ВОЗМОЖНОСТИ ОПЕРАЦИОННОЙ СИСТЕМЫ FREEBSD НА РАБОЧЕМ МЕСТЕ И В СЕТИ

Локтев А.А.
АНОУ «МФЮА»

В работе рассматривается система FreeBSD, занимающая в мире второе место по распространенности среди операционных систем с открытым кодом и предоставляющая программное обеспечение для использования в любых целях, без дополнительных ограничений. Используя этот программный продукт можно избежать дополнительных сложностей и правовых коллизий, присущих коммерческому использованию GPL-продуктов. Каждый проект внутри FreeBSD поддерживает публично доступное дерево исходных текстов программ. При помощи Concurrent Versions System (CVS) можно получить доступ к коду проекта, его документации и вспомогательным файлам. Это позволяет пользователям в любой момент получить копию дерева любого из проектов или системы в целом. В проекте FreeBSD пользователям может предоставляться три различных варианта системы: Current — версия для разработчиков, Release — версия для конечных пользователей, Stable — версия FreeBSD, являющаяся логическим продолжением версии Release. По мере того, как в Release обнаруживаются ошибки и в дерево CVS вносятся изменения, дистрибутив переходит на стадию Stable.

Как и любая сетевая операционная система, FreeBSD предлагает набор программ, которые позволяют превратить компьютер в узел, предоставляющий Интернет - услуги: электронная почта, веб- или FTP-сервер, сервер доменных имен, средства для маршрутизации и сетевой трансляции адресов, прокси-сервер, сетевой экран и т. п. Кроме того, свободно распространяемый пакет Samba позволяет организовать на базе FreeBSD файловый сервер и сервер печати в сети Microsoft либо создать Primary Domain Controller для сети Windows. На FreeBSD могут поддерживаться коммерческие и свободно распространяемые версии SQL-серверов, также может быть организована работа в составе кластера. Кроме количества доступного программного обеспечения, большое значение для выступающей в роли сервера операционной системы имеет устойчивость в работе, по этому критерию система FreeBSD занимает одно из лидирующих мест.

FreeBSD также подходит на роль рабочей станции, которая используется для работы в Интернете и для обработки текстовой информации. В этой системе, как и в большинстве Unix-подобных операционных систем, графический интерфейс представлен средой X Window System, которая не обладает большим количеством функциональных возможностей, поэтому для упрощения работы пользователей используются оконные менеджеры, которые по удобству сопоставимы со средствами Windows. Хотя все же у части программ, работающих под этой операционной системой во FreeBSD нет аналогов.

В работе делается достаточно подробное сравнение FreeBSD, Windows и Linux, приводятся классы задач для успешного решения, которых рекомендуется та или иная операционная система. Значительная часть работы посвящена описанию команд FreeBSD для администрирования самой системы, управления консолями, файловой системой, процессами и работой с пользователями. Изложенный материал позволит пользователю понять принципы и отличия использования системы FreeBSD, освоить более быструю и надежную работу с приложениями без графического интерфейса.

СЕТЕВЫЕ ИНТЕЛЛЕКТУАЛЬНЫЕ ГЕОИНФОРМАЦИОННЫЕ СИСТЕМЫ

А.Н. Шогин

ВИНИТИ РАН

INTELLEGENT NETWORKING GEOINFORMATION SYSTEMS

A. Shogin

Введение

В последние годы в области геоинформатики начаты исследования, направленные на создание распределенной геоинформационной среды. В частности, задача разработки распределенной среды GeoWeb поставлена в 2006 г. ведущей в области геоинформатики компанией ESRI (ArcReview № 1 (36) 2006). Работы в данной области в России ведутся с 2004 г. в рамках программы Президиума РАН «Электронная Земля». Наиболее существенным результатом этого проекта является разработка базовой версии сетевой геоинформационной среды [1,2].

Одним из важнейших элементов геоинформационной среды «Электронная Земля» являются распределенные сетевые аналитические ГИС ГеоПроцессор 2, КОМПАС V и ГеоТайм II. Функциональность этих систем ориентирована на решение двух типов задач: (1) Просмотр многодисциплинарной географической информации (ГИ) и оценивание связей между ее компонентами; (2) Нахождение многомерных зависимостей в ГИ, прогнозирование, обнаружение и распознавание целевых заранее неизвестных стационарных и динамических свойств изучаемой среды.

Распределенная технология создается с учетом требований, основанных на нашем опыте в разработке и применении сетевых аналитических ГИС для поддержки исследований в науках о Земле, экономике, социологии и в других дисциплинах.

Первое требование состоит в необходимости выполнения комплексного геоинформационого анализа разнотипных многодисциплинарных данных и использовании экспертных знаний для поиска устойчивых решений. Сложность получения устойчивого решения во многих случаях обусловлена тремя факторами: (А) взаимодействием рассматриваемых процессов, (Б) невозможностью прямых измерений их ключевых характеристик, (В) недостаточным объемом наблюдений и воздействием на результаты измерений шумов, уровень которых соизмерим с уровнем полезного сигнала. Второе требование связано с необходимостью использовать при анализе одновременно ГИ из различных хранилищ данных. Часто эти хранилища находятся в местах сбора и актуализации данных: на пространственно распределенных серверах и на ПК пользователя. Следующее требование обусловлено тем, что в научных исследованиях часто необходимо иметь возможность подключения специализированных плагинов к базовой конфигурации системы, которые, как и данные, могут находиться либо на пространственно распределенных серверах, либо на ПК пользователя. Четвертое требование относится к обеспечению возможности сохранения результатов пользователя, которые включают в себя полученные при работе геоинформационные слои и ГИС-проект, состоящий из метаданных исходных и полученных данных и из метаданных плагинов. Пятое требование основано на существенной роли наглядности при исследовании ГИ. Наглядность обеспечивается высокой интерактивностью управления процессом анализа и высококачественным

картографическим отображением данных на каждой операции анализа. И, наконец, последнее требование относится к обеспечению совместимости форматов данных с наиболее распространенными ГИС стандартами.

В современных информационных технологиях перечисленные требования хорошо реализуются либо с помощью сетевых приложений, которые заранее загружаются с сервера на ПК пользователя, либо с помощью апплетов, которые временно загружаются для работы с ними на ПК. Сетевые ГИС ГеоПроцессор 2.0, КОМПАС V и ГеоТайм II написаны на объектно-ориентированном языке Java. Это обеспечивает независимость программы от платформы пользователя, и ее выполнение либо как апплета в Java-машине любого современного браузера, либо как обычной локальной программы.

Рассмотрим более подробно функциональность ГИС ГеоПроцессор 2.0 и два примера ее применения для анализа геолого-геофизических данных.

Сетевая ГИС ГеоПроцессор

Сетевая ГИС ГеоПроцессор предназначена для интерактивного представления и анализа векторно-растровой географической информации. Области применения – комплексный анализ пространственных данных, изучение геологической среды и поддержка принятия решений в таких задачах как оценка природной опасности, прогноз природных ресурсов, оценка экологического состояния среды [3].

Технология ГИС ГеоПроцессор в значительной степени ориентирована на пользователя, не являющегося специалистом в области ИТ. Анализ поддерживается интерактивными операциям, высококачественной графикой и использует интуитивно понятные методы обработки данных. Одной и важнейших функций системы является возможность интеграции географической информации распределенной на сетевых серверах и на ПК пользователя (последнее обеспечивает конфиденциальность данных пользователя). Система также поддерживает возможность сохранения результатов пользователя (ГИС слоев и ГИС-проекта).

ГИС ГеоПроцессор 2.0 написана на языке Java 1.5. Это обеспечивает независимость программы от платформы пользователя, а также дает возможность ее выполнения в любом современном браузере.

ГИС ГеоПроцессор 2.0 поддерживает следующие функции интерактивной обработки и комплексного анализа векторных (точки, линии, полигоны) и сеточных пространственных данных:

1. Операции ввода/вывода:

- Загрузка данных, распределенных на сетевых серверах и на ПК пользователя по XML файлу ГИС-проекта.
- Динамическая загрузка векторных и сеточных (растровых) геоданных в форматах SHP (ESRI), FLT (ASCII), PTS (ASCII), PTN (ASCII) и растровых изображений в наиболее распространенных форматах JPG, PNG, GIF и т.п..
- Сохранение ГИС-проекта и полученных данных.
- Выбор проекции карты, сохранение и вывод на печать.

2. Визуальное исследование растровых и векторных данных:

- Композиция карты из растровых и векторных слоев.

- Изменение размеров и масштаба карты с интерполяцией и без интерполяции сеточных слоев.
- Динамическое изменение закраски, прозрачности, диапазона видимых значений, типа линий и размеров пиктограмм.
- Картографическое измерение сеточных и векторных слоев.
- Построение разрезов сеточных слоев карты по произвольному профилю и измерение значений по разрезу.
- Моделирование освещенности.
- Формирование выборок прецедентов в виде совокупностей единичных точек и/или полигонов с помощью указания объектов на карте и с помощью автоматического выбора прецедентов по сеточному или точечному слою.
- Комплексный анализ по сходству: функция сходства формируется с помощью функции расстояния до прецедентов или как принадлежность к построенным по прецедентам полуинтервалам.
- Оценивание статистик одного или двух сеточных слоев (минимум, максимум, среднее, среднеквадратичное отклонение, гистограмма, корреляция и ошибка аппроксимации).
- Редактирование координатной сетки.

3. Преобразование данных:

- Преобразование векторный слой -> сеточный слой (вычисление сеточных полей расстояний до векторных объектов, близости, плотности и т.д.)
- Преобразование сеточный слои -> сеточный слой с помощью операций растровой фильтрации (вычисление градиента, сглаживание в произвольном скользящем окне, вычисление среднеквадратичного отклонения, выделение аномалий и т.д.) или с помощью вычисления конструируемых пользователем произвольных функций от нескольких сеточных слоев.
- Преобразование сеточный слой и векторный слой -> атрибуты векторного слоя с помощью вычисления статистик сеточного слоя в буферных зонах векторных объектов или с помощью конструируемых пользователем произвольных функций от атрибутов нескольких векторных слоев.

4. Пространственный правдоподобный вывод

- Оценивание функции сходства к выборке прецедентов
- Оценивание функции принадлежности к двум классам
- Оценивание непараметрической регрессии.
- Оценивание функции предпочтения и нахождение логического выражения, объясняющего полученное решающее правило.

Оценка ущерба от землетрясений

Рассмотрим пример оценки возможного ущерба от максимального землетрясения для городов Северного Кавказа с населением более 100000 человек. Использованы данные о пиковых ускорениях при землетрясениях [4] (ресурс получен через Центральный портал геоинформационной среды «Электронная Земля» http://eearth.viniti.ru).

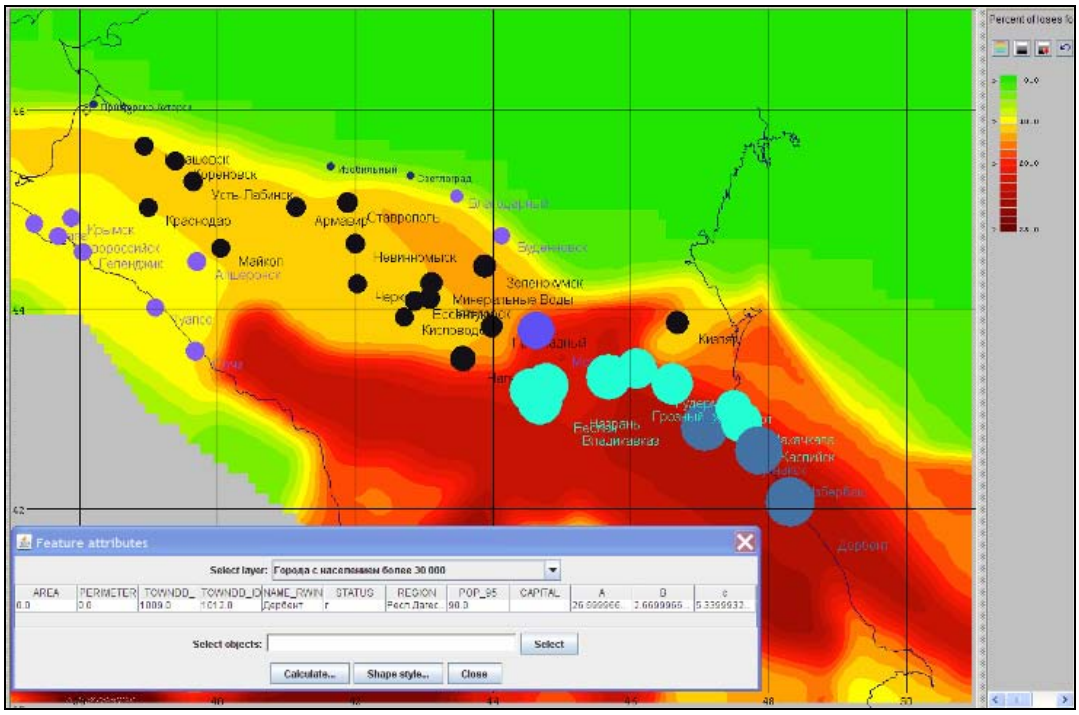

Рис 1. Разрушения застройки 7КП от максимальных землетрясений в %.

С помощью преобразований «*Сеточные слои* ⇒ *Сеточный слой*» по сеточному слою пиковых ускорений *A* вычислено поле максимального балла землетрясений $I=(\lg A-0.014)/0.3$ [5] . По полю *I* вычислен сеточный слой *V* доли разрушения зданий типа 7КП (тип зданий выбран только для иллюстрации метода): $V=0$ при $I<7$, $V=3.5\%$ при $I=7$, $V=11.9\%$ при $I=8$, $V=37\%$ при $I=9$. Затем применено преобразование «*Сеточные слои и Векторные слои* ⇒ *Атрибуты векторного слоя*». С его помощью вычислены доли разрушения застройки в 5 км окрестностях городов. Полагая застройку сооружениями типа 7КП однородной для выбранного размера зоны, можно принять результат в качестве оценки ущерба от максимального землетрясения. На Рис. 1 показаны сеточный слой разрушений построек 7КП в процентах, величинами кружков обозначены значения разрушений для городов. Внизу дано значение доли разрушения в г. Дербент, равное 27%.

Оценка сейсмической опасности

Рассмотрим пример выделения зон возможных очагов землетрясений (ВОЗ) с магнитудами $M>6.5$ для Кавказа. Использован ресурс http://www.iitp.ru/projects/geo/, GeoProcessor 2⇨ArmEast, разработанный по данным [6].

Следуя [3, 6], предполагалось, что очаговые зоны сильнейших землетрясений приурочены к пересечению областей неоднородности земной коры с зонами надвиговых и сдвиговых разломов, активных в Кайнозое.

Вначале с помощью аналитических преобразований и визуального исследования был выполнен разведочный анализ исходных и преобразованных данных. Для иллюстрации метода выбрано наиболее простое решение, использующее только поле модуля градиента скорости вертикальных движений за постсарматское время (признак X_1), и активные в кайнозое надвиговые разломы. С помощью преобразования «*Векторный слой* $\Rightarrow$ *Сеточный слой*» вычислен признак X_2 - поле расстояний до надвиговых разломов. Далее использован метод индуктивного логического вывода [3]. Полученное правило имеет вид:

ЕСЛИ градиент скорости вертикальных тектонических движений за постсарматское время (X_1) больше10 усл. ед.

И расстояние до надвиговых разломов (X_2) менее 20 км

ИЛИ X_1>6.4 усл. ед. **И** X_2<9 км,

ТО возможны очаги с M >6.0.

Полученные по данному правилу зоны ВОЗ и эпицентры с магнитудами M >6.0 показаны на Рис. 2. Прогноз зон в море, на юге и юго-востоке не производился из-за отсутствия геолого-геофизических данных.

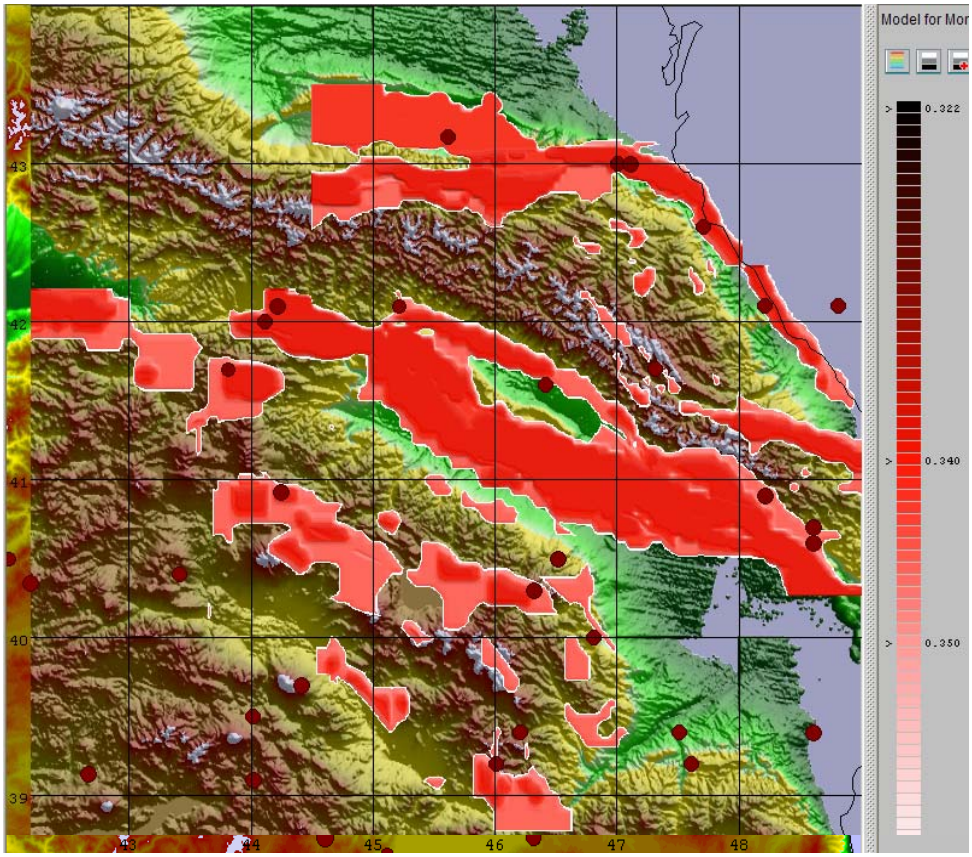

Рис. 2. Зоны возможных очагов землетрясений (ВОЗ) с M >6.0.

Заключение

Мы конспективно рассмотрели некоторые элементы технологии сетевых аналитических ГИС, составляющих основной аналитический ресурс геоинформационной среды, разработанной по Программе ПРАН «Электронная Земля». При анализе информации эти системы могут загружать для обработки данные и плагины, распределенные как в телекоммуникационных сетях, так и на ПК пользователя. Рассмотренные примеры пространственного и пространственно-временного анализа свидетельствуют об эффективности применения указанных систем для решения достаточно сложных исследовательских задач в науках о Земле.

В настоящее время увеличивается количество новых источников цифровой ГИ, которая имеет и пространственную и временную компоненты. Это связано в первую очередь с развитием средств мониторинга природных и социально-экономических процессов. Одной из важнейших инициатив в этой области является рассчитанный на предстоящие десятилетия Проект по созданию Глобальной системы систем наблюдений за Землей (GEOSS), которая начинает разрабатываться в целях понимания и решения глобальных проблем окружающей среды и экономики. Эти тенденции требуют своевременного создания новых геоинформационных технологий хранения, передачи и обработки огромных массивов данных, представляющих исследуемые процессы в пространстве и во времени. В базовой версии геоинформационной распределенной среды «Электронная Земля» сделаны первые шаги в данном направлении.

Работа выполнена при поддержке программы Президиума РАН «Разработка фундаментальных основ создания научной распределенной информационно-вычислительной среды на основе технологий GRID», направление: «Электронная Земля: научные информационные ресурсы и информационно-коммуникационные технологии».

Литература

1. Арский Ю.М., Гитис В.Г., Шогин А.Н. Электронная земля – сетевая среда поиска, интеграции и анализа геоданных // М., ПИК ВИНИТИ, Смирновский сборник, 2007 (в печати).
2. Arskyi, Yu., Gitis, V., Shogin, A., Weinstock, A. Network Geoinformation Environment for Analysis of Spatial and Spatio-Temporal Data // Abstracts of the IUGG XXIV General Assembly, Perugia, Italy. 2007.
3. Гитис В.Г., Ермаков Б.В. Основы пространственно-временного прогнозирования в геоинформатике // М., ФИЗМАТЛИТ, 2004. 256 С.
4. Giardini, D., Grunthal, G., Shedlock, K., and Zhang, P. Global Seismic Hazard Map assembled by, GSHAPE. 1999.
5. Trifunac, M.D. and Brady, A.G. On the correlation of seismic intensity with peaks of recorded strong ground motion // Bull. Seismol. Soc. Amer., 65. 1975. 139-162.
6. Gitis, V.G., Ermakov, B.V., Ivanovsnaia, L.V., Osher, B.V., Trofimov, D.P., Schenk, V., Shchukin, Ju.K., Jurkov, E.F. The Information Technology of the GEO system for Prediction M_{max} of Earthquake // Journal of Earthquake Prediction Recearch. 1993. v. 2, n. 2. 221-228

ТЕОРЕТИКО-МЕТОДОЛОГИЧЕСКИЕ АСПЕКТЫ ИНФОРМАЦИОННОГО РАЗВИТИЯ ХОЗЯЙСТВЕННЫХ СИСТЕМ

А.Э. Калинина
Волгоградский государственный университет

THEORETICAL AND METODOLOGICAL ASPECTS OF AN INFORMATION DEVELOPMENT OF ECONOMICAL SYSTEMS

A.E. Kalinina

В условиях социально-рыночной трансформации хозяйственной системы обмен информацией предполагает формирование такого информационного пространства, в котором складываются новые формы соединения факторов, качественно отличные от ранее существовавших в отечественной экономике. Необходимость уточнения теоретической модели хозяйственной системы с учетом развития информационного фактора связана с одной стороны, с развитием коммуникационных технологий, обусловливающих изменения институциональных и организационных взаимодействий в обществе. С другой стороны, фундаментальные положения экономической теории недостаточно учитывают отражение пространственно-временных изменений информации как трансакционного фактора хозяйственной деятельности. Современная экономическая теория нуждается в исследованиях изменений, происходящих в хозяйственных системах всех уровней под воздействием повышения интенсивности и расширения масштабов информационных взаимодействий.

Эффективность управления конкретной хозяйственной системой прямо зависит от механизма поиска, переработки, хранения и использования информации. Это предполагает определение и рационализацию внутренних и внешних связей и отношений информационных пространств данных систем, а при возрастании их сложности обеспечивает устойчивость координации их деятельности и сопряжения с системами более высокого и более низкого уровня. При неразвитом информационном пространстве в хозяйственной системе существенно возрастает количество слабо формализуемых задач принятия управленческих решений. В первую очередь это обусловлено неадекватностью имеющихся данных для диагностики состояния и динамики экономических субъектов и объектов, отражения и выявления тенденций и закономерностей их развития, выработки рациональной стратегии и тактики их функционирования в различных хозяйственных системах.

Современные преобразования в хозяйственной системе России, развитие предпринимательской деятельности и рыночных структур, необходимость оперативного и компетентного принятия решений сопровождаются ростом спроса на качественную, достоверную, оперативную информацию, удовлетворение которого возможно только на основе широкомасштабной информатизации как общества в целом, так и отдельных субъектов различных уровней, и, в первую очередь, региональных.

Информатизация хозяйственных процессов является закономерным этапом процесса эволюции хозяйственной системы общества. Современная концепция эндогенных факторов производства определяет информацию как один из ключевых факторов, обусловливающих экономическое развитие национального мезоуровня[1].

---

[1] См.: Иншаков О.В. Теория факторов производства в контексте экономики развития / О.В. Иншаков // Научный доклад на Президиуме МАОН (Москва, 29 ноября 2002 г.). – Волгоград: Изд-во ВолГУ, 2002.; Иншаков О.В. Знакомая и неведомая мезоэкономика / О.В. Иншаков, Н.Н. Лебедева // Экономика развития региона: проблемы, поиски, перспективы: Ежегодник. Вып. 2. – Волгоград: Изд-во ВолГУ, 2002.

Необходимость оценки масштабов и уровня развития информатизации становится особенно актуальной, когда информация трактуется как важнейший и редкий ресурс, фактор, элемент экономического потенциала региональной хозяйственной системы, который должен эффективно использоваться для достижения ее целей[2].

Возрастающая роль информатизации в развитии экономики регионов как подсистем национального хозяйства мезоуровня, необходимость комплексного исследования ее тенденций, закономерностей, условий и перспектив эффективной реализации обусловливают актуальность и необходимость разработки, апробации и совершенствования методов оценки уровня развития региональных информационных процессов.

Системный подход к изучению роли информации сформировался и нашел отражение в трудах Э.Шредингера, П. Тейяра де Шардена, Н.И.Кобозева, Н.Н. Моисеева, Н. Винера, Г. Громова, Д. Львова, Дж. Неймона, К.Шеннона, У. Эшби и др.

В современной отечественной экономической науке в последние годы XX века и в начале XXI века информационному фактору уделяется особенно большое внимание. В этом, несомненно, большая заслуга таких ученых как С. Авдашева и А. Яковлева, Г.Т. Артамонова, Н.А. Кузнецова, Н.Л. Мусхелишвили, Ю.А. Шрейдера, Е. Майминас, А. Мовсесяна, Р.М. Нижегородцева, И. Пещанской, А.А. Сидорова, В.Ф. Байнев, А.А. Штрик и др.

Значительный вклад в исследование формирующихся моделей информационного пространства и эффективных взаимодействий в нем внесли работы таких зарубежных авторов, как М. Лэхманн, Д. Норт, Г. Селла, К. Шапиро, Б. Эндмондс, а также отечественных авторов, в числе которых Т. Закупень, А. Киселев, Т. Кричел, И. Ладенко, Т. Нестеренко, С. Паринов, С. Старовит, В. Телерман, И. Швецов, Ю. Шрейдер. Понятийный аппарат изучаемой проблемы активно исследуется И. Андреевым, И. Дзялошинским, Н. Кривошеиным, В. Лопатиным, С. Нехаевым, С. Яскевичем и др.

Проблемы информатизации и трансформаций различных видов деятельности в условиях информационной экономики отражены в работах Р. Аткинсона, Дж. Брэдфорда, Дж. Гэлбрейта, Р. Друкера, К. Лаудона, М. Кастельса, К. Келли, Д. Москелла, В. Михайловского, Н. Моисеева, А. Ноува, Д. Подольны, К. Пэйджа, В. Старбука, Э. Тофлера, К. Фримана, а также Т. Богомоловой, Г. Клейнера, Д. Львова, М. Фрумкина, М. Хайдеггера, Ю. Хохлова, , Р. Цвылева, Ф. Широкова, А. Шадрина, К. Ясперса и др.

Влияние информационных технологий и механизмов на устойчивое развитие экономических систем стран и регионов исследуются М. Аттинджером, Д. Беллом, М. Бандеманном, М. Вебером, Б. Генкиным, А. Гранбергом, К. Голдспинк, В. Дрожжиновым, О. Иншаковым, Н. Лычкиной, М. Мизинцевой, Н. Моисеевым, Т.

---

[2] См.: Лаудон К., Старбук В. Организационная информация и знания / К. Лаудон, В. Старбук // Информационные технологии в бизнесе / Под редакцией М. Желены. – СПб.: Питер. – 2002. – С. 284–300.; Вальтух, К.К. Информационная теория стоимости /К.К. Вальтух. Новосибирск: Наука, 1996. 412 с; Яковец, Ю.В. Рента, антирента, квазирента в глобально-цивилизационном измерении /Ю.В. Яковец. – М.: ИКЦ «Академкнига», 2003. С. 141.; Урсул, А.Д. Природа информации. М.: Изд-во политической литературы, 1968. С.17.; Тамбовцев, В.Л. Пятый рынок: экономические проблемы производства информации. М.: Изд-во МГУ, 1993. С. 5-6.; Эрроу, К. Информация и экономическое поведение // Вопросы экономики. 1995. № 5. С. 100.; Стоуньер, Т. Информационное богатство: профиль постиндустриальной экономики // Новая технократическая волна на Западе / Т.Стоуньер. – М.: Прогресс, 1986. С. 393.; Goldfinger, Ch. Financial Markets as Information Markets: Preminary Exploration/ Paper prepared for the ENSSIB Conference Ëconomie de l'information». Lyon, May 20.

Шауэром, А. Штриком, А. Урсул, Э. Фогелем, Й. Хергетом, Б. Эдмондсом, К. Эрроу и др.

Вместе с тем возникает необходимость теоретического обоснования организационных и институциональных взаимодействий в рамках информационного пространства, разработки механизмов управления ими и использования в экономической теории. Не вполне определены экономические формы объектов информационных взаимодействий, противоречия и тенденции развития отношений между их субъектами, исследование содержания внутренних и внешних эффектов информатизации экономики. Недостаточно исследованы способы представления аналитической информации в различных базах знаний предметных областей экономического профиля.

Анализ степени научной разработки данной предметной области свидетельствует о недостаточной изученности информационного аспекта жизнедеятельности хозяйственных систем регионального уровня. В настоящее время не сложилась теория организации специализированных информационных систем отдельных сфер, сегментов, секторов и рынков, не изучен потенциал и механизмы их интеграции в единое информационное пространство экономики регионов. Возможности использования информационных методов анализа для оценки информационного развития хозяйственных систем, влияние информационных потребностей на развитие экономики региона, формирование условий, обеспечивающих эффективность информатизации, не получили достаточного освещения в отечественной экономической науке.

Требуется создание и внедрение как специализированных, так и комплексных трансакционных технологий и механизмов использования баз данных в регулировании региональной экономики. Нужны новые концептуальные представления и инструментарий для построения и анализа моделей, информационно отражающих функционирование экономики региона.

Эффективное управление процессом региональной информатизации зависит от механизмов сбора, переработки, хранения и использования информации. В условиях становления информационного общества это связано с использованием информационных технологий и автоматизированных информационных систем, которые позволяют аккумулировать и эффективно обрабатывать большие объемы аналитической информации, снижают трудоемкость процессов управления и использования информационных ресурсов, повышают надежность и оперативность передачи и получения информации. В этом случае обеспечивается возможность проведения диагностики состояния и динамики процесса информатизации, снижение трансакционных издержек, получение эффективных рычагов воздействия на хозяйственную систему при решении отраслевых и территориальных управленческих задач.

Применение современных компьютерных технологий оказывает влияние на организацию информационных ресурсов и предъявляет новые требования к технологиям их обработки, что обусловливает необходимость совершенствования методов их анализа и оценки информационных процессов.

Информация, являясь эндогенным фактором человеческой деятельности, формирует как общественное в целом, так и экономическое пространство. Внутри последнего представляется возможным определять информационное пространство по критериям уровня и масштаба, объекта и субъекта, состояния и динамики, фазы и зрелости и др. Более совершенное информационное пространство, снижая неопределенность хозяйственных взаимодействий субъектов, становится условием роста эффективности использования ими ресурсов и факторов производства. Развитие Internet позволяет вовлекать в обмен все больше информационных

ресурсов, использование которых обеспечивает количественное и качественное изменение структуры затрат и предъявляет новые требования не только к факторам и технологиям, но и к институциональной и организационной составляющей экономического прогресса. Необходимость общего сокращения трансформационных и трансакционных издержек, связанных с использованием информации, обусловливает формирование единого информационного пространства хозяйственной системы[3].

Институциональная специфика информационного пространства выражается в нормах и правилах, определяемых региональными особенностями рынков, их субъектов и объектов. Повышение эффективности взаимодействия субъектов разных сфер, сегментов и секторов хозяйственных систем регионов предполагают обмен адекватной информацией для рыночной координации их деятельности.

Системное стратегическое построение информационного пространства региона адекватное современным технологиям управления региональными хозяйственными системами предполагает выявление и оценку потенциала влияния информационных процессов на хозяйственную систему в целом, анализ и оценку их эффективности.

Становление рыночной структуры российского общества связано с переходом на новый, более сложный уровень организации и подчиняется общим законам развития хозяйственных систем. Одним из видимых проявлений этих законов является усложнение структуры информационного пространства как отражения экономической пространства. Информация, являясь эндогенным фактором человеческой деятельности, формирует как общественное в целом, так и экономическое пространство. Внутри последнего представляется возможным определять информационное пространство по критериям уровня и масштаба, объекта и субъекта, состояния и динамики, фазы и зрелости и др.

Ввиду сложности структурного строения во всякой большой системе процесс формирования любой из ее подсистем идет по разным структурным направлениям неравномерно и в значительной степени стихийно. В такой ситуации первостепенное значение имеет уровень информационной сложности системы.

Обмен информацией и другие виды информационной активности, являются основой реализации экономических процессов. Более совершенная среда информационной активности, снижая неопределенность хозяйственной деятельности субъектов, позволяет им достигать более высокого уровня возможной эффективности использования ресурсов. Развитие Internet позволяет вовлекать в обмен все больше информационных ресурсов, использование которых обеспечивает количественное и качественное изменение структуры затрат и предъявляет новые требования к технологической составляющей экономического прогресса. Необходимость сокращения трансформационных и трансакционных издержек, связанных с использованием информации, обусловливает формирование информационного пространства.

В этой связи становится необходимым решение ряда актуальных научных задач: формирование и дальнейшее развитие системы важнейших понятий, разработки теоретических основ управления информационными ресурсами, создание типовых моделей информационных систем и баз знаний для различных предметных

---

[3] См. подробнее: Калинина А.Э. Развитие информационного пространства региональной хозяйственной системы / А.Э. Калинина. – Волгоград: Изд. ВолГУ, 2005. – 367 с.; Калинина А.Э. Прикладные аспекты развития информационного пространства региона. (APPLIED ASPECTS OF DEVELOPMENT OF REGION INFORMATION SPACE) / А. Э. Калинина // «Фундаментальные исследования» Научно-теоретический журнал. Москва: «Академия естествознания», 2006. - № 6. С.86-89.

областей, адаптации концепции формирования и развития единого информационного пространства России к ее региональным особенностям. В соответствии с данным пониманием проблемы, основные цели формирования и развития единого информационного пространства региона можно сформулировать как: создание и поддержание необходимого для устойчивого развития хозяйственных систем уровня информационного потенциала на основе предметной ориентации информационных ресурсов, запасов и резервов, что позволит обеспечить согласованность решений, принимаемых органами государственной власти различных уровней.

Разработанная методологии исследования информационного пространства региональных хозяйственных систем позволяет учесть вышеназванные положения на этапах построения концептуальной модели предметной области, проектирования на ее основе информационной системы и соответствующей базы знаний и обработки полученной информации (процесс управления информационными ресурсами) посредством многопараметрической оценки эффективности информатизации и рейтингового позиционирования хозяйственных систем.

Использование концепции эндогенных факторов производства для решения ключевых задач управления на мезоуровне, предполагает разработку адекватного механизма оценки пространств, формируемых каждым из факторов в процессе их функционирования[4]. Эта оценка возможна и необходима для стратегического взаимодействия трансформационных и трансакционных факторов различных аспектированных хозяйственных систем. В частности, структуризация информационного фактора приводит к необходимости формирования информационного пространства, что обусловливает, в свою очередь, снижение трансакционных издержек.

Создание научных принципов исследования и проектирования информационных механизмов базируется на методологическом осмыслении взаимосвязи социально-экономических законов развития общества и процессов развития хозяйственных систем, познании закономерностей этих процессов.

Теоретический анализ позволяет установить, что в процессе своего развития информационные механизмы хозяйствования подвержены тем же трансформациям, которым подвергалась хозяйственная система в процессах эволюции, революции и инволюции рыночных отношений, изучение которых позволяет выявить содержание и противоречия формирования информационных механизмов управления хозяйством в конкретных формах и на различных уровнях его современной региональной организации с учетом специфики переходного состояния.

Анализ функций управления хозяйственных субъектов и правила их разнообразной композиции позволяет не только сформулировать, но и реализовать эволюционный функционально-структурный подход к созданию эффективных информационных механизмов хозяйственных систем. Информационный механизм развития хозяйственных систем представляет собой преобразование внешней и внутренней информации об условиях и среде, ресурсах, субъектах и объектах, процессах и результатах их функционирования в информационный фактор, осуществляемое на трех стадиях (привлечения, трансформации и реализации), адекватное целям, параметрам и направлениям деятельности хозяйственных систем. В качестве основных способов такого преобразования выступают информационные технологии, а средствами – информационные системы организации. Результатом действия информационного механизма выступает

[4] См.: Иншаков О.В. «Ядро развития» в контексте новой теории факторов производства / О.В. Иншаков // Экономическая наука современной России. – 2003. – № 1. – С.11-25

интегрированное и структурированное единое информационное пространство хозяйственной системы.

Комплексный, междисциплинарный подход к формированию информационного механизма развития хозяйственных систем различного уровня и оценке эффективности действия этого механизма, предполагает изучение объектов воздействия (информация и информационные ресурсы), форм, методов, средств (информатизация и информационные технологии), результатов (информационное пространство), критерия оценки эффективности информационного механизма (оценка уровня информационного развития).

Информационное пространство целесообразно определять как вид экономического пространства, выделенный на основе признания эндогенности информационного фактора производства, включающий отношения хозяйствующих субъектов по поводу, как этого фактора, так и соответствующих условий, ресурсов и продуктов их деятельности. Применение структурной рекурсии к хозяйственным системам позволяет получить для каждой из них на разных уровнях внутрисистемное информационное пространство, отражающее аспекты их внутренней среды. Таким образом, возникает иерархия «вложенных» информационных пространств различных уровней, которая отражает гетерархию аспектированных подсистем внутри хозяйственной системы каждого уровня, что обусловливает и обеспечивает их функционирование.

Трансакционные издержки являются параметрами, определяющими специфику отношений между субъектами регионального экономического пространства, для функционирования которых характер информационного обмена как неотъемлемый атрибут их взаимодействий имеет решающее значение. Информационный фактор как завершающий эндогенный наряду с организационным и институциональным обеспечивает трансакции, многообразно связывая производителей и потребителей в пространстве воспроизводственных процессов. Это будет выражаться в характере и структуре обменов и взаимодействий между отдельными информационными системами агентов в информационном пространстве региональной экономики.

Характеристики информационного пространства региональной экономики должны системно выделять и отражать его аспекты в рамках территориального, функционального и эволюционного подходов, при доминировании каждого из которых оно соответственно определяется как:

- статическое, или насыщенное, метрически характеризующее совокупность дифференцированных субъектов и объектов территории;
- динамическое, или деятельное, характеризующее интеграционные свойства функционирующих объектов и субъектов в зависимости от степени их структуризации в предметно ориентированной информационной системе;
- развивающееся, или необратимо изменяющуюся реальную предметную область, представленную в адекватной понятийной модели, отражающей направленное взаимодействие его субъектов и объектов во времени.

Применение к региональному информационному пространству указанных подходов дает возможность уточнить инструментарий для анализа взаимодействий, складывающихся между экономическими субъектами мезоуровня. Накопление и структурирование информации как эндогенного фактора происходит в регионе кооперационно с другими факторами, функциями и структурами в развитии хозяйственной системы в целом, отдельных ее подсистем и субъектов.

Эволюционный подход позволяет построить систему информационного пространства региона, где каждый объект выступает как модель, отражающая взаимодействия его субъектов. Эффективность развития региональной хозяйственной системы зависит от качества, масштабов и интенсивности

информационного обмена между его агентами, а также от ресурсов и условий распространения потоков информации в данном информационном пространстве. В процессе трансакций за счет кросс-взаимодействий информационного фактора его субъектов происходит увеличение информации, как условие и результат необратимости эволюции хозяйственной системы региона. Информационные образы субъектов и объектов региональных хозяйственных систем раскрываются в соответствующей его информационному пространству предметной информационной системе.

Ментальная модель информационного пространства региональной экономики, как подсистемы мезоэкономического пространства, может быть представлена в форме понятийной модели предметной области. Эта модель, накапливая информацию, видоизменяется, оказывая обратное влияние на эволюцию экономического пространства региона. Чем точнее информационная модель отражает содержание региональной хозяйственной системы, тем больше потенциал эффективной деятельности его субъектов, что обеспечивается принципом соответствия информационного пространства другим факторным видам хозяйственных пространств.

Концептуальное моделирование процесса региональной информатизации позволяет комплексно представить ее как систему, элементы которой на любом уровне декомпозиции определяют состояние информатизации региона. Уровневая структуризация информатизации как объекта управления позволяет выделить в качестве первого иерархического уровня стратегические мероприятия по обеспечению информатизации региона, декомпозиция которых обусловливает тактические мероприятия второго уровня – основных подсистем концептуальной модели информатизации (информационные потребности, информационный потенциал и инфраструктура). Элементы третьего, оперативного уровня управления, фактически определяют состояние подсистем верхнего уровня и, являясь индикаторами, характеризуются наибольшими динамическими изменениями вследствие воздействия внешней среды. Состояние всех элементов концептуальной модели на каком-либо уровне декомпозиции определяет уровень информатизации региона как объекта управления.

Наиболее существенные количественно-качественные характеристики состояния региональной информатизации отображает система показателей развития информационного фактора региональной хозяйственной системы, состоящая из двух взаимосвязанных подсистем: показателей уровня информационных потребностей в регионе и степени их удовлетворения. Показатели, представленные в модели, рассчитываются на основе данных Госкомстата и других открытых источников, что стандартизирует методику сбора первичной информации. Первая подсистема отражает существующие информационные потребности общества и связана с уровнем образования и развития научной деятельности в регионе. Вторая подсистема состоит из показателей информационных ресурсов, уровней финансовой и технической обеспеченности, а также уровня использования результатов информатизации, обусловленного степенью ее социализации и профессионализации.

Комплексная система мониторинга информатизации региональной хозяйственной системы состоит из дескриптивного анализа состояния региональной информатизации, многофакторной рейтинговой оценки состояния информационного фактора развития региональной экономики и многофакторной ранговой оценки регионов по уровню информатизации. Уточненная методика рейтинговой оценки состояния информационного фактора развития региональной экономики основана на применении кластерного анализа и рейтинг-технологий.

Результатом многофакторной рейтинговой оценки состояния информационного фактора развития региональной экономики является построение общего и частных рейтингов регионов (макрорегионов), ранговой оценки – присвоение регионам рангов по степени отклонения от стандартного значения показателей. Для определения мер регулирования информационных процессов региональной экономики необходима оценка эффективности информатизации в процессе мониторинга, состоящего из дескриптивного анализа состояния региональной информатизации, который представлен организационно-экономической характеристикой проводимых в Южном федеральном округе мероприятий по информатизации в рамках региональных и федеральных целевых программ.

Рейтинговая и ранговая оценки состояния информационного развития региональных хозяйственных систем позволяют группировать региональные хозяйственные системы, выявлять динамику их развития, оптимизировать управленческие воздействия на различных уровнях его организации и выявить две функциональные группы показателей: определяющие макросостояние процесса региональной информатизации и определяющие направления управленческих воздействий в регионах, нуждающихся в регулировании.

Развитие региональных хозяйственных систем неизбежно будет сопровождаться переходом на более высокий уровень контекстно зависимого абстрактного понимания, определения и исследования их информационной структуры. Решение проблемы информатизации региона требует осознания новых концептов единого информационного пространства и информационного общества. Фиксация нового знания и практическая реализация его в форме новых культурных ценностей, институциональных структур и высоких технологий требует многоаспектного информационного отражения и представления аспектированных хозяйственных систем регионального уровня, интеграция информационных моделей которых обусловливает формирование и развитие единого информационного пространства.

АВТОМАТИЗИРОВАННОЕ ОПРЕДЕЛЕНИЕ ТЕМАТИКИ ПУБЛИКАЦИЙ ПРИ БИБЛИОГРАФИЧЕСКОЙ ОБРАБОТКЕ

В.С. Егоров
Т.Н. Чернобровская

AUTOMATED DETERMINATION OF PUBLICATIONS SUBJECTS IN THE COURSE OF BIBLIOGRAPHIC PROCESSING

V.S. Egorov, T.N. Chernobrovskaya

Определение тематики публикаций – одна из ключевых задач в процессе обработки научно-технической литературы. От точности ее решения зависит качество подготовки информационных продуктов ВИНИТИ и обслуживания потребителей. При точном глубоком индексировании конечный пользователь не будет испытывать затруднений в поиске необходимых ему сведений, несмотря на постоянно увеличивающиеся объемы НТЛ. При хорошем знании Рубрикатора ВИНИТИ он легко может собрать информацию по малознакомой тематике.

Выявление тематики – многоэтапный процесс, который начинается с грубой оценки тематической направленности публикации (тематическая разметка на входе) и заканчивается точным индексированием по Рубрикатору (специалистами в отделах научной информации). На разных этапах обработки для определения тематики доступны различные сведения о публикации и это является предметом сложных процессов аналитико-синтетической переработки в любом

информационном центре. Специалист, выполняющий разметку, должен обладать энциклопедическими сведениями, знать достаточно глубоко не только одну научную отрасль, но и все соприкасающиеся с ней. А на первичном этапе разметки обладать знанием всех отраслей науки и техники. К сожалению, такого рода специалистов не готовит ни одно учебное заведение и процесс вхождения в специальность происходит в ходе практической работы.

ВИНИТИ, имея значительный потенциал в создании автоматизированных технологий обработки научно-технической литературы, уже давно ведет исследования в области автоматизации процессов разметки. Особенно эта проблематика была популярна в 70-х, начале 80-х годов, т.е. в начале массового внедрения компьютерной техники. Работы носили теоретический характер с небольшими модельными экспериментами. Полномасштабного внедрения не было по крайней мере по двум причинам: отсутствия на входе института потока НТЛ в электронной форме и слабость компьютерной техники. Сейчас в условиях появления значительного потока литературы в электронной форме, создания эффективной системы автоматизации процессов обработки входного потока ВИНИТИ (система ВХОД), резкого сокращения штатов интерес к этой тематике приобретает практический интерес. Институту необходимы решения, дающие реальную экономию трудовых ресурсов. Работы в этом направлении начаты в 2000 году, при этом их можно разделить на две составляющие:

- автоматическая классификация, основанная на элементах искусственного интеллекта,

- автоматическая классификация, основанная на использовании классификационных меток проставленных в экземплярах научно-технической литературы при их издании сторонними организациями, реализующая сложившиеся в институте схемы обработки НТЛ, использующая элементы библиографического описания литературы.

Безусловно, первое направление является более сложным и перспективным, однако ожидать реального практического внедрения достаточно сложно, хотя одна из подобных систем уже находится в опытно- промышленной эксплуатации. Настоящая статья посвящена описанию результатов, внедренных в промышленную эксплуатацию по второму направлению. Эта технология является частью Автоматизированной системы комплектования и регистрации входного потока (АСКР), которая эксплуатируется в ВИНИТИ с 2000 г., и позволяет создавать механизмы автоматической тематической разметки на основе библиографических сведений об обрабатываемых изданиях.

Из всех видов обрабатываемой литературы были выделены те из них, которые имеют в исходном документе классификационные рубрики или специальную терминологию, или объединены общностью замысла, тематики, целевым или читательским назначением, выходящим под общим названием и в однотипном оформлении.

Результатом визуальной оценки первоисточников явилось выявление в БО тех элементов данных, которые можно использовать для определения тематики, т.е. выявление тематически нагруженных элементов.

Такими элементам являются:

- Код вида издания
- Классификационные индексы:
- Организации, участвующие в подготовке издания:

Коллективный автор публикации
Место выполнения опубликованной работы
Место защиты диссертаций
Издательство

При наличии признака тематической размети одновременно в нескольких элементов данных, ее выбор осуществляется строго в указанной выше последовательности.

ЭД, участвующие в процедуре автоматической разметки (профилирования), снабжены специальным признаком, значение которого соответствует штампу Отдела научной информации (ОНИ) по соответствующей тематике. Согласно принятой в ВИНИТИ технологии одноразового реферирования существует понятие 1-го ОНИ, в реферативных изданиях которого отражается реферат на первоисточник, и понятие 2-го и последующего ОНИ, которые заимствуют данный реферат.

Рассмотрим технологию автоматической разметки для разных ЭД.

1. Профилирование по коду вида издания

Код вида издания (поле 626) – обязательный элемент БО. В автоматическую разметку включены следующие виды издания.

Картографические виды издания (626=*КГИ*). Автоматическая разметка картографических изданий осуществляется с помощью специального маршрута, который автоматически проставляется им при регистрации. В качестве 1-ого ОНИ этому виду изданий проставляется штамп ОНИ по географии, а в 2-ми ОНИ определены ОНИ по геологии, геофизики (астрономии) и охране окружающий среды.

Издания книжного типа, а именно книжные серии (626=*ИКТ.СЕР*). Выпуски книжных серий составляют заметную часть входного потока изданий книжного типа. Традиционная технология библиографической обработки таких изданий предусматривает заполнение элемента данных «сведения о серии» (название серии, коллективный автор и издательство). При этом каждый раз значение этих полей вводилось заново независимо от ранее поступивших выпусков той же серии. Такая технология приводила к разночтению описания книжной серии и, как следствие, не позволяла группировать выпуски одной серии, а для номерных серий не было возможности следить за их «лакунами».

Между тем описание серии является довольно стабильным. При наличии соответствующего аппарата накопления и хранения сведений о книжных сериях можно, во-первых, существенно сократить трудозатраты на ввод библиографического описания выпуска, а, во-вторых, - использовать общие данные для определения тематической направленности выпусков одной серии. С этой целью в состав описания книжной серии был включен специальный элемент «Признак автоматической разметки» - поле kind_razm и поле “Код профильности” (штамп ОНИ). Значение поля kind_razm = 1 означает, что данное издание должно быть направленно на реферирование в указанное ОНИ.

Периодические издания - (626=*СИ*). На основе статистической обработки периодических изданий были выявлены так называемые «моноразметочных изданий» - такие, которые отражаются в выпусках Реферативного журнала только одного ОНИ на протяжении 3-х лет Аппарат автоматической разметки данного вида издания аналогичен книжным сериям.

2. Профилирование по классификационным индексам.

Автоматическая разметка по классификационным индексам осуществляется для авторефератов диссертаций и государственных и отраслевых стандартов. Авторефераты отечественных диссертаций в качестве обязательного элемента БО содержат специфический элемент «Номер специальности» по классификации ВАК, а стандарты имеют индекс «Общероссийского классификатора стандартов». Для автоматической разметки этих видов изданий были созданы специальные словари «Классификатор ВАК» и «Классификатор НТД». Большая часть этих словарей имеет прямо выраженную тематику, которая проставлена в специальные поля «Штамп ОНИ1» (однозначные) и «Штамп ОНИ2» (многозначные). Наполнение указанных полей осуществляется строго по словарю SUBJECT (Тематика (коды, штампы)). Код ОНИ1 и ОНИ2 проставляется только в том случае, если практически все издания по этой рубрике идут в конкретные ОНИ. В этом случае и происходит автоматическая разметка.

Для изданий, у которых рубрика однозначно не определяет ОНИ, в код ОНИ1 проставляется признак «ручной разметки» - ТО! Например, для рубрики VAK 05.02.01 – «Материаловедение (по отраслям)» осуществляется традиционная ручная разметка.

Авторефераты диссертаций, тематика которых не соответствует Рубрикатору ВИНИТИ (например, «Литературоведение»), автоматически исключаются из дальнейшей обработки, и им автоматически присваивается маршрут «Научные фонды» (НФ).

Если в БО документа была допущена полиграфическая ошибка в индексе классификации, тогда система автоматической разметки определяет его как несуществующий. Для таких случаев были введены в словари «псевдокод», например 00.00.00 в словаре «Классификатор ВАК» и 00 в словаре «Классификатор стандартов», на их основании такой документ передается на участок традиционной ручной разметки.

3. Профилирование организаций

Профилирование организаций распространяется на издательства, коллективных авторов, центры-депозитарии, места защиты и выполнения работы, организаторов мероприятий, так как для них есть основания предполагать специализацию по тематике.

Профилирование организаций основывается на статистической обработке входного потока массива документов, обрабатываемых в ВИНИТИ за последние четыре года, и на принципе представительности выборки достаточного количества наименований публикаций, которые подготовлены к изданию данной организацией.

На основе анализа статистических данных, снятых с таблицы analit.ORG_SHTAMP, были выявлены организации, которые готовят информационную продукцию по одному устойчивому тематическому профилю. Устойчивые тематические профили были занесены в описания объектов-организаций.

Как и в случае с сериальными изданиями и книжными сериями в состав описания объектов-организаций был включен специальный элемент «Признак автоматической разметки» - поле kind_razm. Значение поля kind_razm = 1 означает, что для издания, подготовленного данной организацией, может быть применена процедура автоматической разметки, и для него выбирается соответствующий код профильности. Это в первую очередь место защиты иностранных диссертаций, у которых в этом поле указан факультет.

1. Дивильковская Т.Ю., Козачук М.В., Чернобровская Т.Н. Классификация изданий книжного типа в БД ВИНИТИ //Сб. тр. Международной конф. «Информационное общество. Интеллектуальная обработка информации. Информационные технологии. (НТИ-2002), 16-18 октября 2002, с,126, ВИНИТИ, М. 2002
2. Чернобровская Т.Н., Денисова Л.А., Пономаренко Т.П. Обработка картографических изданий в АСКР ВИНИТИ //Сб. тр. Международной конф. «Информационное общество. Интеллектуальная обработка информации. Информационные технологии. (НТИ-2002), 16-18 октября 2002, с. 376, ВИНИТИ, М. 2002
3. Шапкин А.В. Автоматизированная система комплектования и регистрации входного потока ВИНИТИ. Ч. 1. //НТИ. Сер. 1. - 2005 - № 3. С. 8-19, М. ВИНИТИ
4. Шапкин А.В. Автоматизированная система комплектования и регистрации входного потока ВИНИТИ. Ч. 1. //НТИ. Сер. 1. - 2005 - № 4. С. 16 - 31, М. ВИНИТИ
5. Чернобровская Т.Н., Блат К. Д., Маркова Л. И. Словарная база АС «Вход» //Сб. тр. Международной конф. «Информационное общество. Интеллектуальная обработка информации. Информационные технологии. (НТИ-2002), 16-18 октября 2002, с. 377, ВИНИТИ, М. 2002

SCIENTOMETRICS DEVELOPMENT (REVIEW OF KAZAKHSTAN AND FOREIGN EXPERIENCE)

G.Zh. Alibekova
National center for scientific and technical information of the republic of kazakhstan
Almaty, Kazakhstan

The current stage of Kazakhstan development is characterized by sharp attention to state of scientific potential of the country. USSR collapse brought also downfall of science: outflow of scientists from laboratories to business and other quick-profit spheres, weaken research ethics, getting scientific degrees not for science but for career, less publications in local and foreign editions, less collaboration among CIS and other foreign countries, out-of-date research equipment, etc. All these factors made our R&D potential weaker and lower. Qualitative studies confirm that fact. But quantitative assessment is still not developed.

Nowadays, almost any kind of development is being quantitatively evaluated. Indices of competitiveness and human development, indices of technological achievements and sustainable development are examples of those attempts. One of them is successful and universal; others require additional efforts to be adapted for individual conditions of different countries.

Growing interest in science development and in it’s measuring is due to expanding tendency of the world economy to innovative development concept. This is the next wave of interest to science in the world. The first one was observed in 1940s when the number of science institutions, science employees was growing rapidly and R&D expenditures were increased in large extent. Exceptional role of science in technical progress, upgrade of production and other factors made the subject of science management more actual and important. As a result the new idea of developing Science of Science (naukovedenie) was created. Later science of science divided into several directions: scientometrics, psychology of scientific creativity, sociology of science, history of science, economic

issues of science development. This article is focused on some aspects of scientometrics development.

J. Bernal is considered as a founder of the ‘science of science”. He published his “Social function of science” in 1939, where he emphasized a science as a special object for researches.

E.Garfield and D.Price contributed a lot into development of science of science. In particular, D. Price developed his theory of the exponential growth of scientific literature and tied it to the increased number of indexing and abstracting services. D.Price is recognized as “a father of scientometrics”. Garfield created his Science Citation Index (SCI) - unique multidisciplinary indicator of citations. That was an attempt to quantitatively evaluate potential of science. His Institute for Scientific Information (ISI) issues SCI since 1964/[1]. Creation of SCI made easy to calculate impact factor, which measures importance of scientific journals. He is recognized as one of the founders of bibliometrics and scientometrics.

Scientometrics is “the study of the measurement of scientific and technological progress” (Garfield,1979b). Scientometric research has a strong application-oriented tradition. For example, scientometric studies may help governments and private sectors identify their competitive edges and make strategic plans for future research areas and allocate research funding to key research areas [2].

Nowadays, the most developed field of scientometrics is bibliometrics.

The first studies on measuring science in Russia were made by Nalimov V.V., who suggested measuring scientific potential through number of foreign awards of scholars, number of publications and citations [3]. Later professor Nalimov initiated using statistics methods to describe chemical processes and scientometrics.

Although Nalimov was dismissed from the second year of study at physics and mathematics department of Moscow State University named after Lomonosov in 1930 he could defend his dissertation work devoted to methods of mathematic statistics in studying mistakes of spectral and chemical analysis in 1957 . In 1959 he published his first work in scientometrics jointly with G. Vleduts and N. Styazhkin, scientists had been working at All-Russian Institute of Scientific and Technical Information (AISTI). One of chapters of that work described issues of mathematical models of the world science development, where ideas of D. Price were supported and developed. He also realized that quantitative study of science can not be successful without well-organized information services and paid a lot of attention to questions of accumulating, analyzing and using information. In 1961 he organized research section named “Chemical cybernetics” in frame of which conducting scientometrics researches became possible. Thus, that theme was officially recognized.

In the mid of 1960s at the Institute of nature science and engineering history, operating under USSR Academy of sciences, department of the Science of science was established. Center of S&T potential and science history study was opened in Ukraine Academy of science. At the Soviet-Poland symposium hold in L’vov in 1966 Science of science was recognized as a new field of study.

Nalimov also took active part at the symposium. And during discussions he first introduced “scientometrics” as the new name for the new field of study engaged in quantitative evaluation of science development.

In his article “Quantitative methods of science development study” Nalimov outlined importance of what to measure and how to measure. Not less significant is to be able to evaluate in what extent those quantitative parameters really reflect science development. Nalimov emphasized most widespread parameters are number of publications, number of science staff, scope of science groups, etc. In that article Nalimov defined tasks of scientometrics. Except abovementioned those are drawing and analyzing growth curves for national science, studying research organization structure, studying problems of staff

hiring and training, science development prognosis and management. As for science management it is important to solve two issues: 1) how much should be granted to research development by the government and how those funds should be allocated among different fields of science; 2) how to organize different science fields quantitative monitoring system.

After that article Nalimov jointly with Z. Mulchenko published another articles repeating and developing initial ideas on language of science citations in scientometrics, science study as an informational process, subjects and tasks of scientometrics. In 1967 article of "invisible collective"[5] members was published, where journals were studied as communication channels and contribution of certain countries into the world information flow was evaluated. In that article the information model of science development process is clearly defined and journals are considered as communication channels and bibliographic references - as a special language of science information indicating publication effects on the world information flows development. To evaluate contribution of different countries into the world science information flow it was suggested to consider 1) efforts of a certain country in developing its science and 2) effectiveness of those efforts. Relative number of publications reflected in reference journals was taken as a measure of those efforts. And relative number of bibliographic references to authors of different countries was taken as a measure of the efforts' effectiveness.

As a result it was revealed that USA and Great Britain on the one hand and USSR on the other hand created two information flows overlapping in a little extent [4]. American - Britain publications were about 55% of all publications and USSR publications – about 20%. Indicator of English language publications' citing was found at the same 55%, and USSR publication's citing was at about 3-4%. To compare, German language publications had 5-10% of citing that was rather good indicator and French publications had less than 5% of citing.

Low level of Soviet publications citing was explained by poor organization of information services. According to I. Granovskiy there are other factors of low effectiveness of the Soviet science except poor information services.

In another article "What do we expect from scientometrics?" Nalimov refers to the theme of research group age. Authors thought that success in research field depends on emotional set up of scientists and mostly young people are able to stay long at high emotional state. Thus, effectiveness of research groups depends on their age and inflow of young scholars should have enough speed to keep average age of research group at appropriate level.

As a result scientometrics was expected to make multisided qualitative analyses of science development process. Such analysis is a base for effective science management. But there are a lot of organizational and technological barriers to conduct academic works in this field in Kazakhstan as well as in Russia and other CIS countries. In Kazakhstan there were no any deep researches in scientometrics. In contemporary, Russia has rather long history of scientometrics development.

According to L. Gohberg, Director of Institute for statistics researches and knowledge economy in Russia, there are two kinds of problems in Russia. First, is analysis of those bibliometric data which are in database of for example Institute of scientific information in Philadelphia, European data base SCOPUS and other specialized databases in different science fields– Chemical Abstracts, Inspec, Medline. In this regard it is important to solve questions of those data extraction, interpretation, processing and even linguistic difficulties in identifying authors, organizations, etc.

---

[5] The term of '"invisible collective" was introduced by D.Price and indicated groups of researchers from different organizations and countries working on the same issues.

Second, formation of Russian own sources in credit of scientific community would be recognized as a single database for such kind of analytical works. In this regard a lot of problems up to legal ones (as defining such database status, journals that must be included into it) should be solved.

Another problem is a majority of Russian scientific journals is not referenced. Thus credibility of such publications quality becomes questionable. Moreover, journals are not presented in electronic form that leads to technological problem of entering that information.

Currently, Russian Foundation of Fundamental Researches (RFFR) is engaged in bibliometric studies. It is rather representative sample, but it includes publications of only those who apply to RFFR and therefore it is just a part of Russian researchers. Database of RFFR started to 1994 and reliable statistics come from 1997. In 2001 RFFR had been required to grant publications in electronic form. Alexandr Libkind and his son Iliya are authors of RFFR Index, which was highly evaluated in 2003 by Dr Garfield, the founder of bibliometrics [4].

Nowadays, RFFR Index covers 60% of Russian more active scientists and a half of it overlaps with SCI.

Bibliometrics in Kazakhstan is in its start of development phase. National Center for scientific and technical information of the Republic of Kazakhstan (hereinafter “Center”) is engaged in bibliometric researches.

Bibliographic database of the Center includes data of 2413 journals, where Kazakhstan authors have publications [5]. The database contains more than 50 ths. references including 15,8 ths. references on publications of Kazakhstan authors.

Analysis of the database showed that publications in metallurgy have higher citation index (about 0,67), it means that Kazakhstan authors were more cited than foreign ones. The lowest citation index was observed in geology (0,28). In contemporary, it shows foreign authors were more cited than local authors.

In 2006 study of agrarians’ citation started. Average number of references per article was: in food industry – 2,4; veterinary – 3,1; animal husbandry – 1,8; mechanization and electrification of agriculture – 2,7; farming, plant-growing – 3,3 .

In general, indicators of the Center’s bibliographic database are presented in the Table 1.

Table 1. Database on Citation of Kazakhstan authors (1994-2006)

Field	Journals	Articles	Total number of references	References to KZ authors
Chemistry, environment	1047	5226	30484	7687
Industry	950	4026	18200	6345
Agriculture and food industry	301	1150	3134	1042
Physics and mathematics	115	335	2488	787
Total	2413	10737	54306	15861

Source: National Center for scientific and technical information, Kazakhstan 2006

In general, scientometrics study in Kazakhstan is not developed. Given Bibliometrics research was made by the order of the Ministry of education and science of Kazakhstan. This study was a first attempt of research in this field and it served as a driving force of interest in this field for many of researchers. We hope that this interest will lay the fundamental for wider research in scientometrics, which is recognized as a base of science development.

Reference:

1. Российские публикации и их цитируемость в мировом научном сообществе, http://www.ng.ru/science
2. Mapping *Scientometrics* (1981-2001). Chaomei Chen, Katherine McCain, Howard White, Xia Lin.
College of Information Science and Technology, Drexel University, Philadelphia, PA 19104-2875, USA.
3. Можно ли измерять науку? Исследования В.В. Налимова по наукометрии. Ю.В. Грановский
4. New scientometrics data on Russia and other countries. V. Markusova.
5. “Development of the State System of S&T information”. Republican Target S&T Program.
National Center for scientific and technical information of the Republic of Kazakhstan