

База данных числовой информации

Несмотря на большое количество разработанных электронных библиотек, доступ к числовой информации остается ограниченным. Получается каламбур, цифра до сих пор не оцифрована.

Разные элементы числовой информации обычно плохо систематизированы, не упорядочены по степени важности, труднодоступны. Существуют проблемы с сопровождением и обновлением информации и построением запросов для извлечения данных. Безусловно эти проблемы решаются, но за счет больших затрат времени, иногда приводящих к прекращению деятельности или закрытию отдельного направления деятельности.

К наиболее часто встречающимся проблемам можно отнести следующие:

- Числовую информацию (временные ряды, отдельные показатели, и т.д.) часто трудно найти в силу отсутствия системы их централизованного хранения и сопровождения. Во многих организациях встречается отсутствие такой системы или просто наблюдается, беспорядок (бардак) в вопросах сопровождения числовой информации, часто информационное поле просто отсутствует.
- Обычно итоговая числовая информация находится на различных компьютерах и подвержена риску потери в случае вирусной атаки или механического повреждения диска, а также увольнения сотрудника.
- На поиск числовой информации тратится много рабочего времени, часто поиск не дает результатов или приводит к ошибочным результатам, сопровождаемым эмоциональным и нервным напряжением.
- Распространено мнение, что проблема решается созданием информационного хранилища, поскольку в нем можно найти все. Это не всегда так, поскольку в информационном хранилище трудно выделить уровни иерархии (важности) информации, а возможности поисковой системы ограничены.
- Не всегда удается найти нужную информацию из-за наличия различных версий документов на нескольких компьютерах.

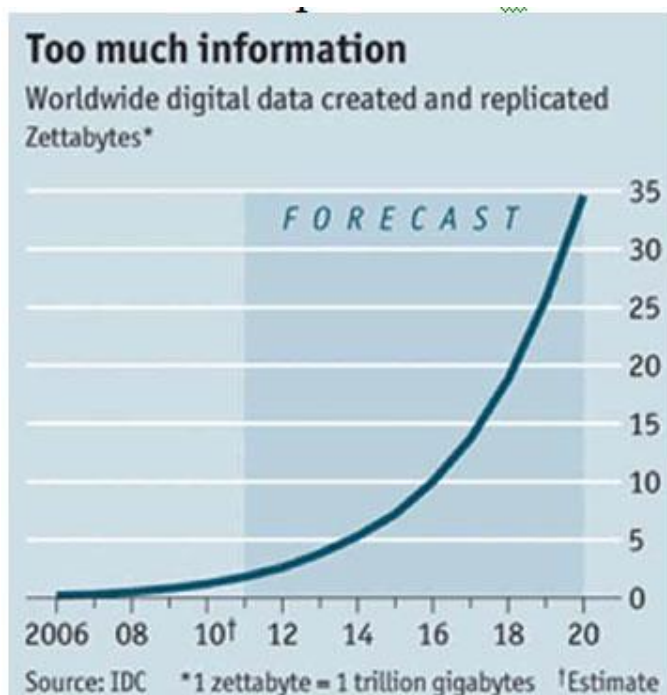
Мы проделываем много работы в поиске и анализе различных источников данных, потом где-то сохраняем результат и снова возвращаемся к поиску уже среди своих разрозненных файлов, вспоминая где мы уже могли сохранить требуемую информацию.

Переход на сложные платформы типа SAS, Oracle т.д. не решают проблему, поскольку зачастую масштабы отдельной задачи существенно меньше минимальных возможностей большой информационной системы, которые сами по себе сложны, кроме того ограничением является стоимость самого программного обеспечения и лицензии.

Но, главным вызовом является стремительный рост информации. С каждым годом количество окружающих нас данных увеличивается, время на поиск необходимой информации растёт и складывается ощущение, что мы ходим по кругу и утопаем в потоке информации.

Статистика говорит, что 90% доступной с сети Интернет информации было размещено в последние 2-3 года, прогнозы подтверждают кратный рост информации (рис 1.)

Рис. 1 Прогноз объема производимой информации



В условиях информационного цунами можно потерять не только время, но и бизнес. Ярким примером служит ситуация с разорением компании Yahoo.

Для сохранения ориентации в быстрорастущем информационном потоке целесообразно сортировать, сегментировать информацию по степени значимости и разрабатывать технологии по ее хранению, сопровождению и анализу. Проще говоря, чтобы навести порядок в данных нужно структурировать имеющуюся информацию и иметь возможность её быстро и легко просмотреть.

Авторами была разработана архитектура базы данных числовой информации (DataViz), позволяющая пользователю самостоятельно, без помощи программиста, работать с числовыми рядами.

К основным свойствам можно отнести следующие:

- DataViz позволяет осуществлять сбор, централизацию и систематизацию числовой информации, упрощает к ней доступ, и обеспечивает функции автоматического резервного копирования, анализа и построения структурированного архива.
- DataViz представляет, как информационную оболочку в формате HTML, так и в формате Excel, позволяющую дальнейшее развитие и дополнение, упрощающую доступ к числовой информации, облегчающую ее классификацию и поиск.
- DataViz дает возможность пользователю самостоятельно, без помощи программиста, сохранять и систематизировать числовую информацию, а также обеспечить к ней доступ как в локальной среде, так и в среде Интернет

DataViz - это простой веб-каталог, который структурирует как личные, так и корпоративные числовые данные, позволяет разложить по категориям и имеет удобный интерфейс просмотра информации сразу в графическом и табличном представлении, а также предоставляет возможность делиться опубликованными данными по ссылке при наличии доступа к Интернет.

DataViz предлагает унифицированный формат для хранения числовых данных в текстовых файлах, которые удобно использовать также и в любом текстовом или табличном редакторе, а не только в DataViz.

DataViz не требует инсталляции и легко запускается на любом компьютере, достаточно скопировать каталог с её содержимым и обладает следующими важными свойствами:

- Доступность - веб-каталог доступен с любого компьютера, у которого есть доступ к сети интернет и можно делиться ссылками на выбранные ряды данных.
- Безопасность - каталог доступен всем на чтение, только владелец может вносить исправления.

Использование DataViz позволит навести порядок в числовой информации, даст возможность управлять ей, что существенно сэкономит силы и сократит время обработки и сопровождение новой информации.

Пример реализации DataViz можно посмотреть на следующем сайте:

<http://www.infoarchives.ru/dataviz/>