернативные) продукты и производители, конкурентные преимущества продукта заявителей приведены.

## 2.5. Риски коммерциализации

Оценка (от 0 до 4): 0

Уровень оценок рисков коммерциализации низкий.

Суммарная оценка по разделу: 14

## 3. План действий по реализации проекта

### 3.1. План действий по реализации проекта

Оценка (от 0 до 5): 4

Этапы работы представлены не полностью (не освещены ниокр, этап реализации продукции).

### 3.2. Работа с инвестором

Оценка (от 0 до 3): 2

Работа с инвестором и продвижение продукта на рынок представлены, однако отсутствует информация по следующим проблемам: объёмы финансирования инвестором.

### 3.3. План по НИОКР

Оценка (от 0 до 3): 2

Смета и календарный план НИОКР обоснованы не полностью (отсутствует информация по следующим проблемам: год выхода на рынок).

### 3.4. Интеллектуальная собственность

Оценка (от 0 до 3): 3

Интеллектуальная собственность не защищена, но защита планируется. Права будут зарегистрированы на фирму.

### 3.5. Риски в плане

Оценка (от 0 до 3): 0

Риски в плане действий по реализации проекта не приведены.

### 3.6. Организация управления проектом

Оценка (от 0 до 3): 1

Организация управления проектом представлена, но не полностью (отсутствует информация по следующим проблемам: управление, менеджмент, директор, собственник).

Суммарная оценка по разделу: 12

## 4. План развития предприятия

### 4.1. Планирование 3-го года развития предприятия

Оценка (от 0 до 5): 4

Темпы развития малого предприятия и достигаемый уровень капитализации на третий год действия заказа представлены не в полном объеме.

4.2. Объем коммерциализации по величине годового оборота и выработки на одного работающего

Оценка (от 0 до 3): 3

В заявке приведены объемы коммерциализации по величине годового оборота и выработки на одного работающего к окончанию срока действия контракта.

### 4.3. Реальные сроки окупаемости

Оценка (от 0 до 3): 3

Сроки окупаемости работы реальные.

4.4. Соотношение обоснованных объемов реализации продукции и запрошенных бюджетных средств

Оценка (от 0 до 3): 3

Объем реализации продукции по проекту значительно превышает объем запрошенных бюджетных средств (бюджетные средства составляют 20% и менее от средств реализации продукции).

4.5. Обоснованность уровня заработной платы

Оценка (от 0 до 3): 0

Уровень заработной платы сотрудников малого предприятия не обоснован (зарплата менее 2 000 рублей в месяц!).

4.6. Появление новой интеллектуальной собственности в результате выполнения проекта

Оценка (от 0 до 3): 3

В результате выполнения проекта будет создана новая интеллектуальная собственность. Права будут зарегистрированы на фирму.

Суммарная оценка по разделу: 16

5. Характеристика команды заявителей

5.1. Уровень квалификации и научно-технический потенциал

Оценка (от 0 до 5): 3

Уровень квалификации и научно-технический потенциал участников конкурса не очень высокий (докторов - 0, кандидатов - 3, в том числе профессоров - 0, доцентов - 0).

5.2. Влияние демографических данных на реализацию проекта

Оценка (от 0 до 3): 3

Демографические данные членов коллектива никак не могут повлиять на реализацию проекта (города и регионы более половины участников проекта совпадают).

5.3. Опыт работы с инновационными проектами

Оценка (от 0 до 3): 1

Опыт работы с инновационными проектами имеется у половины коллектива.

#### 5.4. Информация по переходу заявителей на работу в малое предприятие

Оценка (от 0 до 3): 3

Информация по переходу заявителей на работу в малое предприятие дана полностью (должности: директоров - 1, сотрудников - 11, в том числе на постоянной основе - 12).

#### 5.5. Уровень квалификации директора предприятия

Оценка (от 0 до 3): 3

Выбор директора малого предприятия обоснован (кандидат наук, директор или зам. директора на уже существующем предприятии).

#### 5.6. Уровень качества представления материалов заявки на конкурс

Оценка (от 0 до 3): 3

Все разделы представлены.

Суммарная оценка по разделу: 16

Суммарная оценка по всем разделам: 72 из 100 возможных.

Резюме программы: В заявке есть недочёты; несмотря на то, что представлены все разделы, не приведены риски коммерциализации, не учтены риски в плане действий по реализации проекта; кроме того, не продумана стратегия развития предприятия с точки зрения заработной платы сотрудников. Очевидно, что такое экспертное заключение не может являться основанием для окончательного принятия инвестиционного решения по данному инвестиционному проекту, но вполне обеспечивает объективную информационную поддержку личного собеседования членов экспертного совета программы "Старт" с авторами этого инновационного проекта.

Опытное тестирование программного комплекса показало, что слабым местом разработки являются большие затраты машинного времени на синтез экспертного заключения, синтез экспертного заключения для 50 заявок по программе «СТАРТ» позволил получить следующую оценку затрат машинного времени (таблица 1):

№	Раздел экспертного заключения.	Затраты машинного времени на формирование оценки по разделу (мин.).
1	1.1. Теоретические основы, идеи, суть проекта для коммерциализации	20,12 <sup>1</sup>
2	1.2. Научно-технические риски проекта	0,11
3	1.3. Устранение научно-технических рисков проекта	0,11
4	1.4. Научно-технические публикации по проблеме	0,11
5	1.5. Новизна и аналоги.	0,11
6	1.6. Интеллектуальная собственность – права и защита.	0,12
7	2.1. Перспективность коммерциализации результатов	0,12
8	2.2. Обоснованность стратегии коммерциализации	0,12
9	2.3. Анализ потребностей рынка	0,13
10	2.4. Конкуренты на рынке	0,12
11	2.5. Риски коммерциализации	0,12
12	3.1. План действий по реализации проекта	0,11
13	3.2. Работа с инвестором	0,10
14	3.3. План по НИОКР	0,11
15	3.4. Интеллектуальная собственность	0,10
16	3.5. Риски в плане	0,10
17	3.6. Организация управления проектом	0,10
18	4.1. Планирование 3-го года развития предприятия	0,04
19	4.2. Объем коммерциализации по величине годового оборота и выработки на одного работающего	0,04
20	4.3. Реальные сроки окупаемости	0,05
21	4.4. Соотношение обоснованных объемов реализации продукции и запрошенных бюджетных средств	0,04
22	4.5. Обоснованность уровня заработной платы	0,04
23	4.6. Появление новой интеллектуальной собственности в результате выполнения проекта	0,05
24	5.1. Уровень квалификации и научно-технический	2,1

<sup>1</sup> Оценка получена при следующих настройках программы: Максимальный объем сохраняемого файла: 10 Кб, учитывать 1 страницу результата поиска системой Yandex, анализировать одно-, двух- и трёхсловные термины, количество терминов в запросе (всего 4 запроса): 10, время ожидания загрузки страницы из Интернета: 10 сек. Скорость подключения к Интернету: 5760 Кбит/с (D).

	потенциал	
25	5.2. Влияние демографических данных на реализацию проекта	2
26	5.3. Опыт работы с инновационными проектами	1
27	5.4. Информация по переходу заявителей на работу в малое предприятие	2,1
28	5.5. Уровень квалификации директора предприятия	1
29	5.6. Уровень качества представления материалов заявки на конкурс	0,01
	ИТОГО:	30,38

Таблица 1. Затраты машинного времени на получение оценок.<sup>2</sup>

Ещё одна особенность программного комплекса, обеспечивающая его гибкость – множество настроек, которые позволяют учитывать детали анализа, строить меньше или больше моделей запросов для поискового механизма, учитывать только определённые весовые характеристики терминов, просматривать определённое количество страниц, выданных поисковым механизмом как результат поиска по сформированным моделям запросов и т.д. Эти настройки позволяют варьировать глубину анализа научно-технической составляющей проекта, также как и прочих разделов файла заявки, а, следовательно, сократить время работы программы, требуемое для вынесения окончательного экспертного заключения, что позволяет получить это заключение значительно быстрее экспертов, работающих над заявкой параллельно.

Для работы с программой требуется открыть нужный файл заявки инновационного проекта, при этом его текст появляется в окне программы, и затем нажать на кнопку старт главного меню программы – начнётся анализ файла. Через несколько минут появится окно с результатом, которое будет обновляться, пополняясь оценками по проанализированным разделам и

<sup>2</sup> Прогон программы проводился на компьютере AMD Athlon(tm) Dual Core Processor 4450e 2.30 ГГц, 2.00 Гб ОЗУ, второй процессор не был занят параллельными задачами.

текстовым описанием/обоснованием оценок. Третья строчка (научно-техническая новизна проекта), содержащая самую трудоёмкую оценку, будет содержать процент проделанной работы, по достижении 100% в ней появится окончательная оценка по разделу научно-технической новизны. Существуют настройки программы, доступные из соответствующего пункта главного меню, но для работы программы достаточно настроек по умолчанию.

Авторы понимают, что для повышения эффективности программы необходимо оптимизировать алгоритм выборки терминов из файлов, найденных по совпадению их терминологических портретов с портретом раздела научно-технической новизны проекта, а также основной алгоритм программы. Кроме того целесообразно повысить интеллектуальность принимаемых решений за счёт введения дополнительных правил коррекции оценки.

Одним из направлений совершенствования программы является выведение всех правил и эталонов за пределы программного кода, что позволит менять правила в соответствии с некими рекомендациями в независимых файлах правил, хранящихся в отдельном каталоге, не прибегая к изменению программного кода. Кроме того, авторы рассчитывают, что результаты опытной эксплуатации программного комплекса позволят уточнить требования, предъявляемые к содержанию, структуре и форме информации входных заявок, что позволит повысить интеллектуальность, объективность и эффективность машинного формирования экспертного заключения.

Уже полученный опыт показывает, что создание программного комплекса, решающего задачу автоматизации синтеза экспертной оценки инновационного проекта программы «СТАРТ» является реальной и выполнимой задачей. Этот комплекс может значительно упростить работу эксперта и повысить скорость вынесения и качество принимаемых решений по рассматриваемым проектам.

Модуль сравнения терминологических портретов двух документов, зарегистрирован в Роспатенте в сентябре 2008 года, где планируется регистрация всего программного комплекса после завершения этапа тестирования.