

O. E. Корнеенко

okorneenko@gsu.by

Гомельский государственный университет им. Ф. Скорины, г. Гомель, Беларусь

ИНФОРМАЦИОННЫЕ РАЗРЫВЫ В ИАС

Исследуются вопросы взаимодействия источников информации и ее аналитиков в информационных системах, рассматриваются возникающие при этом проблемы.

При проектировании, реализации и эксплуатации информационных систем решается задача создания и поддержания предметно-ориентированного, интегрированного и зависимого от времени набора данных. Вопрос выявления и преодоления информационных разрывов, возникающих в потоках данных, формирующих подобный набор, является одним из основных при проектировании информационных систем. Полученный набор данных необходим для поддержки анализа и принятия решений пользователями в условиях большого быстрорастущего объема информации.

При современном уровне развития информационных технологий простое перемещение данных из одной БД в другую не представляет затруднений. Но, если мы просто переместим данные из всех источников данных в одну БД, то получим всего лишь несогласованный набор данных. Чтобы создать нечто доступное для анализа конечным пользователем, необходимо согласовать поступающие данные, создать тот самый согласованный, предметно-ориентированный, интегрированный, зависимый от времени набор данных.

Изучаемой системой может быть, например, весь бизнес или некоторая его логически обособленная часть. Выделив в такой системе составные части, изучив их состояние, развитие и связи между ними, можно получить представление о состоянии и развитии системы в целом. Данные, необходимые для такого анализа, должны представлять максимально полный набор измеряемых показателей, отражающих состояние всех составных частей системы в любой момент времени.

Понятие – это элементарная единица мыслительной деятельности, обладающая известной целостностью и устойчивостью и взятая в отвлечении от словесного выражения этой деятельности. Поскольку мы мыслим понятиями, то, столкнувшись с каким-либо множеством объектов, мы разбиваем заинтересовавшее нас множество объектов на классы «эквивалентных» в каком-либо отношении. Таким образом, получается новое множество, элементы которого сходны по определенному нами отношению эквивалентности. Элементы этого нового множества можно считать единными объектами.

Понятию для обеспечения обмена присваивается имя – идентификатор. У понятия есть некий концепт – смысл, который это понятие выражает. Другими словами, это перечень свойств, на основании которых было определено данное понятие. Оно обладает неким денотатом, что обозначает объём понятия (совокупность элементов в конкретном речевом произведении). Чем обширнее набор свойств, тем уже класс объектов, удовлетворяющих этим свойствам, и наоборот, чем уже содержание понятия, тем шире его объём.

Очевидно, что работая с данными, желательно получать как можно более точный анализ и максимально полное представление о состоянии бизнеса: оперируя понятиями, подобными таким, как «карточка клиента» или «филиал», можно лучше проанализировать состояние объекта, если собрать больше параметров для анализа его состояния. Иными словами, чем больше содержание (набор свойств) понятия, тем более «обширный» концепт можно рассмотреть. При этом меньшее число объектов обозначается этим понятием, и охватывается меньший объем (денотат).

На рисунке 1 представим пример стандартного для любой информационной системы набора данных (совокупность структурированных данных, представленная в виде таблицы).

Дата	Расчетный счет	Сумма остатка в валюте счета
15.09.2017	BYSSIS381XXXXXXXXXXXXXX	12345,67

Рисунок 1 – Набор данных информационной системы

Имена понятий и языковые знаки образуют более сложные языковые конструкции – знаковые системы, языки. Очевидно, что показатель «Остаток по расчетному счету BYSSIS381XXXXXXXXXXXXXX на дату 15.09.2017» связан с понятиями «Остаток», «Расчетный счет BYSSIS381XXXXXXXXXXXXXX» и «Дата 15.09.2017», которые также связаны с другими понятиями и т.д.

Более того, каждый из трех элементов понятия входит в несколько различных систем связей с соответствующими элементами других понятий, то есть образует связи «многие–ко–многим». В совокупности такие связи образуют понятийную систему. Отношения между рассмотренными элементами понятия в виде ER–диаграммы можно представить на рисунке 2.

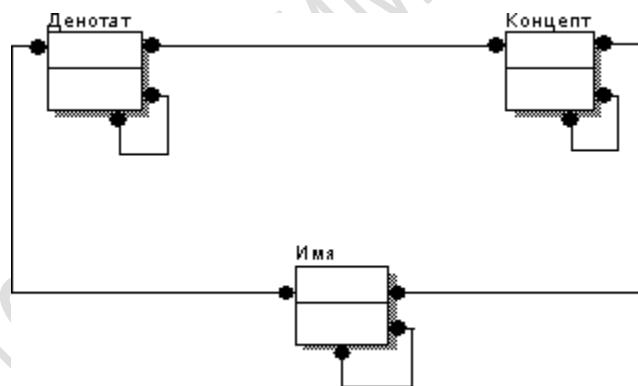


Рисунок 2 – Отношения между элементами понятия

Понятно также, что любая знаковая система (язык) не может описать весь мир полностью. Поэтому разработчики, создавая информационную систему для определенных задач, опираются на неполную систему понятий, которая, по их мнению, наилучшим образом подходит для решаемой задачи. Таким образом создается некий набор данных информационной системы.

При этом и имя понятия, и его концепт определяются достаточно произвольно в рамках различных систем понятий. Кроме того, в разных информационных системах одному и тому же денотату понятия могут быть присвоены разные имена, а одни и те же имена могут быть присвоены понятиям с различными концептами и, следовательно, иметь пересекающиеся (в сложных случаях и непересекающиеся) денотаты. Также понятия и каждый из их элементов могут иметь различные системы связей с другими понятиями. Такая ситуация называется семантическим информационным разрывом между системами.

Пространственно-временная точка, в которой производятся измерения и регистрация набора данных, называется точкой регистрации, а пространственно-временная точка, в которой производится анализ – точка анализа. Эти точки могут быть достаточно удалены друг от друга в пространстве или во времени, чтобы не иметь возможности взаимодействовать между собой напрямую помимо канала связи.

Ни один из видов анализа не применим к одному единственному измерению показателя. Поскольку, получив информацию из единственной точки регистрации и ничего не зная более, аналитик в точке анализа не может ничего расчленить, не может выявить общие свойства и связи и не может построить классификацию. Если регистраторов будет несколько, то вместе они регистрируют уже не один показатель, а столько, сколько их самих. Например, добавляются координаты точки регистрации (местоположение и время). В этом случае аналитик получает информацию по нескольким каналам связи. Получается некоторая система сбора информации для анализа. И теперь аналитик может, например, классифицировать поступающую информацию по одинаковым координатам, времени или имени показателя. Может попробовать выявить общие свойства и связи между показателями, то есть в различной степени применить все виды анализа.

Информационно-аналитические системы (ИАС) предназначены для сбора и анализа всевозможных показателей. И в любой ИАС подобную классификацию разработал человек, аналитик, участвовавший в создании системы, который сумел превратить процесс анализа в алгоритм. Данному алгоритму теперь и следует система. Но есть и другая часть работы системы – предоставление информации пользователю для дальнейшего анализа. Если считать, что БД ИАС это некое пространство, в котором существуют свои координаты и протекают во времени свои процессы, то точка регистрации – это поле в таблице БД. Регистратор – это тот запрос, который пользователь применяет к БД. Пересылаемые по каналу связи показатели с именами – это результирующий набор данных, формируемый запросом и, наконец, аналитик – сам пользователь. Если учесть, что многие системы сами выступают в роли пользователя по отношению к другим системам, то может выстраиваться множество цепочек из любого количества звеньев, полностью описывающее продвижение информации от самого первого замера показателя к конечному пользователю.

Информационный разрыв может возникать там, где существуют связи «многие ко многим» между элементами понятия или понятиями в системе понятий. Например, один регистратор передает информацию о сумме остатка на счете, а другой – о зачисленной сумме. Если один передает «Сумма – 30», а второй «Сумма – 30», то аналитик столкнется с семантическим информационным разрывом: имена понятий одинаковые, а их концепты и, следовательно, наполнение – денотат – разные.

Канал связи также может влиять на передаваемую информацию, так как передает данные от регистратора к аналитику: отфильтровывать часть данных, вызвать задержку в их передаче. Поскольку каждое передаваемое аналитику понятие на самом деле является еще и набором данных, то задержка в передаче части системы или потеря таковой приведут к разрыву, так как теряются связи. Например, передается понятие «*Остаток по счету № BYSSIS381XXXXXXXXXXXXXX* на дату 15.09.2017». Если ранее или одновременно с этим понятием к аналитику не поступило понятие «*Номер счета BYSSIS381XXXXXXXXXXXXXX*», то имя понятия просто не может быть интерпретировано. Это «асинхронный» разрыв.

Информационный разрыв – это синоним непонимания информации, которая заключена в данных. Непонимание может быть полным, частичным, разным у разных пользователей или, например, у пользователей и «производителей» данных. Информационная система содержит большие объемы разнородных данных и, обычно, большое количество «разнородных» потребителей.

Для интенсификации общения специалистов, повышения эффективности анализа и сокращения сроков принятия решений, ИС обязательно должна обладать средствами

исключения информационных разрывов или минимизации времени на их преодоление. Это может быть весьма полезно в процессе анализа решений и проектирования подсистем управления метаданными.

Литература

1. Корпоративный менеджмент [Электронный ресурс] / Интернет–проект «Корпоративный менеджмент». – Москва, 2017. – Режим доступа: <http://www.cfin.ru/> – Режим доступа: 03.09.2017.