

Информационные технологии и управление данными

Базы знаний, основанные на стандартах, как рабочая среда в нефтегазовой отрасли Standards-based knowledge environments for the oil and gas industry

Труди Кертис (Trudy Curtis), руководитель Ассоциации по открытым моделям данных в нефтяной отрасли (Public Petroleum Data Model Association, PPDM) рассказывает об основах ассоциации и ее работе.

Xорошее планирование и осуществление управления данными, основанные на стандартах, позволяют добиться того, что такое управление будет нацелено в будущее. Ассоциация по открытым моделям данных в нефтяной отрасли (ОМДН) уверена, что интегрированное хранилище

знаний, информации и данных, созданное по международным стандартам, позволит нефтегазовым компаниям достичь своих целей в бизнесе.

Введение

Биологическими видовым признаком человека является желание понять, описать и классифицировать все на свете. Тысячелетиями мы, с большими или меньшим успехом, собирали, упорядочивали и складировали информацию. Может показаться, нам хочется описать и классифицировать те аспекты окружающего мира, которые связаны с нашими интересами. Две тысячи лет назад римляне, собираясь покорять мир, сосредоточились на военном деле и средствах связи. У

Сегодня добывающие отрасли, нефтегазовая в особенности, действуют в обстановке, когда объекты поиска сложны как для понимания, так и для разработки, а правовая база становится все более жесткой. Как итог, возникла потребность упорядочивать и использовать обширные объемы Знаний, Информации и Данных (ЗИД) в обстановке интенсивных внешних процессов, для которой характерны несколько важнейших признаков.

Области разделения труда

Для успеха в конкурентной борьбе компании используют разделение труда, так, чтобы работники с определенными навыками могли выполнять свои обязанности наиболее эффективно. Например, буровой отдел занимается проходкой и обслуживанием скважин, а специалисты по управлению собственностью работают над приобретением или ликвидацией активов. К сожалению, побочным эффектом разделения труда, особенно в крупных организациях, часто оказывается изоляция и фрагментация знаний.

Развитие условий труда

В нынешних условиях с сильной конкуренцией и быстро меняющимися условиями труда предприниматели должны быстро и точно реагировать на изменения на рынке, в законодательстве, в технике и в общественном сознании. Акционеры ждут прибылей, партнеры требуют эффективных действий, власти ужесточают правила – все это заставляет искать в наших данных ответы на вопросы, о которых прежде и не думали.

Долгий жизненный цикл скважины

Жизненный цикл скважины долг – 30-50 лет и более. Информация, полученная в начале цикла, вдруг становится жизненно важно в самом его конце. Часто, однако, она уже утрачена или недоступна. К тому же оказывается, что электронные средства хранения не так долговечны, как нужно бы. Результаты технического проекта будет невозможно использовать через несколько лет, когда нужные инструменты будут недоступны, или через десять лет, когда они исчезнут совсем.

Важность стандартов

Нефтегазовые компании лидируют в выработке стандартов и групп стандартов с целью снижения влияния названных факторов на свою работу. Во многих отраслях требования стандартов являются в настоящее время стимулом к развитию. А всего 15 лет назад компании, если и пользовались стандартами, то часто создавали собственные внутренние стандарты. Техника развивалась быстрее, чем политика компаний в области снабжения работников средствами производства, что часто приводило к росту затрат и труднопреодолимым затруднениям из-за несовместимости оборудования, форматов данных, программ и методик.

В 1990-е годы в компаниях появилось понимание того, что затраты труда и капитала в соответствии с внутренними стандартами не дают результата с точки зрения бизнеса. Ценность самих стандартов, однако, по-прежнему признавалась. Результатом были попытки

Об Ассоциации PPDM

Ассоциация PPDM – некоммерческая организация, разрабатывающая и поддерживающая стандарты в энергетической отрасли. Ассоциация, в которую входит более 100 нефтяных компаний и поставщиков данных, программ и услуг, организует круглые столы, где эксперты совместно могут выработать полезные для бизнеса стандарты.

В соответствии с законом PPDM не имеет права получать прибыль; финансирование, предоставляемое PPDM ее участниками и отраслью в целом, полностью идет на создание, развитие, публикацию и поддержку стандартов. PPDM финансируется за счет членских взносов и пожертвований предприятий отрасли. Участники PPDM, использующие ее стандарты, говорят, что многократно окупили свои членские взносы,

Информационные технологии и управление данными

полной стандартизации всех аспектов технологической инфраструктуры компании, от компьютеров на рабочих местах, сетей и операционных систем до управления данными и пакетов программ. Эти проекты обошлись в миллионы, но и сэкономить позволили не меньше. Затем многие организации решили вложить свои средства и опыт в разработку общих стандартов в сфере бизнеса и технологий; в результате в 1989 г. появилась Ассоциация по открытым моделям данных в нефтяной отрасли. С тех пор Ассоциация совместно с компаниями разрабатывает стандарты управления данными в энергетическом секторе.

В 2004 г. компания Gartner Group начинает исследования с целью пересмотра стандартов в нефтегазовой отрасли. «Высшим приоритетом для этого исследования, по мнению наших ведущих клиентов, является помочь в вопросах интеграции данных, в том числе – создание стандартов. До конца 2004 г. в тематику этого исследования нефтегазовой отрасли будут включены принятые в энергетике стандарты, влияющие на интеграцию данных. Компания Gartner считает, что отрасли с сильными и обширными стандартами имеют хорошие экономические результаты независимо от общекономических условий. Для нефтегазовой отрасли сейчас самое время заняться этим вопросом, и клиенты Gartner хотят знать, как на них повлияют эти отраслевые

Группа Cambridge Research Energy Associates (CERA) в работе «Цифровая модель месторождения будущего: исследование для многих» (*The Digital Oil Field of the Future; a Multi-client Study*) также подчеркивала важность стандартов: «Исследование, проведенное CERA показало, что, хотя использование Цифровой модели месторождения будущего (ЦММБ, DOFF) позволило бы дополнительно извлекать до 125 млн баррелей,... отсутствие единых стандартов сбора и распределения данных может замедлить разработку и внедрение ЦММБ, а также увеличить затраты на эти цели.»²

Значение рабочей среды, основанной на знаниях

При имеющихся в нефтегазовой отрасли больших объемах сложных данных очень важно обращаться с ними так, чтобы они были доступны для решения будущих задач, в том числе таких, возникновение которых сейчас трудно предвидеть. Новые возможности или скрытая в данных ценность часто раскрываются лишь спустя годы после проведения съемки.

Хорошее планирование и осуществление работы с данными на основе стандартов позволяют добиться того, что такая работа будет нацелена в будущее. Ассоциация PPDM уверена, что интегрированное хранение знаний, информации и данных, созданное по международным, совместно разработанным, стандартам, позволит нефтегазовым компаниям достичь своих целей в бизнесе лучше, чем хранение по внутренним стандартам или стандартам разработчика базы данных. Рабочая среда, основанная на знаниях, различает информацию, нужную и полезную на текущем этапе (например,

полевые данные), и информацию, которая должна храниться и быть доступной многие годы и десятилетия. Понимание рабочих процессов, в которых создается и используется информация, позволит нефтегазовым компаниям планировать работу с данными от колыбели до могилы. Хранение важных данных с длительным периодом использования в форматах, не зависящих от производителя (XML, TIFF, база данных PPDM) обеспечивает постоянный доступ к ним.

Правильный порядок движения информации между пользователями, отделами, программами и потребителями обеспечивает создание копий важных данных в соответствии с верными процессами управления, и сохранность всей совокупности данных в течение всего требуемого срока хранения.

Характеристики рабочей среды, основанной на знаниях

Хорошая рабочая среда, основанная на знаниях имеет много параметров и составляющих. Сила такой среды – в ее полноте, уровне интеграции рабочих мест и доступности конечному пользователю всех ключевых элементов информации в нужный момент на любом этапе. Большинство составляющих среды лучше разрабатывать совместными усилиями управленцев, пользователей и специалистов по информатике, причем все они должны иметь ввиду интересы бизнеса и долгосрочные стратегические потребности.



Политика, практика, процедуры (ППП)

Соглашения о ППП устанавливают цели организации и явно формулируют стратегию и методы проведения корпоративной политики. Компании следуют иметь такое соглашение по вопросам управления данными. Так устанавливаются основы и средства проверки всей системы. Ассоциация PPDM совместно с организациями отрасли вырабатывает соглашения о ППП в управлении данными для энергетического сектора.

Правила внутреннего распорядка

Правила внутреннего распорядка, воплощенные в определенном наборе вычислительных средств, обеспечивают постоянное и верное исполнение соглашения о ППП. Например, внутренние механизмы безопасности систем управления базами данных (например, Oracle) или операционных систем (например, Windows, LINUX или UNIX) можно использовать для ограничения доступа лиц, на то не уполномоченных, к определенным видам данных. Можно создать проверки и

Информационные технологии и управление данными

встроенные процедуры, позволяющие осуществлять допуск автоматически. Другие правила позволяют следить за теми, кто имеет право использовать, создавать, изменять или удалять информацию в базе данных.

Рабочая группа Ассоциации PPDM разрабатывает такие правила внутреннего распорядка для энергетики.

Среда Знаний, Информации и Данных (ЗИД)

Каждая частичка знаний, информации или данных, важная для организации в течение длительного времени, относится к среде ЗИД, независимо от формата, формы представления или назначения. Для этой среды следует разработать и применять стратегию дублирования, восстановления и хранения. Сейчас слишком многие компании хранят слишком много важных данных в документах Excel или в иных неструктурированных формах, которыми невозможно пользоваться в среде ЗИД. Рабочие группы Ассоциации PPDM создали богатую, практическую и полезную реляционную модель данных, включающую 45 предметных отраслей. Этой моделью широко пользуются в нефтегазовой отрасли для управления данными, входящими в среду ЗИД.

Не нужно, не удобно и не желательно хранить все данные и информацию в единой монолитной системе. С точки зрения среды ЗИД для организации гораздо важнее просто знать, какие данные доступны, где они хранятся, и кто их изначально предоставил. Допущение хранения крупных наборов данных (таких как скважины) в многопользовательских системах и рабочих группах ведет к неоднозначности и дорогостоящему и сложному перекрытию. Система управляемого воспроизведения позволит избежать проблем с несинхронным хранением данных и сохранит полноту и полезность данных.

Работа PPDM

Ассоциация разрабатывает стандарты совместно с предприятиями отрасли для учета потребностей современного бизнеса. Создание рабочих групп и проектов инициируется участниками PPDM в ответ на специфические проблемы, с которыми они сталкиваются или могут столкнуться. Участие предприятий и экспертов со всего мира гарантирует, что продукты PPDM практичны и полезны при использовании в рабочей обстановке. Служба поддержки, которая помогает предприятиям отрасли понимать и использовать стандарты PPDM, дает группам по внедрению ценные советы, которые экономят время. Стандарты PPDM публикуются на ее сайте. Многие

Рабочая среда

Каждый раз, когда пользователи общаются между собой или используют программное обеспечение, происходит создание или использование ЗИД. Некоторые из этих ЗИД будут цепны длительное время, и, значит, относятся к среде ЗИД. Другие данные являются промежуточными или временными. Определить, относится информация к текущей или к долгосрочной, может быть затруднительно. Конечные пользователи, создающие или использующие данные в начале их жизненного цикла, не всегда знают, будут ли эти данные ценностями долгое время; в результате важные данные

в формате производителя или в неструктурированной форме документов Word или электронных таблиц, на бумаге или в ином неудобном для использования виде, а то и уничтожены.

Решения по стратегии управления данными того или иного типа являются в большой степени стратегическими, чем часто пренебрегают под воздействием ради текущих технических или деловых факторов. Наилучший порядок принятия решений включает все стороны вопроса, включая пользователей на всех этапах жизненного цикла данных, технических специалистов, стратегических и бизнес-аналитиков. Процесс принятия решения должен быть основан на корпоративной политике и практике; они, в свою очередь, приведут к выработке приемлемых процедур и правил. Например, корпоративная политика, требующая данных отличного качества, повлияет на процедуры и правила работы с полевыми материалами.

Среда обмена данными

В среде ЗИД любой нефтегазовой компании данные должны перемещаться между пользователями, системами, устройствами хранения, государственными органами и др. Многие компании тратят крупные суммы на разработку и поддержку разнообразных сложных наборов механизмов и методов. Пониманию и эффективному управлению этими процессами нужно долго учиться.

В результате появляется группа узких специалистов, ответственных за работу с важными данными в дорогой нестандартизированной среде. Открытые форматы данных (например, семейства XML, разработанного компанией W3C) дают гибкие и мощные стандартные средства обмена. Ассоциация PPDM приняла формат W3C XML за основу ряда схем, отвечающих нуждам отрасли, для обмена информацией по нефти и газу.

Планирование рабочей среды, основанной на знаниях

Построение и использование полезной среды ЗИД требует времени, желания и средств. Будьте реалистами; вряд ли удастся достичь всех целей сразу в рамках одного проекта. Какое-то время уйдет на анализ текущей ситуации и принятие решения, что делать. Стратегия успеха состоит в поэтапном развитии проекта, который на каждом шаге будет все более полезен для конечных пользователей. Проекты, которыми хорошо управляют в рамках общей системы, имеют больше шансов на успех, чем те, где решения принимаются каждый раз, на основе лишь общего

С самого начала проекта следует привести концепцию управления в соответствие с идеей интегрированной среды ЗИД и с потребностями конечных пользователей и технических служб. Нужно также уяснить, в рамках каких ППП работаете вы и другие сотрудники. Финансируовать будут скорее тот проект, который основан на целях компании и в котором есть потребность.

Разработайте целевую среду ЗИД, которая опознает, защищает важные, не промежуточные, данные и входит в общую среду ЗИД компании. В общей среде может быть много типов носителей, много баз данных и систем, но копирование данных должно быть сведено к минимуму. Обеспечьте доступ к данным и возможность их повторного использования.

Информационные технологии и управление данными

Интеграция правил разработки и правил внутреннего распорядка

Создайте такие правила внутреннего распорядка, которые позволяют отслеживать, как создаются и поддерживаются данные в среде ЗИД. Такие правила, включенные в систему управления базой данных, будут более надежны и устойчивы, чем просто записанные в книге, которая лежит где-то на полке. В ряде случаев может понадобиться продемонстрировать, что применение правил позволяет соблюсти важные требования проверяющих органов.

Например, в базах данных (такие как PPDM 3.7), где хранится информация по добыче и запасам, часто требуется обеспечить возможность добавления в базу новых данных, но исключить возможность изменения данных после того, как они внесены в базу. Изменения в объемах вводятся как поправки, с сохранением всей истории обновляемого набора данных и с указанием, когда и почему было принято решение о вводе поправок. Полезно показать проверяющим органам, что ввод данных по объемам имеют право осуществлять только лица, на то уполномоченные, а изменять или удалять данные, уже попавшие в базу не может никто; дата ввода каждой поправки ее обоснование при этом сохраняются.

Создание единообразной рабочей среды

При создании рабочей среды возникает много вопросов. Какие программы будут использоваться? Какие программы будут обращаться за данными в главное хранилище (и, если будут, то как они их получат?) или будут вносить новые данные в главное хранилище? Создаются ли промежуточные данные и какова их ценность? Какие действия необходимы для установления связи между программами? Осознайте требования к системе безопасности на каждом уровне системы, как в ходе текущей работы, так и в длительной перспективе.

Из всех прочих, рабочая среда наиболее подвержена неуправляемому копированию данных. Следует помнить, что каждый раз, когда в программе и в рамках проекта создается копия данных, понадобится создать также процедуру и правила для поддержания главного хранилища в актуальном состоянии. Слишком часто компании оставляют важную для всех информацию во внутреннем формате данных отдельных программ; в результате важные знания оказываются в форме, в которой их трудно и дорого совместить и использовать с другими, или, еще хуже, данные теряются или

Проверяйте все новые программы на соответствие общей конструкции среды ЗИД и остерегайтесь систем, которые не поддерживают простых и открытых механизмов интегрирования в общую систему. Для программ, дублирующих действия или данные, следует оценить возможность избегать дублирования или

Важность стратегического мышления

С каждой частицей данных или информации, имеющей долгосрочную ценность, следует обращаться так, чтобы в среде ЗИД она могла быть полезной длительное время.

Следует учитывать все аспекты ваших долговременных интересов, ставя ряд продуманных стратегических вопросов. Знаете ли вы всех возможных пользователей этих данных за весь их (данных) жизненный цикл, и для чего они будут эти данные использовать? Слишком часто учитывается лишь первый пользователь или создатель информации, а последующие пользователи вынуждены искать пути поиска или восстановления недостающих важных данных, или данных для них просто

Как ППП и правила внутреннего распорядка каждой организации могут быть интегрированы в общий стратегический план корпорации? За последние 10 лет корпорации децентрализовали многие процессы управления данными и их хранения, передав их конечным пользователям и рабочим группам. К несчастью, лишь немногие из них хорошо понимают требования, которые закон, контролирующие органы или корпорация предъявляют к принятию решений. Чем больше функций по хранению данным передается подразделениям или отдельным работникам, тем больше проблем возникает у корпорации с безопасностью данных, потерей данных и порчей носителей или

Как интегрировать эти данные с другой связанный с ними информацией? Смогут ли пользователи без помех перемещаться между связанными блоками информации и программами? Чем раньше создана стратегия интеграции, тем лучше удастся сложить вместе все куски в единую среду, основанную на знании. Есть много видов стратегий интеграции, применимых в каждом конкретном случае. Модель данных PPDM 3.7 поддерживает прямую интеграцию между структурированными, неструктурными и частично структурированными средами. Другие механизмы, такие как классификация (таксономия), описание связей (онтология) и пакеты средств интеграции дают другие возможности интеграции.

Хранятся ли данные на носителях, которые можно использовать в течение всего жизненного цикла данных? Бумага может сгнить, магнитную ленту зажмет, компакт-диск не архивного качества через несколько лет просто не будет читаться, а у винчестера ограниченный срок эксплуатации. Хранятся ли данные в аппаратно и программно независимых форматах, не привязанных к конкретным операционным системам, программам, или внутренним форматам? Документ PDF, задуманный как стандарт ISO, через 10 лет можно будет прочитать с большей вероятностью, чем документ конкретного текстового редактора. Популярный формат TIFF (Tagged Image File Format), имеет как минимум шесть вариантов, из которых многие уже не используются или скоро выйдут из употребления.

Хранятся ли базы данных, включая базы по проектам, в нейтральном формате, который будет доступен в будущем? Многие форматы данных в науках о земле увязаны с конкретными программами, что приводит к проблемам с потерей данных и при интеграции данных из разных источников.

Информационные технологии и управление данными

Список вопросов будет, конечно, расти по экспоненте; помнить следует, что вопросы эти по сути своей в основном стратегические, а не текущие. К решению этих вопросов с самого начала следует привлекать людей со стратегическим пониманием целей корпорации и планов управления.

Построение среды, основанной на знаниях

Продвигаясь к цели, проверяйте каждый новый компонент на совместимость с целым. Нужно понимать, какие данные использует или создает компонент, и где они хранятся. Следует четко зафиксировать правила внутреннего распорядка, которые устанавливают, кто на что имеет полномочия, какому пользователю или программе принадлежат данные определенного типа, как идет копирование данных или обмен ими, и как данные проверяются.

Компоненты, дублирующие функции или данные, следует оценить с точки зрения возможности избежать дублирования или управлять им. Беспечного копирования следует всегда избегать; неуправляемое копирование дорого стоит и рискованно. Управляемое копирование по автоматизированной технологии и четким правилам может стать эффективным и надежным методом совместного использования проверенной, стандартизованной и точной информации на различных ресурсах.

Выводы

Избежать риска в ходе создания интегрированной среды ЗИД не всегда просто. Поставщики и исполнители могут давать неосуществимые обещания или предлагать решения, которые не выдержат испытания временем. Взвесьте каждую идею, руководствуясь здравым смыслом, и убедитесь, что ваша среда ЗИД будет надежной, удобной и доступной через десятки лет.

Несколько мыслей, с чего начать

1. Подумайте, кому в принципе может понадобиться ваша информация или знания за все время их существования. Это могут быть государственные

Стандарты PPDM

Ассоциация PPDM работает вместе с предприятиями отрасли над стандартами по четырем главным направлениям. Информация о продуктах и услугах PPDM размещена на сайте <http://www.pppdm.org/products/index.html>.

Модель данных PPDM

Реляционная модель PPDM, версия 3.7, основана на начальном уровне (entry level) стандарта SQL 92. Она может быть реализована в любой СУБД, полностью поддерживающей SQL*92 (Oracle, SQL*Server, POSTGRE др.). 1238 таблиц новейшей версии PPDM покрывают 45 предметных областей. Каждая предметная область разработана техническими экспертами и экспертами по бизнесу многих стран в энергетической отрасли. В открытом доступе имеются схемы и документация высокого уровня. Участники Ассоциации имеют доступ ко всем деталям.

Семейство схем XML PPDM

органы, партнеры, финансовые организации, суды и общественность. Все ли группы будут удовлетворены этим решением? Какие дополнительные усилия или затраты могут потребоваться?

2. Какой правовой режим действует сейчас и как он может измениться в будущем? В Австралии уже наметилась тенденция к ужесточению требований государственных органов к управлению данными. (<http://frli.law.gov.au/>).
3. Прежде, чем отбросить данные, рассмотрите их текущую ценность и потенциальную ценность в будущем. Такие решения нужно принимать на стратегическом уровне, а не в ходе текущего управления; данные, не нужные специалистам в одной области, через несколько лет могут быть жизненно необходимы в другой.
4. Храните как можно больше деловых и технических данных в открытых, а не во внутренних форматах. Такие организации, как ANSI, ISO и W3C создают и распространяют технические стандарты, которые вам помогут. Такие организации по стандартам, как PPDM, POSC, OGC или OASIS разрабатывают стандарты для бизнеса в открытых форматах.

Литература

1. Email between Arthur Boykiw (chairman of the PPDM Association Board of Directors) and Dan Miklovic (Gartner Group), April 2004.
2. CERA, *Digital Oil Field of the Future Multi-client Study*. www.cera.com, 13 December, 2002.
3. PPDM Association, 'The PPDM Way', www.pppdm.org/whypppdm/ppdmway.html.

XSD от апреля 2001 г. Эти схемы покрывают широкий диапазон тем, связанных с энергетической отраслью. Схемы, примеры и техническая документация имеются в открытом доступе на сайте PPDM.

Содержание данных и контрольные цифры

Стандартные контрольные цифры упрощают совместное использование данных различными программами и организациями. Ассоциация PPDM совместно с предприятиями отрасли работает над формированием и созданием полезных набор контрольных цифр, которые могут быть цепны в любой конкретной работе.

Большинство таких списков имеются в открытом доступе на сайте PPDM.

Распределенные базы данных

PPDM разработала и опубликовала на своем сайте решения для работы с распределенными базами данных сложной структуры, в том числе – с PPDM.

Пользователи, которым нужна упрощенная версия для эффективной работы с геоинформационными системами работы, могут применять PPDM Lite 1.0 (также есть на