

МОДЕЛИРОВАНИЕ СИСТЕМ ИНФОРМАЦИОННОГО МОНИТОРИНГА КАК ПРОБЛЕМА ИНФОРМАТИКИ¹

*И.М. Зацман, Г.Ф. Веревкин, И.В. Дрынова,
О.А. Курчагова, Н.В. Ларин, Т.П. Норекия*

В мае 2004 года было принято постановление Правительства РФ о переходе на новые принципы организации бюджетного процесса, которое предусматривает введение среднесрочного финансового планирования в стране. Эти принципы предполагают создание, развитие и использование систем информационного мониторинга для оценки результативности бюджетных расходов во всех социально значимых сферах деятельности, включая сферу науки. Основной целью статьи является описание методологических аспектов моделирования систем информационного мониторинга, предназначенных для оценки инновационного потенциала направлений, программ, проектов научных исследований, включая количественную оценку влияния направленных научных исследований на процессы патентования. Обосновывается необходимость полидоменного моделирования на этапе эскизного проектирования новых и модернизации существующих систем информационного мониторинга, сопоставляются математическое и полидоменное моделирование этих систем.

1. Введение

В работе [1] рассмотрены основные функции и области применения информационных систем, обеспечивающих реализацию новых принципов организации бюджетного процесса. Приведённые в этой работе примеры наглядно демонстрируют необходимость разработки теории информационного мониторинга. В настоящее время эта теория находится только в начальной стадии формирования, на которой необходимо очертить комплекс

основных проблем мониторинга, включая постановку проблемы верифицируемости данных, предоставляемых системами информационного мониторинга пользователям (далее — системы мониторинга) [2].

Основной прикладной задачей теории информационного мониторинга является разработка принципов создания, развития и использования систем мониторинга для оценки результативности бюджетных расходов. Создание этих систем предусмотрено в Концепции реформирования бюджетного процесса в Российской Федерации в 2004–2006 годах (далее — Концепция). Концепция одобрена постановлением Правительства РФ от 22 мая 2004 г. № 249 «О мерах по повышению результативности бюджетных расходов».

Суть Концепции выражена в стенограмме выступления заместителя Председателя Правительства РФ А.Д. Жукова на VI Международной научной конференции «Модернизация экономики и выраживание институтов» следующим образом: «В прошлом году мы начали работу по бюджетированию, нацеленному на результат. Для этого, прежде всего, требуется чётко определить, какие Правительство ставит перед собой стратегические цели и задачи. После того, как эти стратегические цели и задачи определены, нужно определить, какое место каждое ведомство занимает в решении этих стратегических целей и задач, определить для этого ведомства показатели, по которым можно было бы судить: эффективно или неэффективно достигается решение этих задач. А уже после этого ведомство само должно расписать свои бюджетные деньги по тем программам, которые должны обеспечить достижение этих показателей в решении общих задач» [3].

Таким образом, основной идеей Концепции является распределение средств бюджета принимаемых обязательств между действующими и (или) вновь принимаемыми бюджетными программами с учётом оценок их результативности на основе программно-целевого планирования. Ядром новой Концепции должны стать широко применяемые за рубежом принципы организации бюджетного процесса, ориентированные на результаты, которые должны быть релевантны долгосрочному прогнозу, эксплицированным, то есть сформулированным в явном виде, стратегическим целям и задачам среднесрочного финансового планирования.

¹ Работа выполнена при частичной поддержке РГНФ, проект № 06-02-04043а.

В соответствии с Концепцией, по мере накопления опыта применения новых методов бюджетного планирования и развития систем мониторинга результативности бюджетных расходов, сфера применения новых методов программно-целевого планирования будет расширяться и структурироваться. Предусматривается поэтапное создание систем мониторинга в разных сферах деятельности, в том числе, в сфере науки.

Из всего спектра вопросов разработкой теоретических основ создания, развития и использования систем мониторинга, настоящая статья посвящена вопросам методологии моделирования этих систем, создаваемых и применяемых в сфере науки, со поставлению математического и полидоменного моделирования, а также обоснованию использования полидоменных моделей для количественной оценки влияния результатов научных исследований на процессы патентования.

2. Методологические основы полидоменного моделирования

Понятие полидоменных моделей, включающих лексико-семантический, информационный, алгоритмический, математический и другие компоненты, было введено в работе [2]. Подобные сочетания перечисленных компонентов предназначены для моделирования человеко-машинных информационных систем, ресурсы которых используются одновременно для решения информационно-справочных и аналитических задач в слабо формализуемых и институционально сложных сферах деятельности, например, в сфере науки.

В полидоменных моделях формализованное описание охватывает одновременно и предметную область мониторинга, включающую события и/или объекты институциональных систем, от ну уровня до макроуровня [4], и саму систему мониторинга. Для сферы науки объектами мониторинга на нуоуровне могут быть результаты деятельности конкретных учёных и специалистов, на микроуровне — научных лабораторий, на мезоуровне — научных институтов и других учреждений сферы науки, а на макроуровне объектом мониторинга могут быть результаты деятельности РАН или всей сферы науки в целом.

Так как целевыми объектами мониторинга могут быть одновременно множество уникальных объектов нуоуровня и массо-

вые события макроуровня, то способ формализованного описания, как предметной области мониторинга, так и самой системы мониторинга, должен, в общем случае, одновременно обладать свойствами разных категорий моделей — математических, логических, алгоритмических, информационных и лингвистических. Потребность в объединении свойств перечисленных категорий моделей и стала отправной точкой для описания новой категории моделей, названных в работе [2] *полидоменными*.

Использование полидоменных моделей позволяет объединить возможности математических, логических, алгоритмических, информационных и лингвистических моделей. Это объединение представляет собой формализованное описание, сочетающее, с одной стороны, структурированное описание лингвистических ресурсов, содержащих описание единичных объектов и событий, например, множество научных результатов в виде структурированных текстов научных статей на естественном языке, которые могут включать математические формулы, таблицы, другие небергаальные объекты [5, 6], и, с другой стороны, формальное описание характеристик массовых событий, например, статистические кривые возрастного распределения специалистов, получающих эти научные результаты. В этом заключается первое отличие полидоменных моделей от других категорий моделей, описывающих, как правило, либо массовые, либо единичные объекты и события предметной области [7].

Второе отличие полидоменных моделей заключается в том, что они позволяют в формализованной форме эксплицировать связи между характеристиками (атрибутами) единичных объектов и событий и массовыми объектами и событиями предметной области [2].

Третье отличие полидоменных моделей заключается в том, что в них, в общем случае, учитываются и сочетаются синтаксические, семантические и прагматические аспекты формализованного описания единичных и массовых объектов и событий [2, 7]. Под прагматическими аспектами в данном случае понимаются цели и способы использования моделей.

Остановимся подробнее на идее сочетания перечисленных аспектов формализованного описания в информатике объектов, событий и процессов предметной области. Идея сочетания синтаксических, семантических и прагматических аспектов, предложенная С. Горном, является ключевой составляющей его методологического подхода, сформулированного следующим обра-

зом: одно из главных интуитивных представлений специалистов в сфере информатики заключается в том, что любой процесс, который можно точно определить, может быть смоделирован в символично-знаковой форме, поскольку точная спецификация процесса уже является некоторой разновидностью символично-знакового моделирования этого процесса [8, с. 132].

Далее С. Горн пишет о том, что в приложениях информатики используются не только числа. В зависимости от предметной области могут также использоваться:

- аналитические выражения из математики и физики,
- одно-, двух- и трёхмерные структурные формулы химии,
- структурно-графические представления полипептидных цепей и двойной спирали ДНК,
- партитуры симфоний,
- обозначения балетных движений по Р. Лабану с помощью системы хореографических символов,
- спецификации схем микропроцессоров и т. п.

После перечисления столь разных символично-знаковых форм С. Горн отмечает, что специалисты из большинства сфер деятельности согласны с тем, что информатика обладает полезными для них возможностями, несмотря на возможные смысловые различия между их пониманием своих символично-знаковых форм и пониманием коллег из области информатики [8, с. 133].

Затем С. Горн говорит о формализованном описании объектов и событий предметной области следующим образом: «Информатика должна соотносить себя именно с прагматическими вопросами, от которых, как мы уже убедились, не должны зависать математика и физические науки в той части, которая касается, если не методов, то результатов. В этом отношении информатика имеет большее сходство с лингвистикой, психологией, биохевоистическими науками, философией и различными профессиями, имеющими к ним отношение». Далее С. Горн пишет, что изучение символично-знаковых форм и операций с ними, независимое от их смыслового содержания или прагматического контекста, называется *синтактикой*. В изучении вопросов синтактики и синтаксических описаний информатика является наиболее формальной дисциплиной, тесно связанной с математикой и её методами. Изучение отношений между символично-знаковыми формами, например, словами, и их значениями (смыслом), независимое от целей и способов использования форм, то

есть с исключением из рассмотрения прагматических вопросов, называется *семантикой* [8, с. 132].

Приведённые основные положения методологического подхода С. Горна и стали отправной точкой для разработки идеи полидоменного моделирования [2, 7, 9]. Основываясь на этом подходе, для интеграции синтаксических, семантических и прагматических аспектов в процессе моделирования в работе [7] был предложен следующий набор компонентов для построения полидоменных моделей:

- 1) спецификация символично-знаковой системы, используемой для описания предметной области мониторинга;
- 2) информационный компонент модели в форме набора схем информационных ресурсов системы мониторинга;
- 3) математический компонент в виде множества векторных функций и аналитических отношений, используемых для вычисления значений индикаторов;
- 4) алгоритмический компонент, необходимый для определения значений параметров векторных функций и аналитических отношений на основе информационных ресурсов системы мониторинга;
- 5) коммуникационный компонент модели, определяющий связи между параметрами векторных функций, теми информационными ресурсами и алгоритмами, которые необходимы для определения значений параметров;
- 6) лексико-семантический компонент, включая используемые классификаторы, тезаурусы, системы онтологий системы мониторинга;
- 7) аналитический компонент в форме набора критериев, построенных на основе индикаторов, нормативно-правовых документов, методики построения критериев и правил их применения пользователями систем мониторинга.

Следствием и развитием методологического подхода С. Горна является включение в состав полидоменной модели информационного, математического, алгоритмического и коммуникационного компонентов в качестве синтаксической составляющей модели, спецификации символично-знаковой системы и лексико-семантического компонента в качестве семантической составляющей и аналитического компонента в качестве прагматической составляющей модели [2, 7, 9].

3. Особенности моделирования систем мониторинга в сфере науки

Основное отличие процесса разработки систем мониторинга объектов и событий в сфере науки от процесса разработки систем мониторинга промышленных и технических объектов заключается в том, что разработчики первых систем часто не могут пренебречь единичными событиями и объектами нанотуринга промышленными объектами (или системами наблюдения за техническими объектами). Например, увольнение одного рабочего с конвейера промышленного предприятия практически не должно повлиять на результативность и эффективность этого предприятия. Увольнение же одного доктора наук из исследовательского института может заметно изменить результативность и эффективность этого института в том случае, если им подготовлены, но ещё не опубликованы монография и несколько статей, если его регулярно приглашают с докладами на научные конференции, если он является руководителем проектов, финансируемых отечественными и зарубежными научными фондами, и т. д.

Причина этого отличия заключается в том, что в сфере науки базовой единицей, генерирующей знания и фиксирующей их в объективированной форме, является учёный как личность. При этом, «знание не пересаживаемо из головы в голову в силу одного простого онтологического обстоятельства: *никто вместо другого не может ничего понимать* (выделено авторами настоящей статьи), понять должен сам ... И этот акт понимания „самим“ не выводит ни из какой цепи обусловливания этого понимания, он должен совершиться или не совершиться, т. е. знание не перекачиваемо в другую голову, как в некую пустоту перекачивалась бы жидкость» [10]. Поэтому, когнитивная и креативная деятельность учёного и его новые фундаментальные результаты, представленные в объективированной форме, являются, как правило, значимыми событиями и уникальными объектами в сфере науки.

При традиционном математическом подходе к построению модели, события и объекты, отношения между ними, а также связи с внешними объектами в этой модели, значимые для определения индикаторов, должны быть описаны в символической форме. Для сферы науки математические формулы и модели

применимы в тех случаях, когда значимые для определения индикаторов признаки или характеристики уникальных объектов предметной области являются *количественными* и носят, как правило, массовый характер.

Например, исходными данными для определения широко известных и часто используемых индикаторов цитируемости одного автора или группы авторов являются число ссылок на их журнальные статьи. В сфере науки число ссылок на конкретную статью является важной *количественной* характеристикой этой статьи. Число ссылок на неё, которые встретились в определённом перечне журналов, определяется в течение заданного периода времени. После идентификации статей одного автора или группы авторов, на которые есть ссылки, и статей, содержащих эти ссылки, индикаторы цитируемости определяются с помощью простого деления общего числа ссылок на статьи этого автора или группы авторов на общее число ссылок в пределах некоторого перечня журналов.

В работе [7] дано более подробное описание индикаторов цитируемости, вычисляемых Институтом научной информации США (ИНИ США) на основе перечня журналов, включённых в указатели Science Citation Index (SCI) и Social Sciences Citation Index (SSCI)². По данным ИНИ США, в этих журналах в 2003 году было опубликовано 698 726 статей, в которых было сделано 4 339 511 ссылок на работы, опубликованные в период 1999–2001 годах, в том числе, 32 176 ссылок на работы российских учёных, опубликованных в этот период времени [11]. Следовательно, значение индикатора цитируемости учёных России в 2003 году было равно 0,74% ($32\,176 / 4\,339\,511 \cdot 100$), если учесть, что по определению этого индикатора учитываются ссылки на статьи, опубликованные в трёхлетний период (1999–2001 гг.), который заканчивается за два года до момента определения индикатора (2003 г.). Для США значение этого индикатора в 2003 году было равно 42,39%, для Германии — 7,09%, для Китая — 1,51%, для Южной Кореи — 0,94% и для Индии — 0,73%.

Приведём ещё один пример, но не из сферы науки, а из сферы инновационного производства. Для количественной оценки уровня технологической эффективности предприятий одной отрасли

² В 2006 году этот перечень состоял из 5698 журналов, включая 3781 журнал из указателя SCI и 1917 журналов из указателя SSCI.

четырёх действий арифметики, как это имеет место для индикатора цитируемости, недостаточно. Например, эволюция распределения предпрятий отрасли по уровням их эффективности описывается системой дифференциально-разностных уравнений, называемой моделью Полтеровича–Хенкина. Эта модель позволила ввести понятие скорости распространения новых технологий. Распространение новых технологий происходит в результате двух процессов: инновации и имитации технологий предпрятий, которые имеют более высокий технологический уровень [12, 13].

В сфере науки традиционно используются и такие индикаторы, которые не требуют предварительной идентификации статей или построения математических моделей. Например, значения индикатора процентной доли расходов страны на проведение фундаментальных исследований, прикладных исследований и опытно-конструкторских разработок (см. рис. 1) определяются простым делением статистически учитываемой величины каждого из трёх перечисленных на диаграмме видов расходов на их сумму с последующим умножением на 100, чтобы получить долю расходов в процентах.

Однако все приведённые примеры из сфер науки и инновационного производства объединяет то, что исходными данными вычислений служат только числовые величины. В первом примере исходными данными для вычисления индикаторов цитируемости является число ссылок на статьи учёных страны. При этом важно отметить, что процессы идентификации статей одного автора или группы авторов и определение числа ссылок на каждую статью находятся за пределами используемой расчётной модели. Во втором примере, исходными данными являются численные значения параметров модели Полтеровича–Хенкина, в третьем примере — численные значения перечисленных видов расходов. По этой причине для получения результата было достаточно провести вычисления в рамках традиционного математического подхода.

Потребность в построении полиномиальных моделей возникает в тех случаях, когда исходные данные являются не только числовыми величинами. Например, если бы в первом примере в перечень исходных данных для определения индикаторов цитируемости был бы включён сам массив из 698 726 журнальных статей 2003 года, то модель определения 4 339 511 ссылок на статьи, опубликованные в период 1999–2001 годах, и, в том чис-

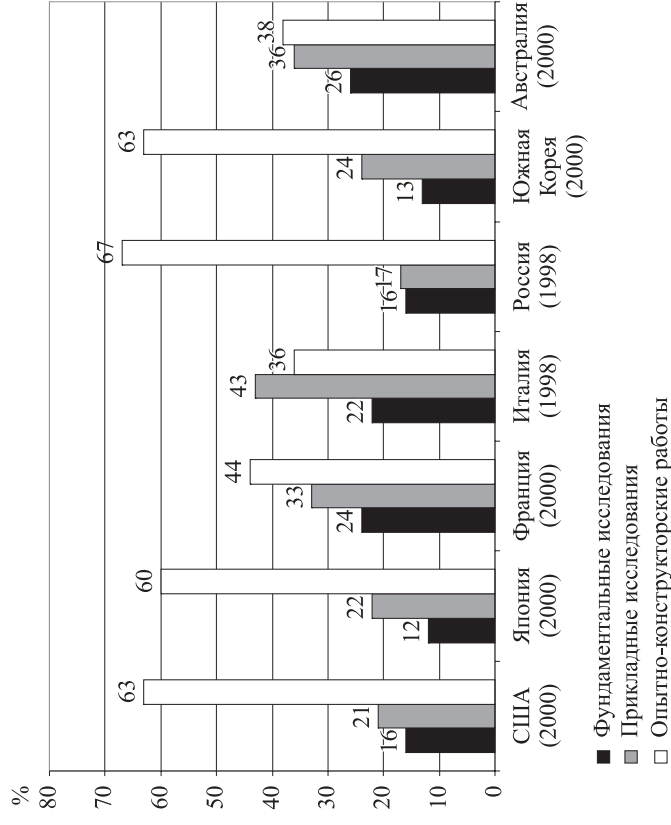


Рис. 1. Распределение затрат на НИОКР по видам работ в 1998 и 2000 гг. [14, V. 1, P. 4–62]. Замечание: об отнесённости 6% НИОКР, производимых в Японии, данные отсутствуют. Сумма трёх чисел для Франции и Италии не равна на диаграмме 100% и в том источнике, откуда была взята эта диаграмма. Источник: Science & Engineering Indicators — 2004

ле, 32176 ссылок на работы российских учёных, опубликованные в этот период времени, должна была предусматривать локализацию, структурирование и обработку текстов библиографических описаний, включая нормализацию названий журналов, идентификацию каждого номера журнала и каждой статьи с учётом того, что каждый номер журнала и каждая из 698 726 статей являются уникальными объектами в сфере науки.

4. Зависимость состояния макрообъектов от нанообъектов

Если исходными данными моделей являются не только числовые величины, то сложность моделирования будет во многом зависеть от тех уровней институциональной системы, которые являются источниками исходных данных. В примере с инди-

каторами цитирования источником исходных данных является нануровень сферы науки, на котором и появляются научные статьи как объективированные формы представления научных знаний.

Важно отметить, что статьи, как объекты нануровня, являются исходными данными для индикации состояния макрообъектов, а именно, всей сферы науки страны в целом. Другими словами, чтобы количественно оценить состояние макрообъекта, целевыми объектами мониторинга должны быть уникальные объекты нануровня. Если провести аналогию с моделированием полёта самолёта и считать самолёт макрообъектом, а самые маленькие его детали, например, его заклёпки, считать нанобъектами, то для индикации состояния самолёта в полёте было бы необходимо регистрировать уникальные состояния каждой его заклёпки. Поэтому, при моделировании технических макрообъектов объекты нануровня, как правило, не учитываются, что наглядно иллюстрирует пример с полётом самолёта, или учитываются опосредованно через интегральные характеристики.

Для индикации состояния сферы науки страны в целом необходимо учитывать множество уникальных объектов нануровня, для описания которых используются средства естественного языка или, в общем случае, символично-знаковых систем.

4.1. Нечисловые исходные данные модели. Продемонстрируем влияние нечисловых исходных данных на методологию построения полидоменных моделей на примере количественной оценки влияния направлений научных исследований на процессы патентования. Когда вычисляются и используют подобные оценки, то обычно делают оговорку, что речь идёт о получении только косвенных количественных оценок влияния результатов научных исследований на патентование.

Наиболее известным примером из этой области является косвенная оценка влияния научных исследований на патентование в США (см. рис. 2). На рисунке изображены две кривые. Верхняя кривая показывает изменение среднего числа ссылок на все публикации, встретившиеся в одном патенте США в период с 1987 по 2002 годы, а нижняя кривая показывает изменение только для научных статей, но не для всех публикаций [14, v 1, p. 5–51].

Приведённые на графике кривые косвенной оценки влияния научных исследований на патентование в США относятся ко

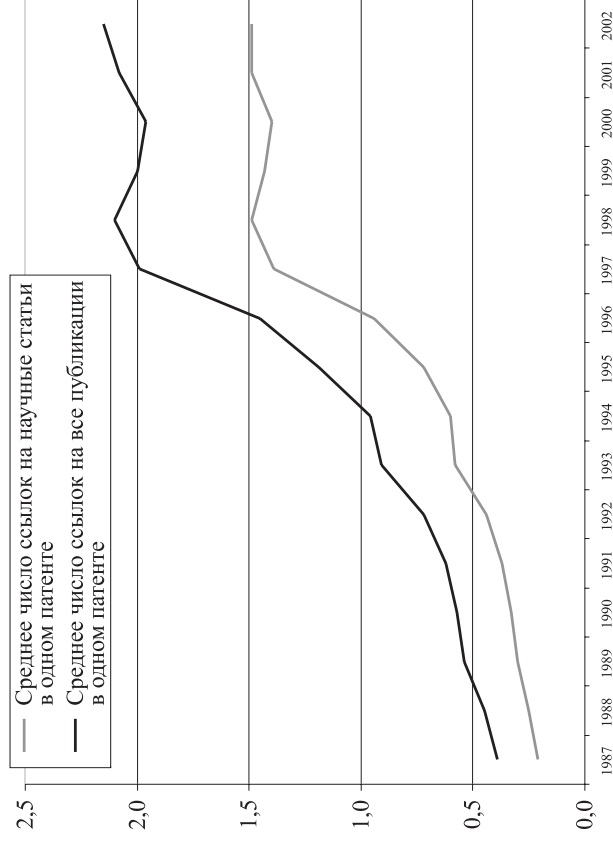


Рис. 2. Среднее число библиографических ссылок в одном патенте США: 1987–2002 гг.

всем направлениям научных исследований. Если же необходимо оценить влияние отдельных направлений научных исследований на процессы патентования, то задача получения подобных оценок, включая процесс моделирования, существенно усложняется.

В настоящее время модель, необходимая для получения оценок влияния отдельных направлений научных исследований на процессы патентования, находится в начальной стадии разработки, на которой проводится описание информационных потоков научных статей и патентов, а также структуризация полных текстов патентов с целью локализации в них ссылок на научные статьи. В процессе описания информационных потоков научных статей и патентов были выделены следующие макрообъекты сферы науки:

- 1) сфера фундаментальных исследований;
- 2) сфера прикладных исследований и опытно-конструкторских разработок без их разделения на две сферы, как это было сделано на рис. 1 для определения значений индикатора процентной доли расходов страны на проведение исследований и разработок (далее — сфера прикладных исследований);

3) сфера аккумуляции новых и наследуемых научных и научно-технических информационных ресурсов, включая научные статьи и патенты, а также программ и баз данных (далее — сфера аккумуляции ресурсов).

Отметим, что предлагаемое разделение научных исследований на фундаментальные и прикладные является весьма условными, так как в одном и том же исследовательском институте могут одновременно проводиться фундаментальные и прикладные исследования, а сам институт одновременно может являться депозитарием научных информационных ресурсов. Однако существует разделение на фундаментальные и прикладные исследования в структуре бюджетных расходов. Иначе говоря, в качестве критерия разделения используются коды источника финансирования научных исследований в соответствии с бюджетной классификацией, а не предмет научных исследований.

В примере выделяется сфера аккумуляции ресурсов, что встречается довольно редко, но для количественной оценки влияния результатов научных исследований на патентование выделение этой сферы позволяет учесть отложенное влияние результатов исследований. Важно отметить, что статистика использования фондов библиотек свидетельствует о том, что наследуемые информационные ресурсы могут оказывать существенное влияние на сферу науки страны в течение десятков лет после их создания [15].

Для учёта связей макрообъектов сферы науки с внешними объектами рассмотрим ещё две сферы:

- 1) инновационного производства;
- 2) некоммерческого использования результатов научных исследований — образование, медицина, экология, силовые структуры и т. д. (далее — сфера некоммерческого использования).

Если ограничиться только задачами количественной оценки влияния направленных научных исследований на патентование, то сферу некоммерческого использования результатов научных исследований можно было бы и не рассматривать. Однако её включение в модель может понадобиться при решении более широкого класса задач по количественной оценке влияния результатов научных исследований на некоммерческие социально значимые сферы деятельности общества.

Отметим, что в рассмотренном выше перечне образования не выделяется в качестве самостоятельной сферы, и эта сфера

учтена в рамках единой сферы некоммерческого использования результатов.

4.2. Информационные потоки и связи между сферами.

Для пяти выделенных сфер максимальное число теоретически возможных информационных потоков между ними равно 20, если учитывать, что каждая из них может быть связана с четырьмя другими. Если же учесть дополнительно и те внутренние информационные потоки, которые замкнуты внутри каждой из пяти сфер, то общее число возможных информационных связей возрастает до 25 (рис. 3).

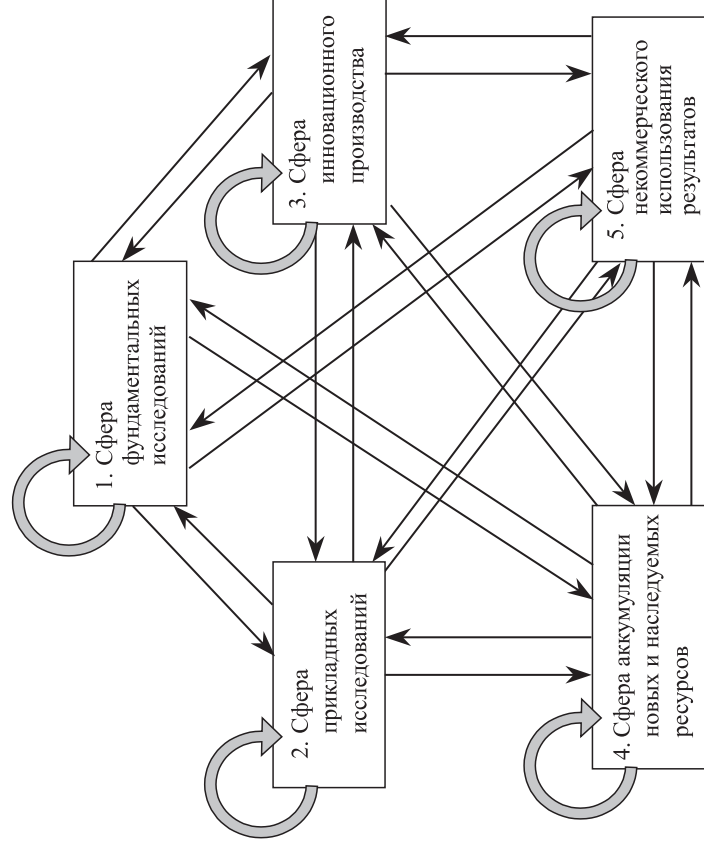


Рис. 3. 25 теоретически возможных информационных потоков

На рисунке условно, в виде прямоугольников, изображён случай пяти сфер и 25 информационных потоков между ними, включая 20 потоков, внешних по отношению к этим пяти сферам, обозначенных прямыми стрелками, и пять внутренних потоков, обозначенных в виде широких дугообразных стрелок.

Широкие стрелки обозначают те потоки информационных ресурсов, которые нашли применение в той же сфере, где они были созданы. Например, применительно к сфере фундаментальных исследований, это те научные результаты, которые используются как исходные данные для получения нового фундаментального научного знания. При этом не исключаются случаи, когда этот же результат может быть использован и в других сферах, что на рис. 1 обозначено 4 прямыми стрелками, выходящими из прямоугольника «1. Сфера фундаментальных исследований».

При решении частных задач моделирования одни информационные потоки могут практически отсутствовать, а другие существенно отличаться между собой по интенсивности. Однако, учитывая предполагаемое многоцелевое использование разрабатываемой модели, предлагается при её разработке учитывать все теоретически возможные информационные потоки между сферами, включая и те, которые при решении частных задач могут отсутствовать. Поэтому геометрия рис. 3 выбрана так, чтобы представить теоретически все возможные информационные потоки между сферами, без графического выделения тех связей, которые являются наиболее актуальными и/или доминирующими для количественной оценки влияния направлений научных исследований на процессы патентования. При решении частных задач этот рисунок можно преобразовать, ограничившись меньшим числом сфер, и расставить графические акценты на отдельных сферах и информационных потоках между ними для наглядного визуального акцентирования тех или иных информационных потоков.

4.3. Схема потоков без обратных связей. Преобразуем рис. 3 для иллюстрации того случая, который широко использовался в середине прошлого века, и получил название «линейной экономической модели инноваций». Согласно этой модели, разработанная фундаментальная идея воплощается затем в прикладных исследованиях. Результаты прикладных исследований служат основой инноваций, в результате реализации которых возникают передовые технологии. В рамках этой модели предполагалось, что обратные связи отсутствуют и при соответствующем финансировании обеспечении количества прикладных исследований и внедряемых передовых технологий прикладных исследований зависимости от объёма проводимых фундаментальных исследований [16, с. 6].

На рис. 4 изображена схема потоков без обратных связей для случая трёх сфер (фундаментальных исследований, прикладных исследований, инновационного производства). Как видно, между сферами имеется только две прямые связи, обозначенные оди-

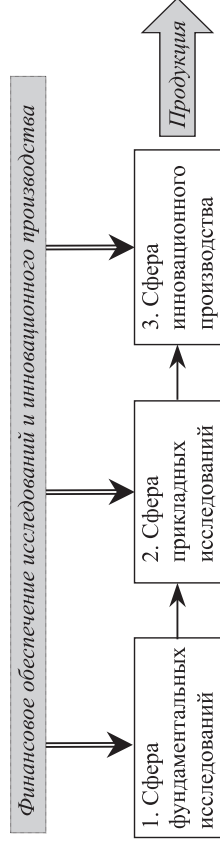


Рис. 4. Схема потоков без обратных связей

нарными стрелками. По сравнению с рис. 3, дополнительно обозначены финансовые потоки (двойными стрелками) и результаты инновационного производства (в виде одной фигурной стрелки со словом «Продукция»), которые в явной форме на рис. 3 не показаны.

4.4. Схема потоков с обратными связями. В последние 30–40 лет модель без обратных связей не является доминирующей, и всё большее внимание уделяется исследованию обратных связей. В методологии построения моделей предполагается, что идеи, которые лежат в основе инноваций (они, как правило, и патентуются), могут генерироваться во всех сферах, а не только в сфере фундаментальных исследований [16, стр. 6, 7].

Далее из пяти перечисленных сфер будут рассматриваться только четыре — фундаментальных исследований, прикладных исследований, инновационного производства и аккумуляции ресурсов (см. рис. 5). Хотя основной целью разработываемой модели является, в первую очередь, решение задачи количественной оценки влияния направленных научных исследований на патентование, рис. 5 акцентирует внимание на том, что в разрабатываемой модели можно будет исследовать одновременно *прямые* и *обратные* связи между четырьмя сферами. Приведены все теоретически возможные информационные потоки между сферами, включая 12 внешних по отношению к ним, которые обозначены одинарными стрелками, и четыре внутренних, которые обозначены в виде широких дугообразных стрелок.

На рис. 5 также изображены финансовые потоки и вещественные потоки, которые в явной форме на рис. 3 не показаны. Фи-

нансовые потоки обозначены двойными стрелками, а результат инновационного производства обозначен фигурной стрелкой со словом «Продукция». Показаны также система правового и нормативного регулирования четырёх сфер и функционирования систем мониторинга. Влияние системы правового и нормативного регулирования на четыре выделенные сферы условно обозначено одинарными пунктирными стрелками, а потоки исходных данных для систем мониторинга — двойными штрих-пунктирными стрелками.

Результаты мониторинга обозначены пунктирной фигурной стрелкой со словами «Услуги систем». Одинарной штрих-пунктирной стрелкой обозначены те результаты мониторинга, которые используются в качестве исходных данных при подготовке правовых и нормативных документов. Точечной стрелкой обозначено правовое и нормативное регулирование процессов мониторинга.

На рис. 5 показана схема из четырёх сфер и максимально возможного числа (16) прямых и обратных видов информационных потоков между сферами. Для решения задачи количественной оценки прямого влияния направлений научных исследований на патентование наибольший интерес представляют следующие пять видов информационных потоков (на рис. 5 они выделены полужирными стрелками):

- потоки внутри сферы фундаментальных исследований (дугообразная стрелка);
- потоки внутри сферы прикладных исследований (дугообразная стрелка)
- прямые потоки между сферой фундаментальных исследований и сферой прикладных исследований (прямая стрелка);
- прямые потоки между сферой фундаментальных исследований и сферой инновационного производства (ломаная прямая угловая стрелка);
- прямые потоки между сферой прикладных исследований и сферой инновационного производства (прямая стрелка).

Регистрация объектов нанотехнологического производства (прямая стрелка). Накопленные потоки являются весьма трудоёмкой задачей. Однако в случае достаточно развитой сферы аккумуляции новых и наследуемых ресурсов, исходные данные для решения задачи количественной оценки прямого влияния направлений научных исследований на патентование могут быть получены из этой сферы (на рис. 5 этот вариант обозначен полужирной двойной штрих-пунктирной стрелкой от сферы аккумуляции ресурсов

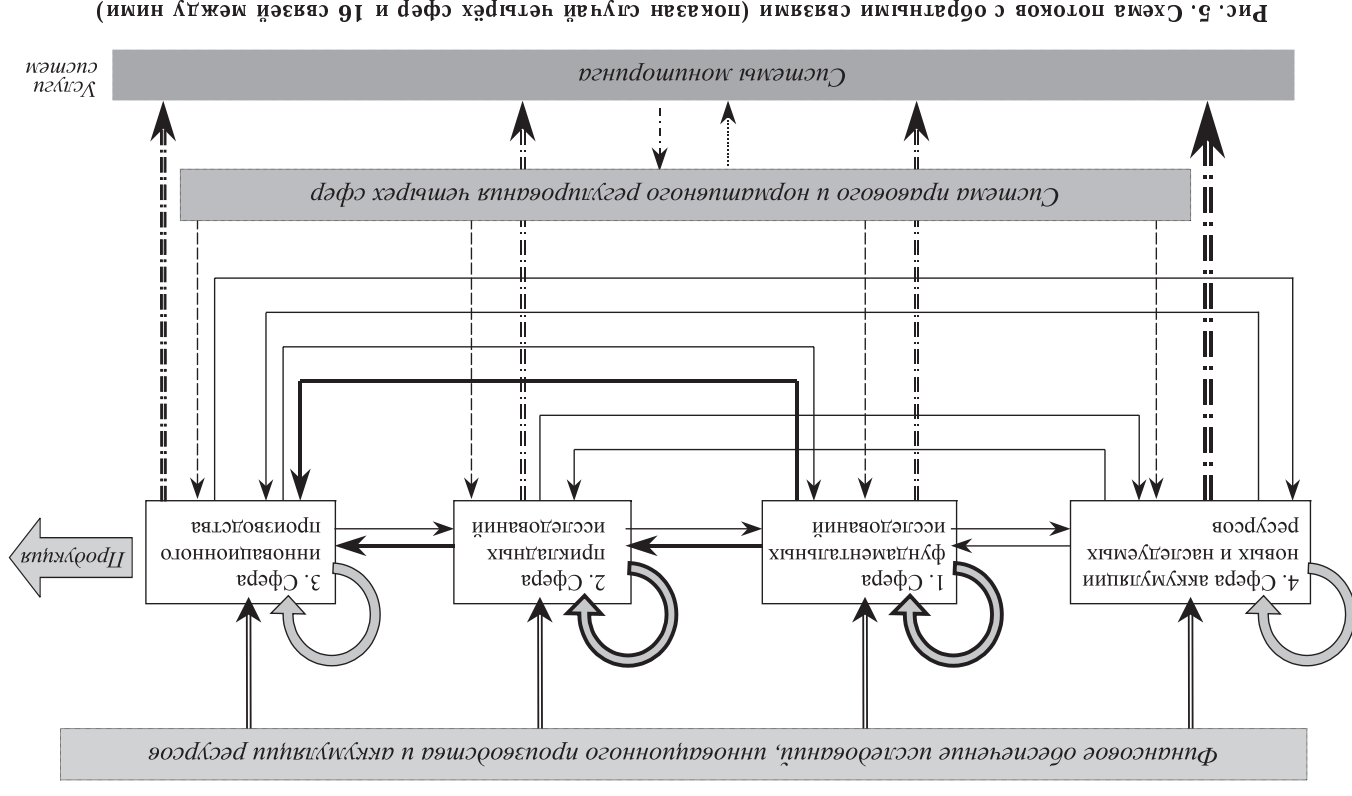


Рис. 5. Схема потоков с обратными связями (показан случай четырёх сфер и 16 связей между ними)

к системам мониторинга). Следовательно, последующие стадии разработки модели во многом будут зависеть от степени полноты информационных ресурсов в сфере их аккумуляции, с точки зрения решения задачи количественной оценки влияния направлений научных исследований на патентование.

5. Заключение

В заключение остановимся подробнее на вопросе полноты информационного потенциала сферы аккумуляции ресурсов для решения задачи количественной оценки влияния направлений научных исследований на патентование. На рис. 5 информационные потоки между сферами и исходные потоки данных для систем мониторинга обозначены по-разному (первые — одиными стрелками, а вторые — двойными штрих-пунктирными). Различие в обозначениях призвано в графической форме отметить тот факт, что реальные информационные потоки между сферами существуют, но они не полностью могут быть представлены в исходных потоках данных для систем мониторинга.

Поэтому основная проблема, возникающая при моделировании и создании систем мониторинга, заключается в определении степени представления (отражения) информационных потоков между сферами в исходных потоках данных для систем мониторинга. Решение вопросов представления информационных потоков зависит от трёх факторов.

Во-первых, заметная часть элементов информационных потоков между сферами и внутри них не фиксируется на материальном носителе (это может быть устное консультирование, обмен персоналом, совещания и т.п. [16, с. 7]), т.е. имеются такие случаи передачи знаний, когда непосредственных «следов» этой передачи, то есть объективированных форм представления знаний в виде публикаций, не остаётся. Отметим, что в одной из классификаций способов передачи знаний, передача без материальных «следов» выделена как один из четырёх основных способов передачи, который в этой классификации получил название «socialization process», что на русский язык можно перевести как «процесс обобществления, или социализации» [17].

Во-вторых, кроме способа передачи знаний без материальных «следов», существует ещё три способа передачи знаний, в которых непосредственные «следы» передачи знаний возникают. Однако и в этих случаях получение оценок для информационных

потоков между сферами является непостоянной задачей, так как не существует централизованной системы, регистрирующей эти потоки.

О степени децентрализации обработки потоков можно судить по результатам обследования крупных информационных служб, приведённым в книге «Инфосфера» [18]. Было подсчитано общее число журналов, обрабатываемых этими службами (48310). Наибольшее число журналов (13912, т.е. 28,8% от общего числа) обрабатывалось службой Chemical Abstracts [18, с. 262–264]. Приведённые результаты иллюстрируют объективную сложность разработки систем мониторинга, обусловленную децентрализацией центров (служб) обработки и аккумуляции информационных ресурсов.

В-третьих, исторически сложившаяся схема информационных потоков отличается от схем информационных потоков между сферами, которым мы пользовались на рис. 3–5 тем, что чёткие границы между сферами определить достаточно трудно, так как одни и те же генераторы информационных потоков могут функционировать одновременно в разных сферах. Поэтому порождаемые ими информационные документы могут одновременно иметь отношение к разным сферам.

Отсюда следует, что *разработка технологии идентификации информационных nanoобъектов по определённому набору признаков и «сортировки» nanoобъектов по сферам-источникам и сферам-получателям, является одним из ключевых вопросов при решении задачи количественной оценки влияния направлений научных исследований на патентование.*

Рассмотренные методологические аспекты моделирования систем мониторинга в настоящее время используются для разработки методов оценки инновационного потенциала конкретных направлений, программ, проектов научных исследований, включая алгоритмы количественной оценки влияния направлений научных исследований на патентование.

Моделирование систем информационного мониторинга является одним из весьма актуальных и перспективных направлений развития информатики. Что же касается прикладного аспекта обрабатываемых методов полидоменного моделирования, то их предполагается опробовать на этапе эскизного проектирования системы мониторинга, анализа и оценки деятельности научных организаций РАН и их подразделений, разрабатываемой специалистами ИПИ РАН и ЦЭМИ РАН.

Список литературы

1. Зацман И.М., Веревкин Г.Ф. Информационный мониторинг сферы науки в задачах программно-целевого управления // Системы и средства информатики. Вып. 16. — М.: Наука, 2006. — С. 185–210.
2. Зацман И.М. Полидоменные модели в системах оценки инновационного потенциала и результативности научных исследований // Труды международной конференции Диалог-2006 «Компьютерная лингвистика и интеллектуальные технологии». — М.: Изд-во РГГУ, 2006. — С. 178–183.
3. Стенограмма выступления Заместителя Председателя Правительства РФ А.Д. Жукова на VI Международной научной конференции «Модернизация экономики и наращивание институтов». — (http://www.hse.ru/temp/2005/files/04_06_2005_jukov.doc).
4. Клейнер Г.Б. Эволюция институциональных систем. — М.: Наука, 2004. — 240 с.
5. Курчава О.А. Когнитивно-лингвистическое моделирование таблиц в научных документах // Системы и средства информатики. Вып. 13. — М.: Наука, 2003. — С. 313–328.
6. Зацман И.М., Курчава О.А. Лингво-семиотический подход к анализу диаграмм // Системы и средства информатики. Вып. 14. — М.: Наука, 2004. — С. 170–185.
7. Зацман И.М., Веревкин Г.Ф., Шубников С.К. Моделирование систем мониторинга. — М.: ИПИ РАН, 2006 (в печати).
8. Gott S. Informatics (computer and information science): its ideology, methodology, and sociology. In: The studies of information: Interdisciplinary messages / Ed. by F. Machlup and U. Mansfield. — N. Y.: Wiley, 1983. — P. 121–140.
9. Зацман И.М. Полидоменные модели электронных библиотек систем мониторинга сферы науки // Труды 8-й Всероссийской научной конференции «Электронные библиотеки: перспективные методы и технологии, электронные коллекции» — RCDL'2006. — Суздаль, 2006 (в печати).
10. Мамардашвили М.К. Классический и неклассический идеалы рациональности. — Тбилиси: Мецниереба, 1984. — 82 с.
11. National Science Board, Science and Engineering Indicators — 2006. Two volumes. — Arlington, VA: National Science Foundation, 2006.
12. Полтерович В.М., Хенкин Г.М. Эволюционная модель взаимодействия процессов создания и заимствования технологий // Экономика и математические методы. 1988. Т. XXIV. Вып. 6. С. 1071–1083.
13. Ташлицкая Я.М., Шананин А.А. Моделирование процесса распространения новых технологий. — М.: ВЦ РАН, 2000. — 50 с.

14. National Science Board, Science and Engineering Indicators — 2004. Two volumes. — Arlington, VA: National Science Foundation, 2004 (V. 1. NSB 04-1; V. 2. NSB 04-1A).
15. Евстигнеева Г.А., Земсков А.И. Анализ спроса депозитарных фондов ГПНТБ России и роль электронных технологий // Труды 4-й Всероссийской конференции «Электронные библиотеки: перспективные методы и технологии, Электронные коллекции» (Дубна, 15–17 октября 2002 г.): В 2 т. Т. 1. — Дубна: ОИЯИ, 2002. — С. 40–50.
16. Инновационная система России: модель и перспективы её развития. Вып. 1. Анализ мирового опыта формирования и функционирования инновационных систем в контексте развития российской национальной модели. — М.: Изд-во РУДН, 2002. — 84 с.
17. Margwick A.D. Knowledge management technology // IBM Systems Journal. 2001. V. 40, № 4. P. 814–830.
18. Арский Ю.М., Гиляревский Р.С., Туров И.С., Чёрный А.И. Инфосфера. — М.: Изд-во ВИНТИИ, 1996. — 490 с.