

# Открытые системы

СУБД

№05  
2012

ISSN 1028-7493

ИТ для бизнеса —  
архитекторам  
информационных систем

[www.osmag.ru](http://www.osmag.ru)



## БОЛЬШАЯ АНАЛИТИКА

Data Science — наука, которой предстоит родиться

• Тестирование облачных сервисов •

Серверы для экстремальных сред • ERP, PLM и BPM в одном флаконе

• Есть ли будущее у ИТ-департаментов?

# TrakCare Lab.

## Advanced technology for breakthrough benefits.



реклама

Лабораторная информационная система InterSystems **TRAKCARE LAB** – продукт, позволяющий автоматизировать большинство процессов в лаборатории. Это многофункциональная система, поддерживающая все направления лабораторной деятельности, а также смежные сервисные функции. Система содержит более 200 драйверов к различному лабораторному оборудованию и позволяет использо-

вать его максимально эффективно. TrakCare LAB подходит как для небольших лабораторий, так и для крупных лабораторных центров.

Система TrakCare LAB установлена в 400 лабораторных учреждениях в 25 странах, включая как отдельные лаборатории, так и общенациональные сети лабораторного обслуживания.

InterSystems  
**TRAKCARE™**

Learn more at [TrakCare.com/lab](http://TrakCare.com/lab)

© 2012 InterSystems Corporation. All rights reserved. InterSystems and TrakCare™ Lab are trademarks of InterSystems Corporation. 1-12 AdvTrakCare

# Зона Больших Данных

«**Б**ольшие Данные — это классика или джаз?» — был задан вопрос на одной из российских конференций, посвященных проблемам обработки огромных массивов данных. Скорее, джаз, а применительно к аналитике, оказавшейся сегодня в зоне Больших Данных, — безусловно: в ближайшем будущем здесь не стоит ожидать спокойного скрупулезного раскладывания данных по полочкам для их последующего неторопливого изучения с целью выявления новых фактов или тенденций прошедшего времени.

Как следует из статей этого номера, посвященного аналитике Больших Данных, впервые за всю историю человечества сырьевой базой для его развития становятся не природные элементы, а, как отмечает Леонид Черняк, искусственно созданные и нематериальные — данные (некоторые авторы даже говорят о данных, как о «новой нефти»). Этот сдвиг парадигмы еще не до конца осознан, не выработана терминология, нет даже определения ее субъектов (например, понятия data scientist), но уже очевидно, что в зоне Больших Данных стабильность и размеренность невозможны, зато востребована импровизация, а сами компании и организации будут больше похожи на джаз-банды, чем на симфонические оркестры.

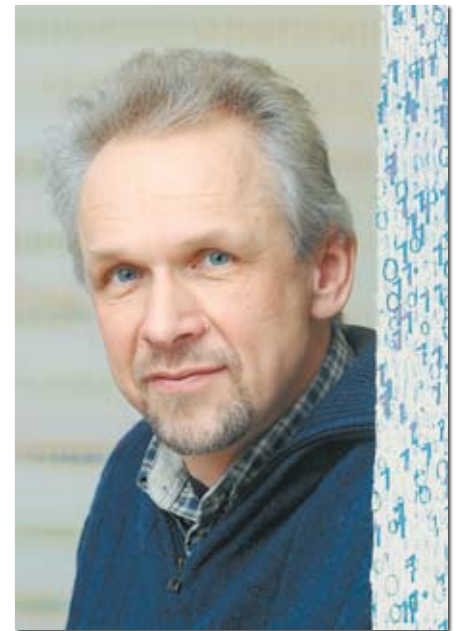
В эру Больших Данных успешными будут компании, способные эффективно пользоваться всеми доступными информационными активами для принятия превентивных решений и получения повышающих конкурентоспособность ценных знаний. Вместе с тем во всем мире сейчас ежедневно генерируется почти 3 тыс. терабайт сырых данных, что усложняет предприятиям работу по поиску, сбору, систематизации и анализу поступающих сведений. Конечно, учитывая некомпетентность большинства, вряд ли стоит рассчитывать, что широко распространенная точка зрения окажется разумной, а следовательно, нужна фильтрация всего потока сведений, тем более что в среднем мы ежедневно обдумываем 60 тыс. мыслей, из которых большая часть — те же самые, что были вчера. Однако и после фильтрации остается еще весьма

обширная зона для применения аналитических технологий.

Организация аналитической работы с Большими Данными для обнаружения нужной информации стала сегодня одной из важнейших задач любой отрасли, однако вопрос не только в том, как обрабатывать, а зачем и какую практическую пользу может принести каждый конкретный тип обработки. Можно ли увеличить ценность данных хотя бы на порядок? Как из мозаичного представления восстановить целостную картину, учитывая данные, которыми ранее пренебрегали или не замечали? Ответы можно найти в этом выпуске журнала.

На стыке Больших Данных, параллельных архитектур и облаков родилось новое направление — специализированные машины для поддержки аналитики обработки крупномасштабных массивов данных. Чтобы быть эффективным в любом деле, следует быть ближе к предмету своей деятельности — для анализа не надо загружать данные в хранилища, чтобы потом их извлекать, а делать это следует либо непосредственно в базе данных, либо в оперативной памяти. Сегодня аналитические машины внедрены почти у половины из тысячи крупнейших предприятий мира, использующих новые аналитические технологии для решения различных задач, направленных в конечном счете на получение Больших Денег. Правда, как отмечает один из авторов этого выпуска, в России главным стимулом выбора технологий для работы с Большими Данными будет все-таки объем данных, которые требуется обработать предприятиям под давлением регуляторов, а решение коммерческими компаниями таких задач, как маркетинг, оценка настроений потребителей путем анализа различных медиа и даже глубокая обработка геологических данных, пока отходит на второй план.

С точки зрения ИТ-службы, новые технологии — это возможность соответствовать стратегическим задачам компании и способствовать росту ее бизнеса, управляя ИТ-активами. Например, считается, что, в эпоху сервисов, ITSM для облаков — это как электричество для компьютера. Однако, как отмечает Филис



Друкер, перевод управления информационными технологиями в новое измерение наталкивается на сопротивление ИТ-департаментов компаний, которые воспринимают появление очередного новшества как угрозу своему спокойствию и безопасности и всячески тормозят применение новых инструментов в корпоративной среде.

Подобно тому как в пору становления США как государства в его плавильном котле были переплавлены европейские национальности, в зоне Больших Данных переплавляются сегодня передовые технологии для получения нового качества. Кстати, надо отдать должное администрации этой страны, уделяющей серьезное внимание самым разным технологиям и недавно выступившей с инициативой «Большие Данные: исследования и разработки». Такие инициативы в прошлом способствовали созданию суперкомпьютеров и Интернета, а нынешняя обещает затянуть в зону Больших Данных все направления ИТ, которые уже стали в развитых странах важнейшим фактором в обеспечении экономической конкурентоспособности. Речь идет о трансформации науки, медицины, образования, общественно-политического устройства путем движения к открытому правительству и, конечно, о национальной безопасности. Большая часть исследований и разработок в области ИТ осуществляется в США частными компаниями, и хотя данная инициатива поддержана пока еще небольшим финансированием, сигнал дан.

*Дмитрий Волков*



IT for Business  
Innovative Technology for Computer Professionals

## COVER FEATURES

### BIG DATA ANALYSIS

#### 16 Tools for the «Big Analysis»

Leonid Chernyak

New IT discipline is being created connecting Big Data, parallel architectures and cloud computing: customized analytic appliances capable of processing large volumes of information.

#### 23 Big Data Analysis Platform

Denis Serov

Big Data is increasingly compared to «new petroleum» nowadays. Like any raw resource, data itself do not have value, so the main task of the industry is to transform raw data to get all the benefits it could bring. To that end, EMC offers an integrated platform for data processing and analysis.

#### 26 Analysis without Limits

Artem Grishkovskiy

Lots of attention is paid to Big Data today. It is often thought that working with Big Data involves lots of tough challenges, but is this really the case? Sybase IQ solution and SAP HANA database, integrated on top of SAP Real-time Data Platform, can perform tasks related to analyzing big data volumes in real time.

#### 28 Big Data and Systems Integration

Maksim Isayev

Russian businesses' demand for Big Data processing solutions is growing, but IT departments do not have enough capabilities to be the source of projects in this area. This opens plenty of opportunities for system integrators that already have experience of building infrastructures for working with big volumes of data.

#### 30 The Big Analytics for Big Data

Alexey Lonshakov, Alexey Mesheryakov, Andrey Svirschevsky,

Current problems for Big Data can't be solve under traditional tools — but the new solutions by the SAS can do it. So it doesn't meter which kind of data used for analytical tasks and which volume can be used.

### PLATFORMS

#### 12 Carrier Class IT Systems

Oleg Kholodniy

The developers of computing equipment for the telecom industry traditionally have to meet a very specific set of customer requirements. Carrier class servers mounted within the standard 19-inch rack is a typical example of today's approach to designing hardware solutions for the telecommunications market sector.

### IT MANAGEMENT

#### 32 Self Service, the Cloud and ITSM...

One Big Happy Family?

Phylis Drucker

Welcome to the new reality of computers. We've lost control! But while we've lost control, we're now capable of providing a more robust and responsive computing environment than ever before. For those organizations willing, ready and able to welcome innovation, the sky's the limit to providing what the business needs to drive competition.

#### 36 The Future of Enterprise in the Cloud

Jamie Erbes, Hamid R. Motahari-Nezhad, Sven Graupner

The widespread availability and adoption of cloud services creates new challenges for enterprise IT and prompts the need for new methodologies, tools, and skill sets for managing a hybrid portfolio of these services and traditional IT systems.

### EXPERIENCE

#### 41 A Key to Risk Assessment

Olga Revyakina

Information on customers, products and markets is especially valuable for insurance companies. Business intelligence systems are their «eyes» that help watch market reality closely, make best decisions and pave the way to success. As AlfaStrakhovanie's experience has shown, the more efficient BI toolset is used by an insurer, the better the company's business goals are met.

#### 44 Virtual Diagnosis Lab

Vladimir Ostrovskiy; Yuriy Kashkarov

The topics of virtualization and remote access are really popular today. They are even discussed in the community of lab information system providers, which have divided into two camps, one opposing and the other being in favor of distributed deployment of laboratory diagnosis resources. A European medicine center has implemented a lab IS project proving that new technology has come even to the conservative area of medicine.

### SOFTWARE ENGINEERING

#### 46 Cloud Service Testing

Natalya Efimtseva

For any software development company, the transition from the traditional model of creating and distributing applications to SaaS leads not only to changing finance and business processes, but to transforming the software building process itself. One of the implications is that cloud services testing becomes the most important stage of the application life cycle. Visual Studio 2012 IDE provides tools for unit, functional and automatic testing of applications deployed in the cloud.

### APPLICATION

#### 52 ERP, PLM and BPM

in an All-in-One Solution

Nikolay Nyrkov

Could all the necessary tasks be performed by just one business software system? This question has the right to be discussed at least in the context of the area of design and manufacturing.

### OPINION

#### 54 US Administration and Big Data

Leonid Chernyak

Big Data Research and Development Initiative proposed by the U. S. government is administrative by nature and completely ignores the personality factor. Let's consider, why.

### OS ACADEMY. IT UNIVERSITIES

#### 56 Data Science, the Research Discipline Yet to be Born

Leonid Chernyak

Objective demand for a science studying data is undeniable, but Data Science as it exists today do not yet have the desirable characteristics of the research field dedicated to studying the subject of data.

### OS ACADEMY. LIBRARY

#### 58 Computer Microbiology, New Computer Interaction Tools, and Digital Video Screens

Sergey Kuznetsov

The March, April and May 2012 issues of IEEE Computer Magazine (IEEE Computer Society, V. 45, No. 3, 4, 5, 2012) are dedicated to using computers in the experimental biology, future computer interaction methods, and digital video screens, accordingly.

<http://infozone.osp.ru>

• Около 400 000 посетителей сайта ежемесячно из России, Украины, Белоруссии, Казахстана и других стран ближнего и дальнего зарубежья

• Свыше 9000 зарегистрированных подписчиков

Открытые системы  
ЖУРНАЛ ДЛЯ АРХИТЕКТОВ ИНФОРМАЦИОННЫХ СИСТЕМ

computerworld

Windows IT PRO

Директор информационной службы

Мир ПК

ЖУРНАЛ СЕТЕВЫХ РЕШЕНИЙ

Сети

Продукты и услуги OSP.RU

События  
Статьи  
Вебкасты  
Конференции  
Аналитика  
OSP TV

Читатели и посетители сайта



Темы

Обеспечение непрерывности бизнеса  
Слияния и поглощения  
Управление бизнес-процессами  
ИТ в здравоохранении  
...  
Десятки тем и зон по продуктам и решениям

Заказчики ИТ



ОТКРЫТЫЕ СИСТЕМЫ  
Open Systems Publications

<http://infozone.osp.ru>



Источник: Rainer Lippert/Wikipedia

## В Германии поставили рекорд выработки энергии на солнечных батареях

В конце мая немецкие солнечные электростанции поставили мировой рекорд выработки электроэнергии, пишет The Guardian: 22 гигаватта в полуденные часы пятницы и субботы. Для сравнения, такую же мощность вырабатывают 20 атомных электростанций.

После прошлогодней аварии на японской АЭС «Фукусима-1» правительство Германии решило отказаться от использования ядерной энергии: в стране немедленно были остановлены восемь АЭС, а остальные девять планируется остановить к 2022 году. Вместо них будут использоваться возобновляемые источники: ветер, солнце и биомасса. Сегодня в стране установлено уже почти столько же солнечных батарей, сколько во всем остальном мире. К 2020 году в Германии надеются сократить выброс парниковых газов в атмосферу на 40% по сравнению с 1990 годом.

В немецком Институте возобновляемой энергии утверждают, что рекордной выработки хватило, чтобы при должном перераспределении выработанной и накопленной электроэнергии в полдень в пятницу и на 50% в субботу. Переход на солнечную энергию имеет и негативные стороны: такая энергия стоит дороже, в результате чего плата за электричество в стране сейчас среди самых высоких в мире — 23 евроцента за кВт×ч.



## Исследователь построил IP-сеть из двух ксилофонов

Аспирант Калифорнийского университета в Беркли Стюарт Гейгер провел демонстрацию IP-сети, соединяющей два компьютера, в которой пакеты передавались путем «проигрыва-

## Нашу галактику ждет неизбежное столкновение с туманностью Андромеды

Через четыре миллиарда лет Млечный Путь прекратит свое существование в его нынешнем виде, поскольку столкнется «лоб в лоб» с туманностью Андромеды, обнаружили исследователи из балтиморского Научного института космического телескопа. В результате образуется совершенно новая гибридная галактика эллиптической, а не спиральной, как у Млечного Пути, формы, предсказывают ученые.

Астрономам известно, что Млечный Путь и туманность Андромеды несутся друг другу навстречу под действием сил гравитации. Но столкновение рассматривалось лишь как вероятность, поскольку была неизвестной скорость поперечного движения туманности Андромеды. Исследователям из Балтимора удалось измерить этот показатель путем систематического наблюдения за определенными регионами галактики с помощью космического телескопа NASA Hubble. Ученые пришли к выводу о неминуемости столкновения. Слияние двух галактик, начавшись через 4 млрд лет, окончательно завершится через 6 млрд лет. Событий такого масштаба еще не было в истории Млечного Пути, формирование которого началось около 13,5 млрд лет тому назад. Наша Солнечная система в результате слияния переместится на новое место, которое будет гораздо дальше от ядра галактики, чем сейчас. Столкновение с какими-либо звездами маловероятно — расстояние между ними, несмотря на слияние галактик, останется огромным. Вместе с тем, отмечают исследователи, процесс, скорее всего, приведет к образованию большого количества новых звезд.



Источник: NASA, ESA

ния» их на ксилофонах. Для этого понадобились два микроконтроллера Arduino, два ксилофона и датчики. Компьютер отправляет пакет в символьном виде на микроконтроллер, который преобразует коды символов в шестнадцатеричный вид. К Arduino подключен ряд светодиодов, каждому из которых соответствует нота на ксилофоне и шестнадцатеричная цифра. Как только загораются светодиоды, человек должен ударить по соответствующим нотам. Прикрепленные к нотам пьезодатчики регистрируют удары и передают принимающему Arduino. Тот преобразует сигнал датчика в шестнадцатеричный код и отправляет его компьютеру назначения.

Скорость передачи составляет один символ в секунду — на передачу целого пакета обычно уходит около 15 минут. Автор утверждает, что эксперимент помог ему по-новому осознать сетевую модель OSI, в которой все уровни изолированы друг от друга. В эксперименте люди выполняли роль самого низкого, физического, уровня OSI, но для компьютеров не имело значения, каким именно способом передавались биты данных.



# Цикл конференций для руководителей и ведущих специалистов современных предприятий



**ОТКРЫТЫЕ  
СИСТЕМЫ**  
Open Systems Publications



Регистрация:  
Тел.: (495) 956-3306, e-mail: [kon@osp.ru](mailto:kon@osp.ru)  
<http://www.ospcon.ru>

**Издательство «Открытые системы»** выпускает 14 различных журналов и газет, которые выходят ежегодным тиражом более 10 миллионов экземпляров. Лидирующие позиции на рынке СМИ обеспечиваются широким спектром изданий для специалистов и руководителей из таких отраслей, как информационные технологии, нефтегазовая отрасль, телекоммуникации, полиграфия и медицина. Активное сотрудничество с мировыми издателями, такими как IDG, PenWell, позволяет нам поддерживать высочайший уровень и обеспечивать актуальность публикаций. Наличие более 30 региональных представительств дает возможность издательству распространять свои журналы по всей территории России и СНГ.

**Агентство корпоративных коммуникаций OSP-Con** – подразделение издательства «Открытые системы», в которое в 2004 году был выделен весь бизнес, связанный с организацией и проведением мероприятий. Сегодня OSP-Con является организатором многочисленных конференций для отечественного ИКТ и бизнес-сообщества. Основной акцент в своей деятельности OSP-Con делает на детальную проработку контентной составляющей проводимых мероприятий, начиная с определения тематики конференций и кончая предоставлением ее участникам собственной аналитической экспертизы. Именно это обстоятельство сделало конференции, проводимые OSP-Con, заметным явлением на российском рынке и позволило агентству занять на нем ведущие позиции.

## ГРАФИК ТЕМАТИЧЕСКИХ КОНФЕРЕНЦИЙ Агентства корпоративных коммуникаций OSP-Con на бизнес-сезон 2012 г.

18	октября	Ethernet-форум 2012, гостиница «Radisson-Славянская»
21	ноября	Второй Российский Форум «Интеграция сложных прикладных систем» (ICAS-2012), Международный информационно-выставочный центр «ИнфоПространство»



# РАССЧИТАНО НА БУДУЩЕЕ, РАБОТАЕТ СЕЙЧАС

## Частное облако от Microsoft

Узнайте подробнее на **[Microsoft.ru/readynow](https://microsoft.ru/readynow)**.



Windows Server



Microsoft  
System Center