

# СПРАВОЧНО- ИНФОРМАЦИОННАЯ БАЗА ДАННЫХ СТАНДАРТНЫХ ЭЛЕМЕНТОВ, ИНСТРУМЕНТА И МАТЕРИАЛОВ



**О**громный объем разнообразной информации по стандартным элементам (на основе ГОСТ, ОСТ, СТП), инструментам, материалам и т.д. попадает в компьютерные системы из самых разных источников. Это могут быть сканированные материалы, информация, введенная оператором с клавиатуры или порожденная в автоматическом либо полуавтоматическом режиме из различных САПР, АСТПП и подобных систем с использованием различных баз данных. В конечном итоге вся эта информация должна использоваться в системах управления предприятием (ERP-системах), системах логистики (заказа и закупки), управления жизненным циклом изделия. Очень важно, чтобы к этому моменту она была однозначно идентифицируема с точки зрения нормативных документов.

К сожалению, это не всегда осуществимо, поскольку нормативные документы оставляют возможность записывать наименование изделия в весьма широком диапазоне, а для идентификации этих изделий в об-

щем случае ничего, кроме наименования, использовать нельзя.

Мы предлагаем подход, основанный на разборе и распознавании наименования, что позволит значительно уменьшить трудоемкость передачи между системами любой информации, содержащей ссылки на стандартизированные изделия.

## Типичные примеры, демонстрирующие полезность идентификации стандартных элементов

### Интеграция САПР и АСТПП

Предположим, что изделие, разрабатываемое в САПР, содержит

**МЫ ПРЕДЛАГАЕМ ПОДХОД, ОСНОВАННЫЙ НА РАЗБОРЕ И РАСПОЗНАВАНИИ НАИМЕНОВАНИЙ, ЧТО ПОЗВОЛИТ ЗНАЧИТЕЛЬНО УМЕНЬШИТЬ ТРУДОЕМКОСТЬ ПЕРЕДАЧИ МЕЖДУ СИСТЕМАМИ ЛЮБОЙ ИНФОРМАЦИИ, СОДЕРЖАЩЕЙ ССЫЛКИ НА СТАНДАРТИЗОВАННЫЕ ИЗДЕЛИЯ.**

стандартные элементы, для которых в САПР имеется библиотека 2D- или 3D-моделей. Далее результаты используются в автоматизированной системе технологической подготовки производства (АСТПП), имею-

щей свою базу данных по стандартным изделиям (как часть номенклатуры). При передаче информации необходимо, чтобы каждый стандартный элемент отображался на соответствующий элемент номенклатуры. Сейчас, как правило, используется подход, при котором таблицы номенклатуры АСТПП в качестве дополнительных параметров содержат коды элементов в САПР и другую информацию, внутреннюю для конкретной системы. При необходимости использовать в модели стандартный элемент система автоматизированного проектирования должна обратиться к АСТПП, выполнить поиск по своему коду, добавить (если это нужно) элемент в номенклатуру АСТПП, получить и запомнить уже код АСТПП элемента, который будет использован при экспорте спецификации.

Такой подход не свободен по крайней мере от трех недостатков:

- каждая система должна "знать" о другой. Если используется не одна, а несколько САПР, сложность задачи возрастает на порядок;
- для обеспечения работы САПР нередко требуется, чтобы АСТПП работала на том же компьютере, а это уменьшает надежность системы в целом;

- для взаимодействия необходимы либо мощный API с обеих сторон, либо специально проделанные "дырки" в обеих программах.

Взамен существующего метода предлагается следующий механизм взаимодействия. Каждая из систем работает автономно. В модуль экспорта спецификаций и/или ведомостей добавляется код, который нормализует наименование стандартного элемента в импорте и ищет элемент номенклатуры, нормализованное наименование которого соответствует данному. Чтобы ускорить импорт, рекомендуется (хотя это и не обязательно) нормализовать наименования номенклатуры в АСТПП. Для реализации этого подхода потребуется доработать только переносимой модуль.

### Интеграция ERP-системы и АСТПП

Предположим, что технологические процессы, использующие стандартные элементы, далее экспортируются в систему управления предприятием. Чтобы технологическая информация была привязана к управленческой, она должна содержать ряд управленческих параметров, причем большая их часть приходится именно на стандартные изделия (например, группа изделия, единицы измерения и схемы единиц измерения). Связано это с тем, что планирования закупок требуют именно стандартные изделия. Чтобы решить проблему, ERP-система экспортирует в АСТПП часть своих таблиц, содержащих возможные значения управляющих параметров. Технолог, занятый разработкой технологической подготовки производства (ТПП), должен устанавливать значения управленческих параметров, выбирая их из предлагаемых списков. Непосредственно к ТПП это никакого отношения не имеет, а правильность ввода не контролируется. АСТПП должна иметь возможность импортировать описания параметров — включая список значений, которые они могут принимать.

Предлагаемый механизм позволяет отказаться от предварительного импорта таблиц ERP и трудоемкого заполнения соответствующих параметров в интерактивном режиме: необходимые параметры добавляются непосредственно в момент экспорта

данных в ERP-систему. Это вполне осуществимо, поскольку разработанное нами решение способно отождествить наименования элементов в обеих системах. Такой подход позволит и подстраховаться на случай изменения данных в ERP-системе в период между импортом данных в АСТПП и экспортом из АСТПП.

### Создание общей базы данных стандартных элементов из нескольких уже существующих

Такая ситуация может возникнуть при слиянии нескольких предприятий. Если задача объединения БД по таблицам и полям, имеющим одинаковый смысл (объединение по вертикали), относительно проста и решается при помощи стандартных средств СУБД, то объединение записей, обозначающих один и тот же объект, как правило, требует ручной обработки всего массива информации. Учитывая, что человек, выполняющий эту операцию, должен обладать определенной технической квалификацией, при большом объеме данных задача становится достаточно трудоемкой.

Предлагаемое решение — автоматический анализ наименований всех стандартных элементов на базе существующих нормативных документов — позволит отсеять ошибочные варианты записей. Во многих случаях при этом выдается диагностика, помогающая определить, в чем именно допущена ошибка. Если эта ошибка не слишком серьезна (использован неверный разделитель, буква другого регистра или символ с похожим написанием — например, "3" вместо "З") и существует возможность автоматически восстановить, что же имелось в виду, возможно автоматическое исправление ошибки. Кроме того, произойдет объединение записей, имеющих разное написание, но одинаковый смысл — таким образом исключается повторное вхождение одного и того же объекта. В дальнейшем исправленное наименование элемента может использоваться в качестве первичного ключа соответствующей таблицы.

### Создание БД стандартных элементов предприятия

Ранее база стандартных элементов на предприятиях просто отсутствовала: лучшее, на что можно бы-

ло рассчитывать, это бухгалтерская информация в "1С" или снабженческие файлы в Excel. Поскольку эти БД создавались с другими целями, они могут быть использованы только как вспомогательная информация для формирования ограниченного перечня. Итак, если необходимо сформировать базу данных номенклатуры при внедрении АСТПП или PDM-продукта, существующий список наименований нормируется, из него исключаются дубликаты и он заносится как объект в БД продукта, где требуется создать базу номенклатуры. При этом возможны дополнительные опции — например, можно исключить типоразмеры, не рекомендованные соответствующим нормативным документом. В принципе списки применяемости типоразмеров могут быть заданы и произвольно — в соответствии с практикой применения на конкретном предприятии. Затем к созданным объектам автоматически добавляются списки параметров, определяющие геометрические, физические, химические и другие свойства объекта. Источником этих свойств служит как нормативный документ, которым непосредственно определяется стандартный элемент, так и связанные документы (например, марка материала изделия определяет его химический состав и механические свойства). Введенные параметры используются внедряемым программным продуктом для технологических, прочностных и других расчетов.

### Состав и основные возможности предлагаемого продукта

Предлагаемый нами программный продукт можно разделить на три взаимодействующие части:

- база данных стандартных элементов;
- GUI-клиент ("толстый") и web-клиент ("тонкий"), предоставляющие пользователю доступ к функциональности БД. SQL-API и C-API для доступа из других программ;
- программное обеспечение для автоматического заполнения БД по текстовым образам нормативных документов.

База данных может поставляться и работать независимо от других

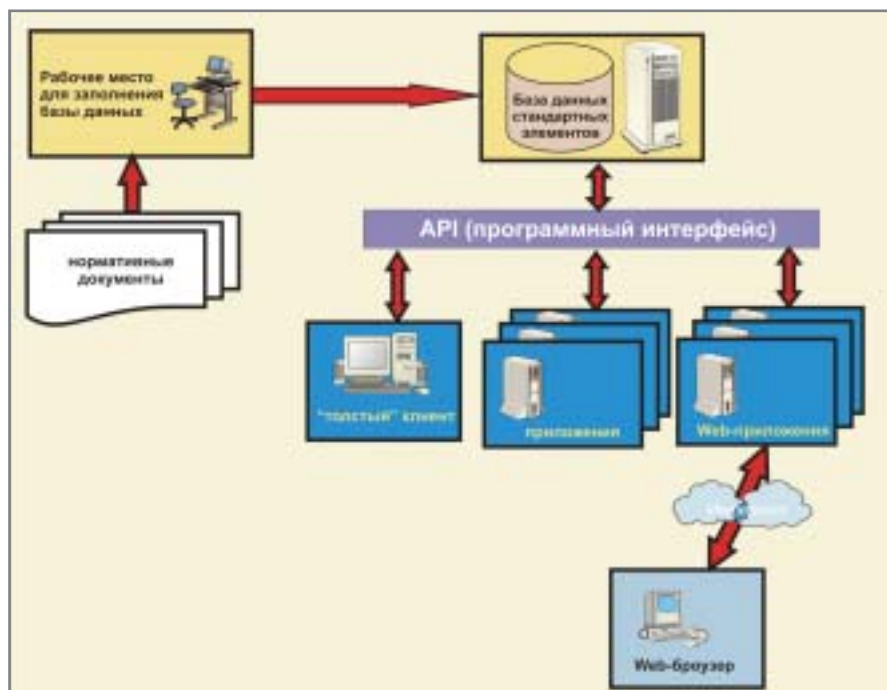


Рис. 1. Структура предлагаемого продукта

компонентов. Второй и третий компоненты требуют БД, но способны работать независимо друг от друга.

#### База данных стандартных элементов

Для стандартных элементов существует несколько типов информационных систем. Некоторые из них представляют собой справочные системы, рассчитанные только на пользовательский интерфейс. Основной формат хранения информации для подобных систем — это гипертекст, а их главный недостаток сводится к отсутствию программного интерфейса. Промежуточное положение занимают системы, использующие базы данных только как контейнер, где хранятся набор параметров, описывающих стандартные элементы. Формат представления параметров и связей при этом специфичен для каждой системы. Как правило, воспользоваться данными можно только с помощью API, предоставляемого системой: использование данных напрямую затруднено необходимостью формирования очень сложных запросов. К слабым сторонам таких систем следует отнести и то, что данные хранятся в ненормализованной форме — как следствие, возрастет пространство, необходимое для хранения, и замедлится доступ к данным. При разра-

ботке нашего продукта мы решили пойти по другому пути:

- максимально использовать функциональность, заложенную в СУБД, вместо дублирования этой функциональности в нашей программе. В частности, вся информация о значащих именах полей, их типах, связях и т.п. хранится штатным для базы данных способом;
- структуру хранения данных максимально приблизить к той, которая содержится в нормативной документации, описывающей стандартный документ. Колонки таблиц базы данных имеют те же имена, что и заголовки таблиц в нормативном документе, при этом исходные таблицы подвергаются ряду трансформаций (объединение, соединение и обратные к ним). Результат представляет собой набор таблиц, которые описывают возможные значения параметров элемента, а также зависимости между ними. Для каждого элемента эти таблицы объединены в схему, имя которой соответствует номеру нормативного документа, определяющего элемент.

Поскольку формат хранения данных достаточно прост, то и извлечь информацию можно при помощи

сравнительно простых запросов. В частности, все возможные комбинации типоразмеров и других параметров содержатся (но не хранятся!) в представлении (view), получаемом простым и естественным соединением (natural join) всех таблиц схемы.

Поверх указанной структуры мы предлагаем дополнительную функциональность, которая расширяет сферу применения данных, содержащихся в системе, и по сути превращает "пассивные" данные в "активные" объекты, обладающие интеллектуальными возможностями, обычно присущими лишь человеку. Эта функциональность позволяет:

- "понимать", что означает наименование элемента; извлекать из наименования значения, содержащиеся в нем и зависящих параметрах, а также проверять, являются ли извлеченные значения правильными. При этом спектр возможных вариантов записи наименований достаточно широк. Правильно понимаются символы, имеющие сходное написание (например, буква "А" русского и латинского алфавитов, прописная и строчная буква "Г", буква "З" и цифра "3" и т.д.). Допускаются сокращенные наименования — например, "Муфта" вместо "Муфта шарнирная". Между полями наименования возможно любое количество пробелов либо их отсутствие. Может использоваться тип разделителя, отличный от нормативного: ".6G" вместо "-6G". Мы привели лишь небольшую часть вариантов, воспринимаемых системой;
- вычислять, определяют ли два варианта названия один и тот же объект;
- определять по наименованиям, является ли один элемент частным случаем другого (к примеру, "Болт М4х6.019 ГОСТ 7805-70" — частный случай "БОЛТ М4Х6 гост 7805-70", но не "Болт М4-6gx6.019 ГОСТ 7805-70"). Возможность такого анализа важна, поскольку в различных приложениях требуется различная степень детализации одних и тех же объектов;
- создавать по выбранным параметрам наименование, полностью соответствующее нормативным документам.

Существенно, что все эти функции внешне реализованы в виде обычных функций языка SQL, а значит и сами они могут использоваться как органичная часть запроса. Так, например, чтобы проверить правильность наименования и, в случае успеха, извлечь значения связанных параметров, достаточно следующего простого выражения:

```
SELECT * FROM g7805.parse
('Болт М4х6.019 ГОСТ7805-70')
NATURAL JOIN g7805.view_main
```

### Разработанные приложения

Прежде всего нам хотелось создать простое приложение, которое, во-первых, предоставляет пользователю большую часть описанной выше функциональности и, во-вторых, без дополнительных трудозатрат стыкуется с любой другой Windows-программой. В качестве механизма взаимодействия был выбран стандартный текстовый буфер обмена, поэтому **Парсер наименований CSoft** может взаимодействовать с любой программой, поддерживающей команды "Копировать" и "Вставить".

В обычном состоянии парсер наименований неактивен и индициру-

ется в "system tray" наряду с другими системными приложениями, но при копировании текста в буфер обмена он автоматически активируется, разворачивая окно свойств объекта, — при этом цвет иконки в "system tray" меняется, показывая статус распознавания.

Ключевые параметры элемента можно выбрать из выпадающих списков (доступны только значения, допустимые по нормативным документам). Как результат в верхнем поле окна генерируется правильное наименование, которое через буфер обмена может быть передано в другое приложение. Таким образом, программа может использоваться и для генерации наименований. Кроме того, в парсер наименований встроены некоторые функции поиска.

Достаточно набрать в верхнем поле наименование элемента, чтобы программа выдала список нормативных документов, определяющих элементы с таким названием. Выбор строки этого списка вызывает переход в режим генерации наименования соответствующего элемента.

Если известен номер документа (например, номер ГОСТ), достаточно набрать этот номер в верхнем поле программы.

Парсер наименований удобен для интерактивной работы. Что же касается обработки больших объемов данных, то для нее предназначено другое приложение — **Анализатор/ва-**

**лидатор стандартных наименований.** Идея этого приложения состоит в следующем: получить информацию из произвольного ODBC — источника данных, проверить и, насколько это возможно, автоматически исправить содержимое после чего записать результат в тот же либо другой источник данных.

При работе приложения доступны следующие возможности:

- в случае неправильного наименования происходит анализ ошибки — с выдачей сообщения в отдельной колонке (рис. 5);
- если ошибок нет, генерируется нормализованный вариант наименования;
- пользователь может самостоятельно определять набор колонок, предназначенных для вывода, а также создавать собственные вычисляемые колонки при помощи скриптов на языках Perl, JScript и VBScript;
- скрипты можно подключать и на этапе, предшествующем обработке (это необходимо, если различные части наименования содержатся в разных колонках);
- исправление неправильного наименования возможно в интерактивном режиме;
- на результаты обработки можно наложить фильтр — например, показывать только ошибочные наименования.

Приложение существует в двух вариантах: с графическим пользовательским интерфейсом и в виде сервиса NT, к которому возможно обращение через API. Второй вариант удобен при создании собственных приложений, которые будут в фоно-



Рис. 2. Парсер наименований в неактивном состоянии

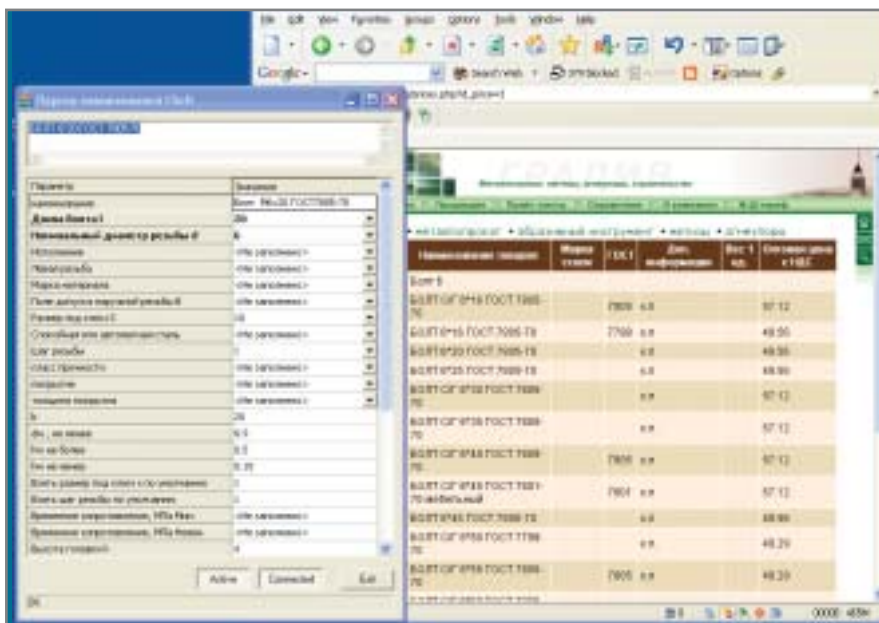


Рис. 3. Пример использования парсера наименований для анализа текста web-страницы

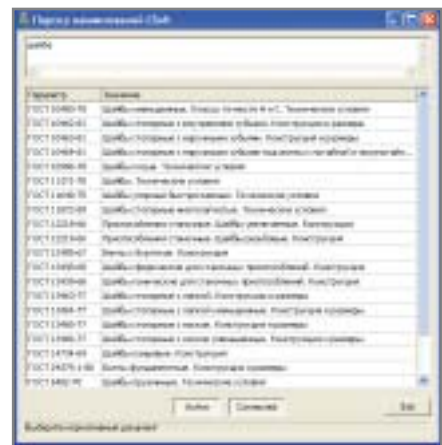


Рис. 4. Выбор схемы стандартного элемента из списка

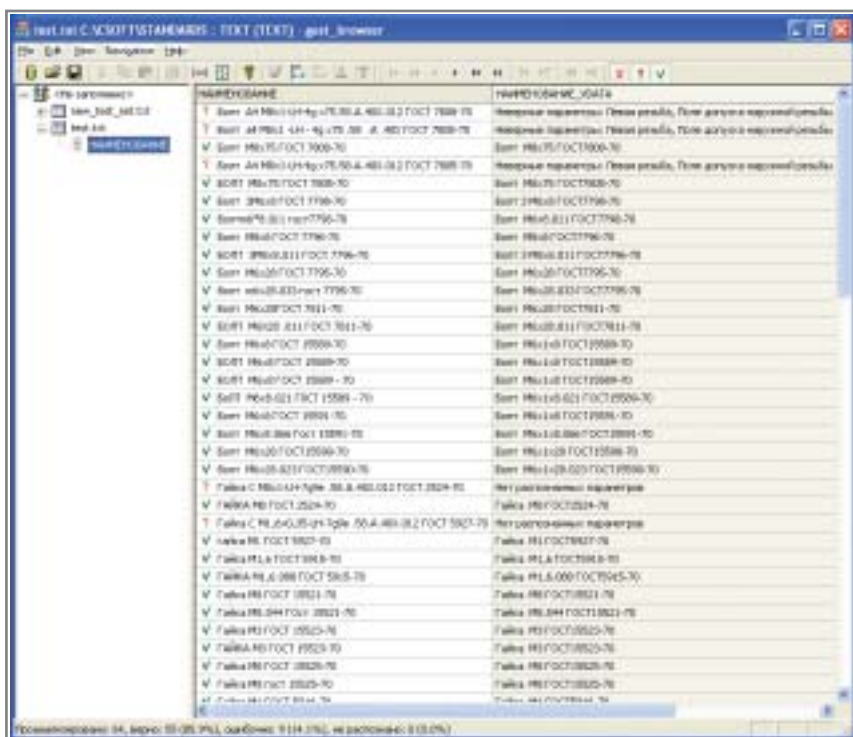


Рис 5. Пользовательский интерфейс приложения *Анализатор/валидатор стандартных наименований*

вом режиме обрабатывать большие объемы информации. На базе этого сервиса разработано web-приложение, предоставляющее все упомянутые выше возможности, причем пользователю не приходится нести дополнительные расходы на администрирование БД стандартных элементов. Это приложение, которое выполняет разбор наименований и их нормализацию, доступно по адресу <http://dev.csoft.spb.ru/ss>.

Перечисленные приложения не исчерпывают всех возможных применений базы стандартных элементов. В настоящее время ведутся работы по включению распознавателя наименований в систему полнотекстового поиска (FTS).

Это позволит создать индекс для технической библиотеки текстовых электронных документов и практически мгновенно выдавать список документов, содержащих или требуемое наименование, или наименование, логически эквивалентное требуемому. Существующие системы FTS такой возможностью не обладают.

Если пользователь работает в программе, которая позволяет под-

ключать функции из внешней динамической библиотеки (DLL), то база стандартных элементов может быть использована для непосредственного контроля пользовательского ввода.

#### Технология заполнения базы данных стандартных элементов

Для поддержания базы данных разработана технология, включающая процесс заполнения БД из нор-

**В НАСТОЯЩЕЕ ВРЕМЯ БАЗА ДАННЫХ СОДЕРЖИТ ИНФОРМАЦИЮ ПРИМЕРНО ПО ТРЕМСТАМ ГОСТам: КРЕПЕЖ, РЕЖУЩИЙ ИНСТРУМЕНТ, В МЕНЬШЕЙ СТЕПЕНИ МАТЕРИАЛЫ. УКАЗАННЫЕ ДАННЫЕ ПОКРЫВАЮТ БОЛЕЕ 80% ПЕРЕЧНЯ КРЕПЕЖА, ТИПИЧНОГО ДЛЯ ПРЕДПРИЯТИЯ.**

мативных документов, существующих в самом разнообразном виде (бумажные, электронные сканированные, электронные текстовые). Этот процесс в значительной степени автоматизирован. Сначала необходимо выверить результаты сканирования и распознавания, затем заполнить конфигурационные файлы (главным образом, определяющие структуру наименования).

Далее автоматически выполняются следующие операции:

- данные проверяются на соответствие типу, отсутствие дублирования и соответствие друг другу наборов значений в разных таблицах. При необходимости могут быть заданы дополнительные критерии проверки — например, возрастание по строкам или столбцам;
- таблицы нормативного документа трансформируются к нормализованной форме так, что их естественное соединение (natural join) обеспечивает полный набор возможных сочетаний параметров стандартного элемента;
- создаются скрипты на языке SQL, выполнение которых приводит к созданию в базе данных структур и объектов, необходимых для работы со стандартным элементом.

Таким образом, все операции, требующие высокой квалификации, — например, проектирование структуры и написание SQL — выполняет программа, а не проектировщик БД. Кроме того, автоматизированная проверка повышает степень достоверности результатов сканирования и распознавания.

#### Текущее состояние справочно-информационной базы данных

- В настоящее время база данных содержит информацию примерно по тремстам ГОСТам: крепеж, режущий инструмент, в меньшей степени материалы. Указанные данные покрывают более 80% перечня крепежа, типичного для предприятия.
- Поддерживаются СУБД MS SQL и POSTGRES (в дальнейших планах — Oracle).
- Существует открытый web-сервис (<http://dev.csoft.spb.ru/ss>), обеспечивающий возможность в тестовом режиме провести проверку и нормализацию различных текстовых документов и/или баз данных.

*Вадим Александров,*  
руководитель программных проектов  
*Семен Козменко,*  
руководитель отдела программных проектов  
*Consistent Software СПб*  
Тел.: (812) 430-3434  
E-mail: [valexandrov@csoft.spb.ru](mailto:valexandrov@csoft.spb.ru),  
[vkozmenko@csoft.spb.ru](mailto:vkozmenko@csoft.spb.ru)



Только вперед

Лучшие проектировщики всегда в курсе новых идей. Новое семейство продуктов AutoCAD 2005 позволит вам с недостижимой прежде эффективностью создавать проектную документацию, управлять совместной работой и обмениваться идеями как с вашими коллегами, так и с заказчиками. Получите исчерпывающую информацию, посетив наш Internet-сайт, и перейдите на новый уровень производительности. [www.autodesk.ru](http://www.autodesk.ru)

Официальный дистрибьютор Autodesk в России **Consistent Software®**  
Тел.: (095) 913-2222, факс: (095) 913-2221 E-mail: [sales@csoft.ru](mailto:sales@csoft.ru) Internet: <http://www.consistent.ru>

**autodesk®**

Autodesk, логотип Autodesk, AutoCAD® и AutoCAD LT® – зарегистрированные торговые марки и торговые марки корпорации Autodesk в США и/или других странах. Все другие названия брендов, продуктов и торговых марок принадлежат соответствующим владельцам. © Autodesk, Inc. Все права защищены.