

УДК 002
ББК 73

О МЕТОДАХ И ПРОЦЕДУРАХ ИНФОРМАЦИОННОЙ ДИАГНОСТИКИ ОБЪЕКТА

© Г.Ф. Гордукалова, 2008

*Санкт-Петербургский государственный университет культуры и искусств,
191186, г. Санкт-Петербург, Дворцовая набережная, 2*

Описана технология рационального соединения специфических приемов и методов для информационной диагностики объектов разной природы. Определены главные условия реализации процедур информационной диагностики. Описаны три основных режима диагностики: ретроспективный анализ объекта, его оперативная диагностика и мониторинговое наблюдение за объектом.

Ключевые слова: информационная диагностика, сетеметрия, ретроспективный анализ, оперативная диагностика, мониторинг.

При специальном рассмотрении вопрос о методах – один из сложнейших для научного направления. Существует множество классификаций методов прогноза, библиометрических методов, общенаучных методов исследования. В каждой отрасли знания складывается собственная типология методов анализа – методы экономического, литературоведческого анализа, исторические, лингвистические методы исследования и т. д. Специалисты отрасли стремятся выделить специфические методы и приемы исследования или профессиональной деятельности, определяя их суть и отличительные черты. Описанные методы обозначаются наиболее распространенным или смысловым наименованием, либо они вводятся под именем впервые применившего их исследователя. Нередко в тематически удаленных областях деятельности возникают содержательно близкие методические подходы к анализу разных явлений. В частности, постоянно применяются без четких разграничений термины «библиометрические», «информетрические», «наукометрические» методы. Как правило, они обозначаются по конечной цели или объекту исследования, поэтому необходимо понимать, что, в сущности, это одни и те же методы, применяемые для анализа потока информации об объекте при решении конкретных задач. В 1970–80-х гг. свои «метрии» стали активно развивать многие отрасли знания – эконометрия, биометрия и др.

Происходит постоянное заимствование методов из других отраслей знания для исследовательских задач в системе информационных коммуникаций. Например, в 1990-х гг. были использованы:

- анализ важнейших событий (critical incidently) – привнесен из теории коммуникаций для вы-

явления межличностных коммуникаций с выделением важнейших успехов и неудач при взаимодействии библиотекарей и пользователей в ходе справочного обслуживания в научной библиотеке [1];

- анализ дискурса (discourse analysis) – заимствован из теории коммуникаций для библиотечно-информационных исследований, в частности для изучения устной и письменной коммуникации и уточнения задач информационных исследований [2]. Он близок к известному ретроспективному анализу документального потока по конкретной проблеме;

- экзистенциальные графы Пирса (existential graphs) – рождены для изучения синтаксических систем, а предлагаются в качестве интерфейсов при вводе в информационные хранилища, выводе естественно-язычной информации и доступе к ней без промежуточных процедур, если будет разработана семантическая структура экзистенциальных графов [3].

Значительный толчок для информационных исследований дан благодаря развитию программного обеспечения систем интеллектуального типа. В частности, в 1997 г. обсуждались первые результаты применения ДСМ-метода (автоматического порождения гипотез) для автоматической классификации полнотекстовых документов – диссертаций по теоретической и прикладной информатике, документов по проблеме машинного обучения [4].

В настоящее время закладываются основы «сетеметрии», и многие из названных методических подходов применяются для анализа информации в сетевом режиме. При этом авторы исследований отмечают, что общий методологический подход сравним с библиометрическим анализом баз данных в целом. Опубликованы результаты его при-

менения к персональным страницам – www: используется статистическая частота слов и словосочетаний, а также авторов, источников публикаций, их цитирования – т. е. в основе вновь лежат приемы библиометрического анализа. Отличие заключается в наличии программных средств для автоматизированного анализа массивов, в употреблении ярлыков, добавлении к информационным объектам кодов языка разметки гипертекста HTML [5].

Следует отметить и еще одну черту современного подхода: после экстенсивного развития методического аппарата в аналитико-синтетической обработке информации (бесконечного дробления приемов анализа данных об объекте, создания множества частных методик вокруг некогда единого подхода), благодаря синтезирующим программным комплексам для обработки данных, активизировался *процесс слияния методов и приемов анализа в общие технологические цепочки*. В результате такого подхода нами и разрабатывалась в начале 1990-х гг. технология информационного мониторинга, способная объединить разнообразные методы библиографической, статистической и интеллектуальной обработки документов (данных) об объекте.

Таким образом, в настоящее время методический аппарат библиометрии становится общенаучным, обогащается методами других отраслей знания, требует не столько дифференциации по области применения, сколько интеграции по иным (процессуальным и содержательным) признакам в рамках общего направления – информационной диагностики объектов. В связи с этим в данном случае речь идет не о типизации распространенных общенаучных и отраслевых методов, которые могут быть использованы в информационной диагностике объектов разной природы, а о рациональном соединении некоторых специфических приемов и методов информационной диагностики объекта в целевые технологические цепочки.

Информационная диагностика – технология анализа информации об объекте, реализуемая с помощью специальных процедур и методов в фиксированном предметном поле по избранным индикаторам для оценки состояния, тенденций и перспектив развития объекта.

Диагностироваться может объект любой природы, если он изменчив с течением времени (или обогащаются наши представления о нем), а сведения об этих изменениях имеются в документальном потоке. Социальные и природные явления, процессы, ситуации; искусственно созданные материалы, технологии и системы; свойства объектов, связи и отношения – т. е. все, с чем связан человек в процессе своей жизнедеятельности, находит свой документальный отпечаток, но фиксируется с разной степенью достоверности, полноты и

оперативности сведений об объекте. Определим главные условия реализации процедур информационной диагностики:

1. Анализируется не сам объект, а известная в обществе информация о нем.

2. Для успешной диагностики должна обеспечиваться определенная полнота сбора документов (данных) об объекте, границы которого заранее описываются информационным аналитиком в виде упорядоченного перечня ключевых слов или проблемно-ориентированного рубрикатора, фиксирующего предметное поле объекта.

3. Диагностические признаки – оценочные индикаторы – могут быть выбраны из числа признаков объекта – его свойств, качеств, физических параметров, особенностей взаимодействия с другими объектами. Они должны быть формализованными, не избыточными, достаточно адекватно отражать изменчивость объекта или знания о нем.

4. Технологическая цепочка методов и процедур диагностики должна позволять информационному аналитику осуществлять диагностику незнакомого объекта с воспроизводимыми результатами оценки объекта.

По хронологическому признаку исследуемого периода в жизни объекта выделим три основных режима диагностики:

- * ретроспективный анализ объекта,
- * оперативная диагностика объекта,
- * мониторинговое наблюдение за объектом.

РЕТРОСПЕКТИВНЫЙ АНАЛИЗ предполагает выявление закономерностей развития объекта за предшествующий период времени. Ориентирован на выявление «критических» временных точек в развитии объекта, существенно изменяющих его состояние или содержание знания о нем. В информационной диагностике целесообразен обратнo-хронологический сбор источников информации об объекте, позволяющий прервать процесс накопления информационного массива при получении достаточных сведений о нем. Используя совокупность процедур информационного моделирования объекта (поиск, отбор, аналитико-синтетическая обработка и систематизация документов/данных) с заданными критериями подобия на хронологической шкале, аналитик может варьировать набор конкретных методических процедур с учетом специфики диагностируемого объекта и его документального шлейфа. В технологическую цепочку ретроанализа легко вкладываются процессы создания ретрорубрикатора, хронособытийная и хронолингвистическая шкалы, являющиеся основным методическим средством диагностики объектов. Построение ретрорубрикатора эффективно вести в прямой хронологии от начальной точки зарождения объекта (например, 1703 г. – по проблеме «Наводнения в Санкт-Петербурге»), априорно задан-

ной точки на временной шкале (наводнения в Санкт-Петербурге последнего пятилетия) или фоновой, после которой существенно изменялось поведение объекта (строительство дамбы, значительно меняющей картину наводнения). Ретрорубрикатор представляет собой хронологическую шкалу, куда вводится каждое новое понятие, предметная рубрика из заглавия или текста профильных документов с указанием даты опубликования работы. Такой рубрикатор студенты готовят за 2–3 часа практических занятий о незнакомом для них диагностируемом объекте. При наличии машиночитаемых библиографических массивов анализируемого периода возможно построение хронолингвистической шкалы с частотностью употребления ключевых слов, а также персон, событий, свойств объекта.

Ретрошкала свойств объекта строится на основе метода поаспектного анализа текстов каталогов, стандартов, рекламных материалов и обзорно-аналитических статей о нем.

Для сложных социальных объектов уместно использовать менее жесткие приемы формализации данных – типологические ряды, номинативные шкалы. Может быть употреблен неординарный метод воссоздания ретросценария прошедших событий с поаспектной фиксацией лиц, событий, следствий на хронологической шкале. Ретроанализ позволяет также оценивать вклад в развитие знания об объекте отдельных стран, фирм, исследовательских коллективов, авторов на основе метода длительности ретроцитирования и коцитирования их работ.

При сборе и интерпретации данных ретроанализа принимается решение о выборе «хронологического шага». В зависимости от величины анализируемого периода могут использоваться методические подходы, существенно различающиеся по стратегии сбора и доле анализируемой информации:

РАВНОМЕРНЫЙ ТОЧЕЧНЫЙ прием предполагает фиксацию индикаторов в *определенных* временных точках на шкале – через десятилетие, пятилетие, ежегодно. Для ретроанализа динамики цен на товары и архива биржевых котировок акций выбирается ежемесячный, еженедельный, ежедневный и даже почасовой «пик». На результаты диагностики весьма повлияют «пробелы» – отсутствие информации об объекте в заданных точках.

КАПЕЛЬНЫЙ метод целесообразен при диагностике стохастических процессов, а также при углубленном ретроанализе объекта в критических точках его развития. Избирательно фиксируется информация об объекте в любом месте хронологической шкалы, сравнивается с предыдущей «каплей» информации до относительно стабильного состояния объекта, которое и принимается за ис-

ходное, фоновое. Периоды времени могут быть существенно неравными. Длительность периода значения не имеет и определяется *отсутствием* событий в жизни объекта.

ПОШАГОВЫЙ метод – длительность «шага» в развитии объекта выявляется эмпирически в обратной или прямой хронологии анализа информации. В отличие от предыдущих подходов, он предполагает неравный по времени, но планомерный по всей шкале ретроанализ объекта. Каждый последующий шаг возможен после синтезирования данных о предыдущем шаге. Это наиболее сложный метод информационной диагностики. Может быть приложен к социальным и иным кризисным ситуациям. Например, развитие августовской кризисной ситуации 1998 г. и ее последствия для российского рынка ценных бумаг. Подобная диагностика была осуществлена нами на основе поаспектного анализа газетных публикаций и факторного анализа полученных данных, выявившего незримые связи политических лидеров страны.

В результате ретроанализа информационный аналитик имеет в своем распоряжении материал для выделения этапов развития объекта, их сводной характеристики, выделения тенденций развития объекта за анализируемый период времени, а при использовании методов информационного прогнозирования может вычленять и перспективные направления его дальнейшего существования в определенных условиях. Параллельно может быть проведена диагностика вклада стран, фирм, авторских коллективов в развитие данного объекта.

ОПЕРАТИВНАЯ ДИАГНОСТИКА – экспресс-анализ состояния объекта в определенный момент его развития. Может реализовываться в любой хронологической точке, но ориентирован, как правило, на оценку современного состояния объекта, сравнительный анализ объектов-аналогов, выбор направлений развития базового объекта из группы альтернативных направлений. Предполагает четко продуманную технологическую цепочку методических приемов сбора и анализа данных под конкретную задачу диагностики родственных объектов. Из числа наиболее эффективных процедур предлагаются следующие:

- *метод «воронки»* – последовательного отсеечения элементов из одного множества по разным признакам. Весьма полезен для решения селективных задач, например выявления ведущих конкурентов на рынке;

- *номинативное шкалирование признаков объекта* (потребительских свойств товара, опубликованных мнений специалистов об объекте и т. д.);

- *методика С.К. Бредфорда*, позволяющая быстро дифференцировать родственные объекты по формальным признакам на 3–5 групп (журналы, фирмы, регионы и др.);

- *построение ризом*, объединяющих взаимодействующие объекты на основе данных цитатного анализа, поаспектного анализа обзорных документов (например, выявлять достаточные условия для реализации прогноза) и т. д.

Приведем пример простейшей технологической цепочки методов для экспресс-анализа рынка труда в регионе. Его проводят студенты на практическом занятии за два академических часа: осуществляется отбор и аналитико-синтетическая обработка объявлений региональных газет; на основе поаспектного анализа снимаются сведения о вакансиях в городе на конкретные профессии и должности; благодаря библиометрическому анализу выявляются требования к претендентам. Сравнительный частотный анализ данных за 2–3 даты текущего месяца с фоновыми данными предыдущего периода позволяет охарактеризовать в обзорной справке тенденции регионального рынка труда, показать изменения в профессиональных и должностных ориентациях работодателей.

МОНИТОРИНГОВЫЙ АНАЛИЗ – непрерывное наблюдение за развитием объекта с определенными временными интервалами в фиксированном информационном поле по заданным индикаторам. Детально технология информационного мониторинга была описана нами в ряде публикаций. В данном случае выделим лишь отличительные черты мониторингового режима диагностики. Как наиболее дорогостоящий вариант диагностики, требующий непрерывного ведения базы данных без каких-либо пробелов и систематического осмысления результатов мониторинга, данная технология может быть принята к использованию лишь по экономически значимым для организации объектам – динамика цен на продукцию, новые профильные технологии и т. д. Именно мониторинговый режим предоставляет реальную возможность информационного прогнозирования объектов. Для одноименного учебного курса создана база данных «Информационный мониторинг», позволяющая обрабатывать массивы библиометрических данных как на уровне библиографических описаний и аннотаций, так и на уровне полнотек-

стовых документов. На этапе статистической обработки данных базовым является факторный анализ признаков объекта, дающий основу для выделения, например, актуальных проблем в области гуманитарных наук, групп политических и научных лидеров по признакам их информационного поведения и т. д.

Краткий обзор технологических цепочек методов и процедур информационной диагностики показывает *возможность сочетания весьма различных процедур и методов анализа информации в рамках наиболее общих режимов диагностики*. В процессе диагностики сложных объектов методический аппарат обогащается процедурами шкалирования качественных характеристик, формально-логическими приемами анализа, специфическими процедурами – информационного прогнозирования.

Список литературы

1. *Radford, M.* Communication theory applied to the reference encounter: An analysis of critical incidents // *Libr. Quart.* –1996. – Vol. 66, N 2. – P. 123–131.
2. *Budd, J. M.* Discourse analysis: Method and application in the study of information / J. M. Budd, D. Raber // *Inform. Processing & Management.* – 1996. – Vol. 32, N 2. – P. 217–226.
3. *Burch, R.* Semantics for Peirce's existential graphs // *НТИ-97 : 3-я междунар. конф. «Информационные ресурсы. Интеграция. Технологии»* (г. Москва, 26–28 нояб. 1997 г.) : материалы конф. – М. : ВИНТИ, 1997. – С. 50.
4. *Виноградов, Д. В.* Алгебраическая модель связанных свойств ДСМ-метода // *НТИ-97 : 3-я междунар. конф. «Информационные ресурсы. Интеграция. Технологии»* (г. Москва, 26–28 нояб. 1997 г.) : материалы конф. – М. : ВИНТИ, 1997. – С. 59 ; К разработке платформенно-независимой версии программной системы, реализующей ДСМ-метод автоматического порождения гипотез / М. И. Забейло [и др.] // Там же. – С. 91–93 ; *Столяров, О. В.* Классификация ДСМ-методом диссертаций по информатике / О. В. Столяров, Р. С. Гиляревский, Н. И. Мельников // Там же. – С. 217–218.
5. *Almid, T. C.* Analyses on the World Wide Web: Methodological approaches to «Wedometrics» / T. C. Almid, P. Ingwersen // *J. Doc.* – 1997. – Vol. 53, N 4. – С. 400–420.

Материал поступил в редакцию 17.12.2007 г.

Сведения об авторе: *Гордукалова Галина Феофановна* – доктор педагогических наук, профессор