ФОРМИРОВАНИЕ ИНФОРМАЦИОННОГО ПРОСТРАНСТВА ДЛЯ ПРОВЕДЕНИЯ МОНИТОРИНГА

Макогон Н.В.*, Харьковский национальный университет городского хозяйства имени А. Н. Бекетова*

В данном докладе под мониторингом (не конкретизируя сферу его применения) будем понимать систематическое наблюдение за исследуемым объектом с целью анализа (идентификации) текущего состояния и прогнозирования поведения (возможно при заданных условиях, или как реакцию на некоторые возмущения) в некоторый будущий период времени. Мониторинг может использоваться для выявления критических состояний объекта, диапазонов устойчивости объекта к внешним возмущениям, факторов влияния, выбора сценария развития и т.д. Для реализации функций мониторинга необходимо создать систему фиксирования и анализа некоторого набора характеристик объекта, основная цель которой состоит в формировании информационного пространства мониторинга. Такое информационное пространство должно обеспечивать взаимодействие субъектов мониторинга, удовлетворяя их информационные потребности по выполнению перечисленных выше функций.

Часто под информационным пространством понимают совокупность банков и баз данных, полученных в результате функционирования различных информационных телекоммуникационных систем, тем самым недооценивая сложность и трудоемкость этого процесса. Как показывает практика, такая совокупность не может быть создана простым объединением, и формирование информационного пространства сопряжено с определенными трудностями. Вот некоторые из них.

Выбор релевантных источников

Подбирая источник данных, или формируя необходимый набор данных необходимо помнить о том, что ключевое требование к источнику: составляющие его данные должны удовлетворять основным показателям качества данных, таким как актуальность, достоверность, оперативность, стоимость, разрешение, релевантность и т.д. Только оценив эти показатели для каждого имеющего набора данных, который планируется включить в информационное пространство, необходимо осуществлять выбор источников. Таким образом, даже на этапе отбора источников данных, на самом первом этапе формирования информационного пространства, как правило, необходимо использовать методы многокритериальной оптимизации, иногда с использованием критериев нечеткой логики.

Несовместимость объединяемых источников

Простое объединение разрозненных банков данных зачастую невозможно, т.к. возникают проблемы несовместимости информационных систем, созданных различными компаниями. Причем такая несовместимость тем выше, чем масштабнее объект мониторинга. Формирование информационного пространства для проведения мониторинга на уровне компании обычно выполняется довольно просто, в то время как переход на региональный или государ-

ственный уровень требует разработки специальных стандартов для новых баз данных или конвертеров для существующих.

Неадекватная и/или негибкая модель данных

Информационное пространство, адекватно отображающее предметную область, редко может быть построено с первого раза. Проектирование информационной модели обычно требует нескольких итераций, – для проверки ее надежности, устойчивости и гибкости. Ошибки в построении модели могут привести к полной переработке проекта.

Замусоренные данные

Во многих случаях при формировании информационного пространства возникают данные, не стыкующиеся с остальными по одному или нескольким параметрам. Иногда это просто сведения об отклонении модели (и тогда их ценность огромна, поскольку они позволяют пересмотреть модель данных без масштабной корректировки проекта). Иногда это ошибочные данные, полученные вследствие неверного ввода или конвертации, или зашумленные показатели, которые встречаются в большинстве исследуемых объектов. Так или иначе, при формировании информационного пространства приходится прибегать к чистке или предварительной обработке данных.

Неполнота данных

Иногда накопление данных наблюдений за тем или иным объектом находится в самом начале, или данные об объекте фиксировались ранее (а бывает и продолжают фиксироваться) не в цифровом виде. Тогда исследователь сталкивается с проблемой недостаточного объема данных, характеризующих объект. Методы решения этой проблемы известны. Это оцифровка существующих данных (например, для пространственно распределенных объектов), идентификация параметров, обусловленных физическими законами (в инженерных сетях) или интерполирование непрерывных состояний (для объектов, у которых фиксировались только дискретные состояния). Выбор методов определяются конкретным объектом и характером данных о нем.

Только после решения всех перечисленных проблем по формированию информационного пространства можно переходить к анализу объекта, поиску закономерностей/зависимостей, идентификации состояний, прогнозированию поведения, визуализации/интерпретации результатов и принятию решений. Таким образом, формирование информационного пространства требует взвешенного подхода, что часто недооценивают при организации мониторинга.

Список литературы

1. Almind T.C., Ingwersen P. Informetric analyses on the World Wide Web: Methodological approaches to «webometrics»//J. Doc. 1997. Vol. 53 (4). P.404–426.

2. Ranking Web of World Research Centers. – http://research.webometrics.info.

3. Ханчук А.И., Наумова В.В. Информационное пространство Дальневосточного отделения РАН // Вестн. ДВО РАН. 2009. № 4. С. 122–129.