



Общероссийский математический портал

В. А. Минин, И. М. Зацман, М. Г. Кружков, Т. П. Норемян, Методологические основы создания информационных систем для вычисления индикаторов тематических взаимосвязей науки и технологий, *Информ. и её примен.*, 2013, том 7, выпуск 1, 70–81

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением
<http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 18.117.230.165

19 ноября 2024 г., 21:29:56



МЕТОДОЛОГИЧЕСКИЕ ОСНОВЫ СОЗДАНИЯ ИНФОРМАЦИОННЫХ СИСТЕМ ДЛЯ ВЫЧИСЛЕНИЯ ИНДИКАТОРОВ ТЕМАТИЧЕСКИХ ВЗАИМОСВЯЗЕЙ НАУКИ И ТЕХНОЛОГИЙ*

В. А. Минин¹, И. М. Зацман², М. Г. Кружков³, Т. П. Норекиян⁴

Аннотация: Анализируется зарубежный опыт вычисления индикаторов тематических взаимосвязей науки и технологий. Цель анализа заключается в разработке принципов создания отечественных информационных систем для вычисления индикаторов взаимосвязей с учетом исторически сложившейся в нашей стране структуры наследуемых научных и патентных информационных ресурсов. Этот вид информационных систем является новым для российской научно-технической сферы. Их создание необходимо для мониторинга и оценивания программ научных исследований и принятия решений на всех этапах программной деятельности. В статье предлагается методология определения индикаторов тематических взаимосвязей в отечественной научно-технической сфере как основа создания информационных систем, предназначенных для вычисления их значений.

Ключевые слова: взаимосвязи науки и технологий; классификация научных направлений; международная патентная классификация; рубрицирование научных документов

1 Введение

1 ноября 2012 г. на заседании Правительства РФ был рассмотрен проект «Государственной программы РФ «Развитие науки и технологий» на 2013–2020 годы», включающей шесть подпрограмм. Для подпрограммы номер 2 «Прикладные проблемно-ориентированные исследования и развитие научно-технического задела в области перспективных технологий» предлагается установить два следующих индикатора⁵, характеризующих достижение цели Государственной программы⁶ в части проблемно-ориентированных исследований:

- (1) коэффициент изобретательской активности (число отечественных патентных заявок на изобретения, поданных в России в расчете на 10 тыс. чел. населения);
- (2) число патентных заявок на изобретения, поданных отечественными заявителями в России

из организаций — участников Государственной программы [1, с. 48].

Из названий этих двух индикаторов следует, что в процессе оценивания научно-технического задела в области перспективных технологий в этой программе учитывается только изобретательская активность. Второй, но не менее важный аспект, а именно: цитируемость в описаниях изобретений публикаций, являющихся непосредственными результатами научных исследований и разработок, — не нашел своего отражения в списке индикаторов этой программы. Здесь важно отметить, что современные методологии, реализуемые в информационных системах индикаторного оценивания процессов трансформации знаний в новые технологии, состоят из двух основных компонентов [2]:

- (1) индикаторное оценивание тематических взаимосвязей результатов исследований и разработок с технологической модернизацией (что да-

* Работа выполнена при поддержке РГНФ, грант № 12-02-12019в.

¹ Российский фонд фундаментальных исследований, minin@rfr.ru

² Институт проблем информатики Российской академии наук, iz_ipi@al70.ipi.ac.ru

³ Институт проблем информатики Российской академии наук, magnit75@yandex.ru

⁴ Институт проблем информатики Российской академии наук, izzittami@gmail.com

⁵ Финансовые индикаторы программы, например удельный вес внебюджетных средств во внутренних затратах на исследования и разработки, в статье не рассматриваются.

⁶ Целью всей программы является «формирование конкурентоспособного и эффективно функционирующего сектора исследований и разработок и обеспечение его ведущей роли в процессах технологической модернизации российской экономики» [1, с. 9].

ет возможность получать экспертные оценки инновационно-технологического потенциала направлений научных исследований);

- (2) индикаторное оценивание изобретательской активности в процессе исследований и разработок (как правило, оценивание ведется по направлениям технологического развития).

Таким образом, в Государственной программе РФ «Развитие науки и технологий» на 2013–2020 гг. используется только второй компонент, но полностью отсутствует первый. Одна из причин заключается в том, что в нашей стране отсутствуют те информационные системы, которые могли бы вычислять индикаторы взаимосвязей научных направлений и технологий, сопоставляя накопленные научные и патентные информационные ресурсы [3, 4].

Возможно, по этой причине разработчики программы и ограничились традиционным индикаторным оцениванием изобретательской активности в процессе исследований и разработок в рамках этой программы, хотя имеющийся зарубежный опыт, описанию которого посвящен третий раздел статьи, свидетельствует о ключевой роли количественных индикаторов взаимосвязей для многоаспектного оценивания научно-технического задела в области перспективных технологий в интересах принятия решений на всех этапах программной деятельности.

Для изменения сложившейся ситуации Российский гуманитарный научный фонд (РГНФ) начал в 2012 г. финансирование инициативного проекта по гранту № 12-02-12019в с целью создания экспериментального образца информационной системы, предназначенного для индикаторного оценивания взаимосвязей результатов научных исследований с развитием информационных технологий (ИТ). Основной целью статьи является описание имеющегося задела и полученных в 2012 г. результатов этого проекта.

2 Проблемы программно-целевого планирования в научно-технической сфере

В течение 2005–2011 гг. сформировалось новое направление исследований «Методы и технологии

информационного мониторинга в научно-технической сфере». Результаты, полученные специалистами ЦЭМИ РАН и ИПИ РАН, опубликованы в подразд. 5.4 монографии «Мезоэкономика развития» [5], двух книгах [6, 7] и 19 статьях [2–4, 8–23], в том числе в трех зарубежных публикациях. В этих работах была сформулирована проблема информационного мониторинга, анализа и оценивания научной деятельности, в том числе программ научных исследований¹ (далее по тексту — проблема мониторинга), а также оценивания влияния их результатов на социально значимые сферы деятельности, в том числе на создание новых технологий.

В процессе постановки этой проблемы учитывались методические документы и нормативно-правовые акты, предписывающие переход к программно-целевым методам бюджетного планирования, ориентированным на повышение результативности использования бюджетных расходов. В частности, постановление Правительства РФ от 22 мая 2004 г. № 249 «О мерах по повышению результативности бюджетных расходов» (далее — постановление № 249) предусматривает переход на среднесрочное бюджетирование, ориентированное на результаты (СБОР), во всех социально значимых сферах деятельности, в том числе в научно-технической сфере. Основные принципы СБОР сформулированы в Концепции реформирования бюджетного процесса в РФ в 2004–2006 гг., одобренной постановлением № 249.

С точки зрения разработки информационных систем мониторинга ключевым методическим документом служат рекомендации по подготовке Докладов о результатах и основных направлениях деятельности субъектов бюджетного планирования на 2006–2008 гг. [24]. Согласно этим рекомендациям используемые системы индикаторов и показателей должны соответствовать следующим 9 требованиям, которые были уточнены в процессе постановки проблемы мониторинга следующим образом:

- (1) *адекватность*: показатель (индикатор²) должен характеризовать прогресс в достижении цели или решении задачи; используемые комплексы показателей (индикаторов) должны охватывать *все существенные аспекты достижения цели и/или решения задач ПД*;
- (2) *точность*: погрешности измерения не должны приводить к искаженному представлению о результатах, эффективности и результативности ПД;

¹ Далее наряду со словосочетанием «программа научных исследований» будет использоваться термин «программная деятельность» (кратко — ПД). Отличие ПД от программ заключается в том, что кроме периода действия программы рассматривается стадия формирования и стадия оценивания результатов программы после ее завершения.

² Термин «показатель» используется в этой статье как родовой по отношению к термину «индикатор». Более подробно вопрос о соотношении этих терминов, включая их дефиниции, рассмотрен в работе [10].

- (3) *объективность*: не допускается использование показателей, улучшение отчетных значений которых возможно при ухудшении реального положения дел; используемые показатели должны в наименьшей степени создавать стимулы к искажению результатов ПД;
- (4) *достоверность*: способ сбора и обработки исходной информации должен допускать возможность проверки точности как собранных, так и обработанных данных в процессе независимого мониторинга и оценивания ПД;
- (5) *однозначность*: определение показателя (индикатора) должно обеспечивать *одинаковое понимание существа измеряемой характеристики*;
- (6) *экономичность*: получение отчетных данных должно производиться с минимально возможными затратами, применяемые показатели должны в максимальной степени основываться на уже существующих процессах сбора информации о ПД;
- (7) *сопоставимость*: выбор показателей следует осуществлять исходя из необходимости непрерывного накопления данных и обеспечения их сопоставимости за отдельные периоды с показателями, используемыми для оценки прогресса в решении сходных (смежных) задач, а также с показателями, используемыми в международной практике;
- (8) *своевременность и регулярность*: отчетные данные о ПД должны поступать со строго определенной периодичностью и с незначительным временным лагом между моментом сбора информации и моментом использования результатов ее обработки;
- (9) *уникальность*: показатели достижения цели ПД не должны представлять собой объединение нескольких показателей, характеризующих решение отдельных относящихся к этой цели задач ПД.

Перечисленные требования представляют собой в совокупности методический фактор, который приобретает особую актуальность для оценивания ПД в тех случаях, когда на основе значений индикаторов планируется формировать и реализовывать стратегию в сфере науки, в том числе распределять бюджетные средства по фундаментальным и технологически ориентированным научным направлениям. С позиции разработчиков информационных систем и технологий мониторинга, с помощью которых должны вычисляться значения индикаторов

¹См. также Акт Конгресса США о результатах и результативности государственного управления (Government Performance and Results Act of 1993), принятый 5 января 1993 г.

ПД, отсутствуют необходимые теоретические основы их разработки, учитывающие *все 9 требований, начиная с требования адекватности*.

Лакуны в современной системе знаний о мониторинге ПД анализировались в отечественных и зарубежных исследованиях неоднократно [22, 23, 25, 26], а результаты этих исследований обсуждались на семинарах.

Например, в материалах семинара по методическим вопросам оценивания федеральных научно-исследовательских программ США, состоявшегося 4–5 декабря 2003 г., приведена сводная таблица из 24 нерешенных задач [25]. Решение некоторых из них требует проведения фундаментальных и прикладных научных исследований.

В эту таблицу включены следующие две проблемы, иллюстрирующие необходимость дальнейшего развития теоретических основ мониторинга:

- (1) отсутствуют модели и методы мониторинга, анализа и оценивания научно-исследовательских программ, обеспечивающие проверку точности данных, используемых для оценки результативности этих программ, т. е. не выполняется требование **достоверности**;
- (2) между экспертами Административно-бюджетного управления и федеральных агентств США нередко возникают конфликты из-за различного понимания смысла индикаторов результатов, эффективности реализации и результативности научно-исследовательских программ (не выполняется требование **однозначности**).

Ряд актуальных задач мониторинга и оценивания ПД был рассмотрен в процессе международной экспертизы итогов мониторинга и оценивания рамочных программ ЕС [26]. Лакуны в системе знаний о мониторинге ПД и, как следствие, **неадекватность** комплексов индикаторов, используемых для оценивания конкретных программ, стали предметом анализа и в отечественных исследованиях. Был проведен анализ комплекса индикаторов одной из отечественных научных программ и показана неполнота используемого в этой программе набора индикаторов результатов, эффективности и результативности [23]. В процессе анализа рассматривались три основные категории результатов ПД [22, 24, 26]¹:

- (1) *непосредственные результаты* исследовательских программ (outputs), например публикации и доклады по итогам выполнения проек-

тов программ (к непосредственным результатам относится только факт публикации или выступления с докладом, но не их содержание);

- (2) *целевые ожидаемые* и фактически полученные *конечные результаты* (outcomes), которые могут быть заданы как для программы в целом, так и для отдельных ее тематических направлений исследований и проектов (к целевым результатам относятся запланированные итоги программ в целом, например увеличение доли молодых ученых среди исследователей, и/или содержательные итоги работ тематических направлений и проектов программ, например выявление закономерностей распространения некоторого вида организмов или синтез нового вещества);
- (3) социально-экономические, технологические и иные изменения, связанные с итогами реализации проектов программ (impacts & linkages), которые представляют собой эффект от применения (влияния) непосредственных или целевых результатов на развитие как самой сферы науки, так и других социально значимых сфер деятельности общества — развитие сфер образования и здравоохранения, технологического и экономического развития, модернизации правовой сферы и т. д. (далее по тексту — *результаты влияния*).

Было показано, что анализируемый комплекс индикаторов оценивания программы не содержит индикаторов результатов влияния [23]. В результате проведенного анализа были сформулированы две новые задачи:

- разработка методов и моделей проектирования новых индикаторов ПД с целью повышения степени адекватности комплексов индикаторов, используемых для оценивания конкретных программ;
- создание информационных систем для вычисления уже спроектированных индикаторов взаимосвязей науки и технологий.

Что касается первой задачи, то методы проектирования новых индикаторов были рассмотрены в работах [27, 28]. Эти методы были разработаны на основе семиотических моделей [29–34]. В данной статье предлагается подход к решению второй задачи. При этом предполагается, что некоторые индикаторы взаимосвязей науки и технологий уже спроектированы и разработчики информационных систем мониторинга могут использовать описания этих индикаторов в процессе разработки технологий вычисления их значений. Это предположение

обосновано тем, что за рубежом исследования взаимосвязей науки и технологий ведутся уже не один десяток лет (см. следующий раздел статьи) и ряд индикаторов спроектирован и опробован на практике. Однако это не означает, что проектирование новых индикаторов оценивания взаимосвязей перестало быть актуальной задачей.

3 Взаимосвязи науки и технологий

Один из наиболее сложных аспектов в исследовании взаимосвязей науки и технологий состоит в индикаторном оценивании процессов переноса знаний, в том числе отраженных в научных публикациях, из разных областей исследований в сферу технологического развития. Причина пристального внимания к индикаторному оцениванию процессов переноса знаний заключается в том, что финансирование научных исследований, ориентированных на развитие научно-технического задела в области перспективных технологий, связано, с одной стороны, с большим риском. С другой стороны, есть риск упустить новые прорывные решения и потерять конкурентоспособность [35]. Это касается конкурентоспособности как отдельных предприятий в конкретной технологической сфере, так и государства в целом [36].

Поэтому и возникла потребность в решении задачи оценивания взаимосвязей науки и технологий в процессе создания информационных систем мониторинга как инструментов обеспечения стратегического планирования и управления финансированием исследований, ориентированных на технологическое развитие. При решении этой задачи один из самых сложных вопросов заключается в том, как зафиксировать факт передачи и использования в технологической сфере результатов научных исследований [35].

В силу указанных причин и стали разрабатываться индикаторы оценивания взаимосвязей, а также методы и алгоритмы определения их значений. Один из предложенных подходов заключается в том, что процессы передачи знаний от науки к технологиям отслеживаются с помощью научных публикаций, цитируемых экспертами в отчетах о патентном поиске и/или авторами изобретений в их описаниях, что конечно не отражает весь спектр взаимосвязей науки и технологий.

По мнению Тийссена с соавторами, основная сложность в отслеживании всего спектра заключается в том, что явные процессы переноса знаний являются многочисленными, а неявные — труд-

но идентифицируемыми [37]. Тем не менее в процессе исследования этой проблемы за последние 30 лет наблюдается явный прогресс, так как был предложен ряд индикаторов, характеризующих взаимосвязи науки и технологий на макроуровне (определяется интегральная интенсивность цитирования результатов всех тех научных исследований, которые связаны с развитием технологий, см. табл. 1) [38] и более детальные взаимосвязи отдельных научных дисциплин и видов технологий [39].

В 1985 г. в процессе сопоставительного анализа статей по биологии и библиографических ссылок на статьи в массивах описаний изобретений по биотехнологиям были экспериментально зафиксированы взаимосвязи между научными публикациями и развитием биотехнологий. В процессе анализа различались патентные ссылки, т. е. ссылки на ранее выданные патенты, и непатентные ссылки, среди которых выделялись ссылки на статьи из журналов, включенных в Указатель научного цитирования (Science Citation Index или SCI), и ссылки на статьи из других журналов. Было обработано более 6500 патентных и непатентных ссылок, извлеченных из описаний 399 патентов [39].

Позже, в конце прошлого века, было проведено широкомасштабное исследование взаимосвязей науки и технологий с использованием более одного миллиона непатентных ссылок, извлеченных из десятилетнего массива описаний изобретений США и Европейского патентного ведомства (ЕПВ). Было показано экспериментально, что *75% научных статей*, цитируемых в этих массивах описаний изобретений по широкому спектру технологий, были подготовлены по результатам, полученным в *некоммерческом секторе научной сферы*. Это дало возможность авторам исследования сделать вывод о сильной зависимости технологического развития от степени государственной поддержки науки в США и европейских странах [40].

В конце прошлого века Мэнсфилд провел серию экспериментальных исследований, взяв в качестве исходных данных сведения об итогах работы ряда американских компаний. Его работы содержат эмпирические данные о другом аспекте взаимосвязей результатов научных исследований и технологических инноваций. На основе анализа этих данных он пришел к выводу, что *10% технологических новшеств* не были бы изобретены или же были бы созданы с большой задержкой, если бы они были сделаны без использования результатов соответствующих академических исследований [41–43].

Один из основных вызовов в современных исследованиях взаимосвязей науки и технологий состоит в том, чтобы разработать методологию количественного и качественного оценивания этих

взаимосвязей как для целых областей знаний, так и для отдельных направлений исследований. Даже при наличии качественных экспертных оценок вклада тех или иных научных направлений в технологическое развитие все равно необходимо вычислять количественные индикаторы взаимосвязей с целью верификации качественных оценок.

Проведенный анализ зарубежного опыта показывает, что вычисление индикаторов взаимосвязей науки и технологий требует автоматизированной обработки больших объемов слабоструктурированных полнотекстовых описаний изобретений и сопоставления результатов обработки патентной информации с названиями источников научных публикаций (журналов или материалов конференций) и с названиями самих статей, хранящихся в научных электронных библиотеках. Это сопоставление дает возможность определить тематику того научного направления, к которому относятся научные публикации, цитируемые в описаниях изобретений. В процессе обработки тысяч и миллионов описаний изобретений используются сложные алгоритмы парсинга, с помощью которых анализируются патентные тексты, в которых выделяются и структурируются библиографические ссылки на научные публикации [38].

4 Информационные ресурсы для определения взаимосвязей

Для сопоставления результатов обработки описаний изобретений с научной информацией используются те патентные информационные ресурсы, в которых цитируются научные публикации. После идентификации каждой научной публикации в описании изобретения определяется источник публикации и ее тематика, что и дает возможность устанавливать тематические взаимосвязи между научными направлениями и технологиями.

Тематика каждого научного направления задается в виде одной или нескольких рубрик выбранной системы классификации областей знаний. Тематика каждой анализируемой технологии задается в виде списка рубрик Международной патентной классификации (МПК). Наиболее часто используются списки рубрик МПК из номенклатуры, разработанной Фраунгоферовским институтом системотехники и инновационных исследований (Fraunhofer Gesellschaft-Institute für Systemtechnik und Innovationsforschung — FhG-ISI). Примеры списков рубрик МПК из номенклатуры FhG-ISI приведены в табл. 1 для пяти видов технологий. Последний столбец содержит данные интегральной

Таблица 1 Коды МПК и интегральные интенсивности цитирования для различных технологий

Название технологий	Коды МПК	Интегральная интенсивность цитирования
Биотехнологии	C07G; C12M, N, P, Q, R, S	138,43
Фармацевтические	A61K	83,71
Полупроводниковые	H01L	56,44
Оптические	G02; G03B, C, D, F, G, H; H01S	21,89
Информационные	G06; G11C; G10L	20,39

интенсивности цитирования результатов тех научных исследований, которые связаны с развитием технологий, указанных в первом столбце. Интегральная интенсивность цитирования определена в табл. 1 как число цитируемых научных публикаций на 100 описаний изобретений [38].

В номенклатуре FhG-ISI предметная область ИТ, для которой планируется разработать экспериментальный образец информационной системы, предназначенной для индикаторного оценивания тематических взаимосвязей, описывается следующими тремя рубриками МПК (см. табл. 1):

G06 — «Вычисление; счет» (эта рубрика МПК включает оптические вычислительные устройства, обработку цифровых данных с помощью компьютеров, аналоговые и гибридные компьютеры);

G11C — «Запоминающие устройства статического типа»;

G10L — «Анализирование или синтезирование речи; распознавание речи».

Таким образом, экспериментальный образец информационной системы должен содержать патентные информационные ресурсы по трем рубрикам МПК: G06, G11C и G10L. Используя эти ресурсы, система должна будет по запросам экспертов определять индикаторы тематических взаимосвязей для предметной области ИТ, т. е. для выбранного экспертом научного направления (дисциплины) будет вычисляться частота цитируемости публикаций этого направления в описаниях изобретений по ИТ, имеющих коды рубрик G06, G11C и G10L.

Используемые за рубежом подходы к определению количественных индикаторов тематических взаимосвязей науки и технологий основаны, как правило, на следующих исходных положениях.

Во-первых, эти индикаторы трактуются как частотные оценки, пропорциональные числу публикаций по тем научным направлениям, которые цитируются в описаниях изобретений, относящихся к рассматриваемой технологии. Так как научные публикации из одной области знаний или одного научного направления могут цитироваться в опи-

саниях изобретений из разных рубрик МПК, то в этом случае результат количественного оценивания является векторной величиной, а для нескольких областей знаний или научных направлений — матрицей (табл. 2) [44, с. 420]. При получении этих оценок учитывается тот факт, что цитирование научных справочников и классических научных трудов в описаниях изобретений, скорее всего, будет указывать на хрестоматийные, а не на новые научные результаты. Кроме того, достоверность частотных оценок во многом будет зависеть от степени представительности используемых массивов патентных и научных информационных ресурсов.

Во-вторых, рассматриваемые частотные оценки отражают тематические взаимосвязи научных результатов с технологиями, но не отражают обратных связей (в частности, влияния результатов технологических разработок на инициирование фундаментальных и прикладных исследований и получение в итоге новых научных результатов).

В-третьих, если в отдельно взятом описании изобретения нет ссылок на научные публикации, то это не должно интерпретироваться как отсутствие связей этого отдельно взятого изобретения с научными результатами, так как не все случаи передачи знаний обязательно эксплицируются в виде цитируемых научных публикаций [45].

Авторы «Третьего европейского отчета по научно-технологическим индикаторам» отмечают особое положение такой области знаний, как математика [44, с. 421]. С одной стороны, в описаниях изобретений к патентам (далее по тексту — в патентах) редко встречаются ссылки на математические публикации (см. табл. 2), с другой стороны, очевидно, что математические методы и модели являются необходимыми во многих сферах технологий (производства).

Остановимся на этом примере процентных индикаторов, определенных с использованием научных и патентных информационных ресурсов, более подробно. Набор индикаторов представляет собой матрицу чисел, характеризующих взаимосвязи

Таблица 2 Индикаторы взаимосвязей сфер технологий (производства) и областей знаний

Сферы технологий (производства)	Науки о Земле и других планетах	Сельскохозяйственные науки	Химия	Медицинские науки	Технические науки	Науки о жизни	Физика	Математика	Междисциплинарные проблемы
Биотехнологии	0,1	4,4	2,0	21,3	0,2	52,8	0,1	0,0	19,1
Фармацевтические и косметические средства	0,0	2,5	5,7	42,4	0,2	34,0	0,1	0,0	14,9
Технологии тонкого органического синтеза	0,0	2,8	10,8	28,8	0,2	40,6	0,1	0,0	16,5
Сельское хозяйство и химическое производство пищевых продуктов	0,1	33,4	2,1	4,2	0,8	48,1	0,0	0,0	11,2
Контрольно-измерительные технологии	0,4	1,5	6,3	29,0	6,6	32,4	10,3	0,0	13,3
Нефтехимическая промышленность и химическая переработка сырья	0,2	8,1	10,6	32,0	1,1	33,3	1,1	0,0	13,6
Технологии производства пищевых продуктов	0,0	15,6	3,9	15,6	3,9	41,7	0,6	0,0	18,7
Полупроводниковая промышленность	0,5	0,5	13,1	1,9	23,8	0,3	58,7	0,0	1,1
Телекоммуникационные технологии	0,6	1,9	1,0	2,5	77,0	0,7	15,7	0,2	0,3
Ядерная техника и технологии	1,6	0,0	8,1	17,7	37,1	4,8	24,2	0,0	6,5
Информационные технологии	1,2	1,0	1,2	6,8	71,1	5,4	11,4	0,2	1,7
Космическая техника	0,0	0,0	20,0	0,0	50,0	0,0	30,0	0,0	0,0
Оптическая промышленность	0,2	0,0	12,0	0,7	22,5	1,7	61,2	0,0	1,4
Медицинские технологии	0,0	1,1	2,8	51,7	4,0	23,4	6,8	0,0	9,9
Технологии обработки поверхностей и лакокрасочные технологии	1,7	0,6	32,8	1,7	16,9	2,8	39,0	0,0	4,5
Технологии химии полимеров	0,2	4,3	42,6	13,3	3,1	26,0	1,0	0,0	9,3
Аудиовизуальные технологии	0,0	0,0	0,0	2,9	63,8	0,0	32,6	0,0	0,7
Металлургическая промышленность и производство материалов	3,4	0,0	29,9	3,4	34,0	2,7	19,7	0,0	6,8
Электротехническая промышленность	0,0	0,4	23,3	3,6	25,7	0,4	42,2	0,0	4,4
Химическое машиностроение	2,9	2,9	42,7	9,7	18,4	9,7	4,9	0,0	8,7

Примечание 1. Междисциплинарные проблемы позиционируются как самостоятельная область знаний.

Примечание 2. Данные таблицы получены в результате обработки европейских патентных заявок, поданных в период времени 1992–1996 гг.

Примечание 3. Сумма значений процентных индикаторов по строкам может быть меньше 100%, так как в первой строке таблицы перечислены не все области знаний.

сфер технологий (производства) и областей знаний, определенную на основе информационных ресурсов ЕПВ. Эта матрица получена в результате выборки и редуцирования данных из упомянутого европейского отчета [44].

В первой строке табл. 2 приведены области знаний. В первой колонке перечислены 20 сфер тех-

нологий (производства). Каждый столбец табл. 2, кроме первого, содержит вектор частотного распределения научных публикаций одной из 9 перечисленных областей знаний по 20 сферам технологий. Для вычисления каждого значения индикатора считалось число тех научных публикаций этой области знаний, которые цитировались в патентах, относя-

щихся к указанной в начале строки сфере технологий.

Например, строка «Биотехнологии» содержит число 52,8. Это число говорит о том, что в обработанном массиве патентов ЕПВ по биотехнологиям 52,8% всех научных цитат являются ссылками на научные публикации по наукам о жизни. Строка «Телекоммуникационные технологии» содержит числа 0,6 и 0,2, которые говорят о том, что в патентах этого вида технологий 0,6% всех цитат являются ссылками на научные публикации по наукам о Земле и 0,2% — ссылками на математические публикации.

Очевидно, что каждая из областей знаний может быть структурирована более детально и для нее может быть получена своя матрица индикаторов, определенных с использованием научных и патентных информационных ресурсов. Пример более детальной структуризации с индикаторами взаимосвязей теоретических и прикладных научных дисциплин с ИТ, вычисленными с использованием патентов США и ЕПВ, был рассмотрен в работе [46].

Отметим, что данные диаграмм, приведенных в этой работе, и табл. 2 иллюстрируют существенные отличия в цитировании научных публикаций для разных теоретических и прикладных дисциплин. Существуют отличия и в региональном разрезе, например доля научных публикаций по компьютерной науке, на которые есть ссылки в патентах США по ИТ, равна 2,35%, что на 2,2% меньше, чем доля научных публикаций по компьютерной науке, на которые есть ссылки в европейских патентах по ИТ [46].

Таким образом, ретроспективное сопоставление научных и патентных информационных ресурсов является основой для индикаторного оценивания взаимосвязей результатов науки и развития технологий. При этом вычисление значений индикаторов не является итоговым этапом методологии оценивания этих взаимосвязей. Заключительным этапом методологии является экспертная оценка вычисленных значений индикаторов. При этом мнения экспертов далеко не всегда будут соответствовать вычисленным значениям индикаторов.

Естественно, возникает вопрос о том, что если используемая методология индикаторного оценивания взаимосвязей имеет объективную основу в виде ретроспективных научных и патентных информационных ресурсов, то зачем нужен завершающий этап — экспертиза вычисленных значений индикаторов. Причина кроется в том, что значения полученных индикаторов являются объективными, но только *косвенными мерами* взаимосвязей науки и технологий. Поэтому для их верификации и используются субъективные, но прямые методы

экспертизы, включающие содержательный анализ сопоставляемых научных и патентных информационных ресурсов. Если косвенные объективные оценки, вычисленные на основе накопленных информационных ресурсов, совпадают с субъективными и прямыми согласованными экспертными оценками, то только в этом случае они могут служить надежной основой для выбора приоритетных направлений ориентированных научных исследований, являющихся ключевыми для создания перспективных технологий.

Таким образом, сопоставление научных и патентных информационных ресурсов является ключевым этапом методологии индикаторного оценивания процессов трансформации знаний в новые технологии.

5 Предлагаемая методология индикаторного оценивания

Как отмечалось ранее, современные методологии оценивания процессов трансформации знаний в новые технологии состоят из двух основных компонентов: (1) индикаторного оценивания взаимосвязей результатов исследований и разработок с технологической модернизацией и (2) индикаторного оценивания изобретательской активности в процессе исследований и разработок. В России до настоящего времени используется только второй компонент, но полностью отсутствует первый, так как в нашей стране нет информационных систем для индикаторного оценивания взаимосвязей.

Важно отметить, что отечественные патентные информационные ресурсы не удовлетворяют требованиям зарубежных вариантов методологии оценивания процессов трансформации знаний в новые технологии (по полноте и структуре ресурсов), что исключает возможность копирования зарубежных вариантов [3, 4]. Поэтому актуальной является задача разработки нового варианта методологии, который учитывал бы исторически сложившуюся структуру и наполнение отечественных патентных информационных ресурсов.

Приведем описание предлагаемого варианта методологии определения индикаторов тематических взаимосвязей науки и технологий в интересах выбора направлений ориентированных научных исследований.

Предлагаемый вариант методологии включает следующие 8 основных этапов.

1. Определение временного периода для вычисления ретроспективного тренда количественных индикаторов, выбор отечественных систем

классификации областей знаний¹, составление списков анализируемых технологических областей, формирование перечня индикаторов, которые необходимы в интересах выбора направлений ориентированных научных исследований.

2. Формирование массивов патентных информационных ресурсов по анализируемым технологическим областям (с использованием списков рубрик МПК из номенклатуры FhG-ISI) для определенного временного отрезка вычисления ретроспективного тренда количественных индикаторов.
3. Формирование базы данных научных публикаций, цитируемых в описаниях изобретений, в виде их библиографических описаний (далее по тексту — указатель цитирования), которая создается на основе сформированных массивов патентных информационных ресурсов по анализируемым технологическим областям. Для каждой публикации указывается идентификатор изобретения в виде номера патента и/или заявки на выдачу патента и все индексы МПК этого изобретения.
4. Автоматизированная структуризация библиографических описаний с одновременной нормализацией и пополнением нормативных списков отечественных и иностранных журналов (для статей), названий конференций (для научных докладов и сообщений) и названий издательств (для книг и трудов конференций с учетом того, что труды ряда конференций публикуются в периодических изданиях) с целью обеспечения автоматизированного рубрицирования.
5. Автоматизированное рубрицирование библиографических описаний цитируемых научных публикаций с использованием названий журналов, издательств, конференций и названий статей, а также с использованием выбранной системы классификации областей знаний.
6. Вычисление индикаторов тематических взаимосвязей как частотных оценок, полученных в процессе сопоставления рубрик МПК описаний изобретений, в которых встретились ссылки на научные публикации, и рубрик системы классификации областей знаний, к которым относятся цитируемые научные публикации, а также визуализация вычисленных значений индикаторов.
7. Получение экспертных оценок для вычисленных значений индикаторов взаимосвязей на

основе аналитических исследований и содержательного анализа сопоставляемых научных и патентных информационных ресурсов.

8. Проверка согласованности экспертных оценок и вычисленных значений индикаторов, фиксация тех пар рубрик МПК и системы классификации областей знаний, в которых мнения экспертов не согласованы между собой (и/или не согласуются с вычисленными значениями индикаторов), и для всех случаев несогласованности — дополнение базы данных цитируемых научных публикаций аннотациями и другими дополнительными полями (с последующим переходом на этап 5).

Сопоставим предлагаемый вариант методологии с европейским вариантом [47]. Для первого этапа отметим одно принципиальное отличие. В предлагаемом варианте имеется возможность выбора систем классификации областей знаний, которая отсутствует в европейском варианте. С точки зрения принятия решений по финансированию направлений ориентированных исследований эта возможность позволяет применять ту систему классификации областей знаний, которая используется при принятии решений (например, ГРНТИ, классификатор направлений фундаментальных научных исследований РАН, классификаторы РФФИ, РГНФ и т. д.).

В европейском варианте применяется только одна система классификации областей знаний (SCI-ISI Journal Classification System), по рубрикам которой распределены журналы Указателя научного цитирования SCI. Эта система классификации используется в процессе рубрицирования цитируемых научных публикаций, которые идентифицируются с помощью пяти полей: (1) фамилия первого автора; (2) источник публикации; (3) год публикации; (4) том; (5) номер первой страницы публикации. Используя название источника публикации, с помощью этой системы классификации определяется тематическая рубрика публикации [47, с. 407].

На третьем этапе отметим второе принципиальное отличие. В предлагаемом варианте методологии предусмотрено формирование указателя цитирования. При этом в указатель включаются и затем используются все научные публикации, цитируемые в описаниях отобранных изобретений (в европейском варианте рассматриваются только так называемые «front-page публикации», которые цитируются на первых страницах описаний отобранных изобретений).

¹Основным критерием для выбора системы классификации областей знаний является ее использование в процессе принятия решений по финансированию направлений ориентированных научных исследований.

На пятом этапе отметим еще одно важное отличие. В европейском варианте для определения тематической рубрики публикации используется только название источника публикации, а в предлагаемом варианте используются и названия источников, и названия цитируемых научных публикаций. Ключевые слова из названий публикаций используются для уточнения рубрики публикации в тех случаях, когда одному источнику соответствует несколько рубрик. С этой целью для каждой рубрики системы классификации областей знаний, которая используется при принятии решений, планируется формировать ее терминологический портрет, используемый в процессе уточнения рубрики публикации.

Еще одно принципиальное отличие касается восьмого этапа методологии, где появляется возможность существенно снизить область несогласованных экспертных оценок для индикаторов за счет использования аннотаций научных публикаций. Ключевые слова аннотаций планируется использовать в тех случаях, когда одни названия источников и цитируемых научных публикаций дают несогласованные экспертные оценки и/или эти оценки не согласованы с полученными значениями индикаторов.

6 Заключение

Предлагаемый вариант методологии позволит вычислять значения количественных индикаторов взаимосвязей науки и технологий как для целых областей знаний, так и для отдельных направлений исследований в целях идентификации технологически ориентированных научных направлений.

Впервые в нашей стране появляется возможность в процессе идентификации ориентированных научных направлений использовать *объективную основу*, а именно: накопленные научные и патентные информационные ресурсы.

Предлагаемый вариант методологии обладает рядом следующих принципиальных отличий от зарубежных аналогов:

- выбор именно тех систем классификации областей знаний, которые используются в процессе принятия решений;
- использование всех научных публикаций, цитируемых в описаниях отобранных изобретений, для определения значений индикаторов взаимосвязей науки и технологий (одновременно планируется вычислять и варианты индикаторов с использованием только «front-page публикаций», что даст возможность сопоставить

полученные отечественные результаты с зарубежными);

- формирование и применение терминологических портретов рубрик отобранных систем классификации областей знаний;
- использование ключевых слов из названий публикаций для уточнения рубрики публикации в тех случаях, когда одному источнику публикации приписано несколько рубрик систем классификации областей знаний;
- использование ключевых слов из аннотаций публикаций в тех случаях, когда названия источников и цитируемых научных публикаций дают несогласованные экспертные оценки и/или эти оценки не согласованы с полученными значениями индикаторов.

В соответствии с предлагаемым вариантом методологии разработаны: концепция и архитектура информационной системы, предназначенной для индикаторного оценивания взаимосвязей науки и технологий; архитектуры функциональных подсистем, обеспечивающих реализацию методологии; основные проектные решения для функциональных подсистем и видов обеспечения этой системы.

Разрабатываемая информационная система индикаторного оценивания взаимосвязей науки и технологий не имеет аналогов в российской научно-технической сфере. Ее создание необходимо для проведения мониторинга, многоаспектного оценивания программ научных исследований и прогнозирования научно-технологического развития страны.

Литература

1. Проект государственной программы РФ «Развитие науки и технологий» на 2013–2020 годы. <http://минобрнауки.рф/документы/2475>.
2. *Архипова М. Ю., Зацман И. М., Шульга С. Ю.* Индикаторы патентной активности в сфере информационно-коммуникационных технологий и методика их вычисления // Экономика, статистика и информатика. Вестник УМО, 2010. № 4. С. 93–104.
3. *Зацман И. М., Шубников С. К.* Принципы обработки информационных ресурсов для оценки инновационного потенциала направлений научных исследований // Электронные библиотеки: перспективные методы и технологии, электронные коллекции — RCDL'2007: Труды IX Всеросс. научной конф. — Переславль: Ун-т города Переславля, 2007. С. 35–44.
4. *Зацман И. М., Курчавова О. А., Галина И. В.* Информационные ресурсы и индикаторы для оценки инновационного потенциала направлений научных исследований // Системы и средства информатики. Доп. вып. — М.: Наука, 2008. С. 159–175.

5. Мезоэкономика развития / Под ред. чл.-корр. РАН Г. Б. Клейнера. — М.: Наука, 2011. 805 с.
6. Клейнер Г. Б., Голиченко О. Г., Зацман И. М. Основные принципы разработки системы мониторинга функционирования исследовательских организаций. — М.: ЦЭМИ РАН, 2007. 62 с.
7. Зацман И. М., Вережкин Г. Ф., Шубников С. К. Моделирование систем мониторинга. — М.: ИПИ РАН, 2008. 115 с.
8. Шубников С. К. Формы документов в системах информационного обеспечения оценки результативности научной деятельности // Системы и средства информатики. — М.: Наука, 2005. Вып. 15. С. 59–76.
9. Зацман И. М. Информационные ресурсы для систем мониторинга в сфере науки // Системы и средства информатики. — М.: Наука, 2005. Вып. 15. С. 288–318.
10. Зацман И. М. Терминологический анализ нормативно-правового обеспечения создания систем мониторинга и оценки результативности в сфере науки // Экономическая наука современной России, 2005. № 4. С. 114–129.
11. Зацман И. М., Вережкин Г. Ф. Информационный мониторинг сферы науки в задачах программно-целевого управления // Системы и средства информатики. — М.: Наука, 2006. Вып. 16. С. 164–189.
12. Шубников С. К., Лоцилова Е. Ю., Косарик В. В. Принципы систематизации и стандартизации описания структур информационных ресурсов в сфере науки // Системы и средства информатики. — М.: Наука, 2006. Вып. 16. С. 190–213.
13. Зацман И. М. Полидоменные модели в системах оценки инновационного потенциала и результативности научных исследований // Компьютерная лингвистика и интеллектуальные технологии: Труды Междунар. конф. Диалог-2006. — М.: РГГУ, 2006. С. 178–183.
14. Зацман И. М. Полидоменные модели электронных библиотек систем мониторинга сферы науки // Электронные библиотеки: перспективные методы и технологии, электронные коллекции — RCDL'2006: Труды VIII Всеросс. науч. конф. — Ярославль: ЯрГУ, 2006. С. 75–81.
15. Зацман И. М., Вережкин Г. Ф., Дрынова И. В., Курчавова О. А., Ларин Н. В., Норежян Т. П. Моделирование систем информационного мониторинга как проблема информатики // Системы и средства информатики. Спец. вып. Научно-методологические проблемы информатики. — М.: ИПИ РАН, 2006. С. 112–139.
16. Зацман И. М., Кожунова О. С. Семантический словарь системы информационного мониторинга в сфере науки: задачи и функции // Системы и средства информатики. — М.: Наука, 2007. Вып. 17. С. 124–141.
17. Zatsman I., Kozhunova O. Evaluating for institutional academic activities: Classification scheme for R&D indicators // 10th Conference (International) on Science and Technology Indicators (STI'2008): Book of Abstracts. — Vienna: ARC GmbH, 2008. P. 428–431.
18. Кожунова О. С. Семантический словарь системы информационного мониторинга в сфере науки и ресурс Eurowordnet: структура, задачи и функции // Системы и средства информатики. — М.: Наука, 2008. Вып. 18. С. 156–170.
19. Архипова М. Ю., Зацман И. М. Основные тенденции патентной активности в сфере информационных и телекоммуникационных технологий // Институциональные основы инновационных процессов: Мат-лы 4-х Друкеровских чтений. — М.: Доброе слово, 2008. С. 201–206.
20. Архипова М. Ю., Зацман И. М., Хавансков В. А. Индикаторы патентной активности РАН // Институциональные концепции менеджмента: Мат-лы 6-х Друкеровских чтений. — Екатеринбург: УрГУ, 2009. Т. 1. С. 141–150.
21. Zatsman I., Kozhunova O. Evaluation system for the Russian Academy of Sciences: Objectives—Resources—Results approach and R&D indicators // 2009 Atlanta Conference on Science and Innovation Policy Proceedings / Eds. S. E. Cozzens and P. Catalan. <http://smartech.gatech.edu/bitstream/1853/32300/1/104-674-1-PB.pdf>.
22. Зацман И. М. Категоризация результатов и индикаторов программ научных исследований в информационных системах мониторинга // Системы и средства информатики. Доп. вып. — М.: ИПИ РАН, 2009. С. 200–219.
23. Zatsman I., Durnovo A. Incompleteness problem of indicators system of research programme // 11th Conference (International) on Science and Technology Indicators (STI'2010): Book of Abstracts. — Leiden: CWTS, 2010. P. 309–311.
24. Методические рекомендации по подготовке Докладов о результатах и основных направлениях деятельности субъектов бюджетного планирования на 2006–2008 годы. http://www.minfin.ru/common/img/uploaded/library/2005/07/metod_270705.doc.
25. Planning for performance and evaluating results of public R&D programs // Meeting the OMB PART Challenge: Workshop Report. — Washington: The Washington Research Evaluation Network, 2004.
26. Special Report No.9/2007 concerning “Evaluating the EU Research and Technological Development (RTD) framework programmes — could the Commission’s approach be improved?” // Official J. Eur. Union C26, 30.01.2008. P. 1–38.
27. Zatsman I., Durnovo A. Program-oriented indicators: Production and application in science // Системы и средства информатики, 2012. Т. 22. № 1. С. 110–120.
28. Zatsman I., Durnovo A. Proactive dictionary of evaluation system as a tool for science and technology indicator development // 17th Conference (International) on Science and Technology Indicators Proceedings. — Montréal: Science-Metrix and OST, 2012. Vol. 2. P. 905–906.

29. *Зацман И. М., Косарик В. В., Курчавова О. А.* Задачи представления личностных и коллективных концептов в цифровой среде // Информатика и её применения, 2008. Т. 2. Вып. 3. С. 54–69.
30. *Зацман И. М.* Семиотическая модель взаимосвязей концептов, информационных объектов и компьютерных кодов // Информатика и её применения, 2009. Т. 3. Вып. 2. С. 65–81.
31. *Зацман И. М.* Нестационарная семиотическая модель компьютерного кодирования концептов, информационных объектов и денотатов // Информатика и её применения, 2009. Т. 3. Вып. 4. С. 87–101.
32. *Зацман И. М., Дурново А. А.* Моделирование процессов формирования экспертных знаний для мониторинга программно-целевой деятельности // Информатика и её применения, 2011. Т. 5. Вып. 4. С. 84–98.
33. *Zatsman I.* Tracing emerging meanings by computer: Semiotic framework // 13th European Conference on Knowledge Management Proceedings. — Reading: Academic Publishing International Limited, 2012. Vol. 2. P. 1298–1307.
34. *Zatsman I.* Denotatum-based models of knowledge creation for monitoring and evaluating R&D program implementation // 11th IEEE Conference (International) on Cognitive Informatics & Cognitive Computing Proceedings. — Los Alamitos, CA: IEEE Computer Society Press, 2012. P. 27–34.
35. *Schmoch U.* Tracing the knowledge transfer from science to technology as reflected in patent indicators // *Scientometrics*, 1993. Vol. 26. P. 193–211.
36. *Computational science: Ensuring America's competitiveness. Report to the President.* — Arlington, VA: National Coordination Office for Information Technology Research and Development, 2005. 104 p.
37. *Tijssen R. J. W., Buter R. K., Van Leeuwen Th. N.* Technological relevance of science: An assessment of citation linkages between patents and research papers // *Scientometrics*, 2000. Vol. 47. No. 2. P. 389–412.
38. *Van Looy B., Zimmermann E., Veugelers R., Verbeek A., Mello J., Debackere K.* Do science-technology interactions pay on when developing technology? An exploratory investigation of 10 science-intensive technology domains // *Scientometrics*, 2003. Vol. 57. No. 3. P. 355–367.
39. *Narin F., Noma E.* Is technology becoming science? // *Scientometrics*, 1985. Vol. 7. No. 3–6. P. 369–381.
40. *Narin F., Olivastro D.* Linkage between patents and papers: An interim EPO/US comparison // *Scientometrics*, 1998. Vol. 41. No. 1–2. P. 51–59.
41. *Mansfield E.* Academic research and innovation // *Research Policy*, 1991. Vol. 20. Is. 1. P. 1–12.
42. *Mansfield E.* Academic research underlying industrial innovations: Sources, characteristics and financing // *Review of Economic and Statistics*, 1995. Vol. 77. No. 1. P. 55–62.
43. *Mansfield E.* Academic research and industrial innovation: An update of empirical findings // *Research Policy*, 1998. Vol. 26. Is. 7–8. P. 773–776.
44. *Third European Report on Science & Technology Indicators.* — Luxembourg: Office for Official Publications of the European Communities, 2003. 451 p.
45. *Nonaka I., Takeuchi H.* The knowledge-creating company. — N.Y.: Oxford University Press, 1995 (пер. на русский язык: *Нонака И., Такеучи Х.* Компания — создатель знания. — М.: Олимп-бизнес, 2003).
46. *Зацман И. М., Кожунова О. С.* Предпосылки и факторы конвергенции информационной и компьютерной наук // Информатика и её применения, 2008. Т. 2. Вып. 1. С. 77–98.
47. *Verbeek A., Debackere K., Luwel M., Andries P., Zimmermann E., Deleus D.* Linking science to technology: Using bibliographic references in patents to build linkage schemes // *Scientometrics*, 2002. Vol. 54. No. 3. P. 399–420.