OSBR.CA

The Open Source Business Resource

SEPTEMBER 2009

SEPTEMBER 2009

PUBLISHER:
The Open Source Business Resource is a monthly publication of the Talent First Network. Archives are available at the website:
http://www.osbr.ca

EDITOR:
Dru Lavigne
dru@osbr.ca

ISSN:
1913-6102

ADVISORY BOARD:
Tony Bailetti
James Bowen
Leslie Hawthorn
Chris Hobbs
Rikki Kite
Thomas Kunz
Steven Muegge
Donald Smith
Michael Weiss

The editorial theme for the September issue of the OSBR is business intelligence (BI). The authors in this issue provide an overview of the importance of BI, the tools which are available, and the role that open source is playing in this increasingly important field.

Sam Selim, Founder and Chief Technology Officer of SQL Power Group, discusses how companies can reduce their costs and dependency on proprietary BI tools by adopting an open source BI model.

John Kemp, a Principal Consultant, and Benjamin Dietz, a BI Consultant, provide a primer on BI, introducing the terms and concepts used throughout the issue.

David Currie, founder of Clearview Informatics, examines the keys to success that allow an organization to extract maximum value from an investment in BI software.

John Kemp takes a look at how buying patterns have changed and what it means for businesses looking at open source BI software.

Steve Holub, a BI Consultant, reports on a recent survey of open source software tools used in BI and data warehousing systems.

Benjamin Dietz and Lily Singh of SQL Power Group compare the features of four popular open source reporting solutions.

Thierry Badard, CTO of Spatialytics, highlights the need for geospatial BI software and the integration of the spatial component in a BI software stack in order to consistently enable geo-analytical tools.

Tom Bondur, Content Development Manager in the Developer Communications group of Actuate Engineering, and Jason Weathersby, the BIRT Evangelist at Actuate Corporation, introduce BIRT and discuss why Actuate chose the open source development model and the benefits that this project brings to the BI user community.

Sam Selim answers the question "What will Oracle do with Sun's open source offerings?".

As always, we encourage readers to share articles of interest with their colleagues, and to provide their comments either online or directly to the authors. We hope you enjoy this issue of the OSBR.

The editorial theme for the upcoming October issue of the OSBR is "arts and media" and the guest editor will be Anthony Whitehead from Carleton University.

Dru Lavigne

Editor-in-Chief

dru@osbr.ca

Dru Lavigne is a technical writer and IT consultant who has been active with open source communities since the mid-1990s. She writes regularly for BSD Magazine and is the author of the books BSD Hacks, The Best of FreeBSD Basics, and the upcoming Definitive Guide to PC-BSD.

For the past decade, traditional proprietary tools from Cognos, Business Objects, Microstrategy, Hyperion, Oracle and Informatica have dominated BI environments throughout North America and Europe. The larger the customer it seemed, the bigger their appetite for spending money on proprietary BI tools. BI vendors have made billions selling software to larger clients throughout North America, much of it in the form of multi-million dollar site licenses. Large customers continue to pay huge perpetual annual maintenance fees, usually 20+ % of the initial license cost. Costs for additional users and new features requiring additional license fees can quickly add up to 40% - 50% of the annual BI budget. This has caused many organizations to scrutinize and question the value of their ongoing BI software expenditures.

In these challenging financial times, cost conscience CIOs and Directors around the world are looking for ways to reduce the ongoing costs of administering their BI program, and a large number of them are turning to open source BI tools. Table 1 summarizes some of the differences between proprietary and open source BI tools.

When compared to proprietary BI solutions, open source alternatives offer significant advantages:

Free end-user deployment: proprietary BI software licenses are typically per-user, thus penalizing organizations for their BI success and for wide end-user adoption. Most open source BI tools offer server-based subscriptions that offer unlimited end-user deployments at no additional cost.

Table 1: Proprietary vs. Open Source BI Tools

	Typical Proprietary BI Tools	Open Source BI Tools
Up-front License Fees	$100,000+ up-front license fees	$0 up-front license fees
Ongoing License Fees	License Fees for each end-user	Free End-user Deployment: No License Fees for each end-user
Ongoing Maintenance Fees	20% Annual Maintenance	Optional server-based subscription fees
Source code access	No access to source code for modifications	Full access to source code for modifications
Product Evolution	Limited input in product evolution	Ultimate input in product evolution
Support Base	Commercial Support only	Community or Commercial Support

Higher quality code: unlike closed source proprietary BI software, open source software is generally more modular, with its code reviewed, critiqued and enhanced by a potentially larger number of code contributors.

Low risk: everyone has access to the source code and developers around the world are able to extend or fork it should the project sponsor go bankrupt or decide to pursue another agenda.

Logical product evolution: open source communities tend to evolve the product based on real life needs and worldwide market demands rather than on finances, sales targets or a marketing agenda.

Product stability: there has been a huge consolidation in the proprietary BI software market over the past few years, so the company you purchased your BI tool from may not be the same company that is now providing you with support and product enhancements.

Wide adoption of open source BI tools: open source BI tools have evolved significantly over the past four years, and now serve to offer a complete low cost alternative to traditional proprietary BI tools. The maturity of these tools, along with the need to reduce annual BI software budgets, is causing hundreds of organizations throughout North America and Europe to start launching proof-of-concepts and adopting them for new departmental projects.

The above arguments for reducing the dependency on proprietary BI tools are gaining worldwide acceptance among CIOs and Directors. Recent examples include the adoption of open source BI Tools at the Bank of America, Motorola, Expedia, the US Air Force and Google. Organizations that are open to using open source BI tools can save close to 50% of their annual BI budget. They can utilize these funds to hire new resources, deliver more BI functionality, or bolster their company's bottom line. This growing adoption of open source BI tools has already taken a bite out of proprietary BI tools vendors' profits and market share. We expect this trend to continue over the next five years, further eroding the popularity of traditional BI tools and the market share of its vendors. The maturity of open source BI tools has also changed the financial landscape of the BI industry. No longer are organizations willing to spend large amounts of money up-front on BI. Instead, most BI program managers are more comfortable with a pay-as-you-play model, usually in the form of annual subscriptions or other forms of software support, from open source vendors. These annual subscriptions are usually in-line with the 20% annual maintenance that proprietary software vendors charge, but without the up-front license fees. Additionally, organizations are freed to switch tools or use multiple tools without incurring any major software investment.

The authors in this issue of the OSBR examine the various open source alternatives to proprietary BI tools, offer in-depth comparisons of the leading open source BI offerings, and provide thoughts on how to assess and select BI solutions.

Sam P. Selim is the Founder and Chief Technology Officer of SQL Power Group Inc. (http://www.sqlpower.ca), a leading Canadian Consulting firm specializing in Data Warehousing and Business Intelligence. Mr. Selim is the lead architect for many of SQL Power's open source data warehousing and business intelligence tools and consulting engagements. He has over 16 years experience designing and developing custom data warehousing solutions for SQL Power clients. Sam has been a presenter at several data warehouse conferences and tradeshows and has written extensively on these topics.

"A strategic Business Intelligence platform puts the right information in the right hands at the right time, and gives managers and executives the ability to test various scenarios for business spending and investments while monitoring important operational drivers of company performance."

The Importance of Business Intellience
http://tinyurl.com/n3xhr5

This article provides a primer on business intelligence (BI) and serves as an introduction to the concepts used throughout the rest of the articles in this issue of the OSBR. We define BI, discuss the components of a BI solution, explain the types of BI tools and provide a brief overview of the evolution of BI.

What is Business Intelligence?

Hans Peter Luhn (http://en.wikipedia.org/wiki/Hans_Peter_Luhn) first defined BI in his article "A Business Intelligence System", published in the October 1958 IBM Journal. More recently, Howard Dresner of the Gartner Group popularized the following definition for BI: An umbrella term to describe "concepts and methods to improve business decision making by using fact-based support systems."

BI, as a practice, has really come to the forefront in the last fifteen years. It is primarily used as a way for organizations to make sense of the mountains of data that have become available to them due to the proliferation of inexpensive storage and the ability to collect large amounts of data via multiple input sites.

What Makes Up a Business Intelligence Solution?

BI solutions grew out of a combination of increased globalization, competition, and pervasiveness of information systems.

The original BI systems were little more than copies of transactional databases. These allowed reports, developed by system programmers, to be run against the data without impacting the performance of production systems.

However, any report changes or new reports required intervention by the information technology group, since the data models used by transactional systems were highly complex and often cryptic. They often utilized space saving techniques to reduce the data storage. On top of that, it was pretty much impossible to combine information from separate systems.

The need for current and complete data to facilitate business decisions, the proliferation of systems and data, and a desire to drive the data to the business user has resulted in the explosion of BI over the past decade.

So, what components are required to build a BI solution? A traditional approach to BI is usually comprised of three major components as shown in Figure 1.

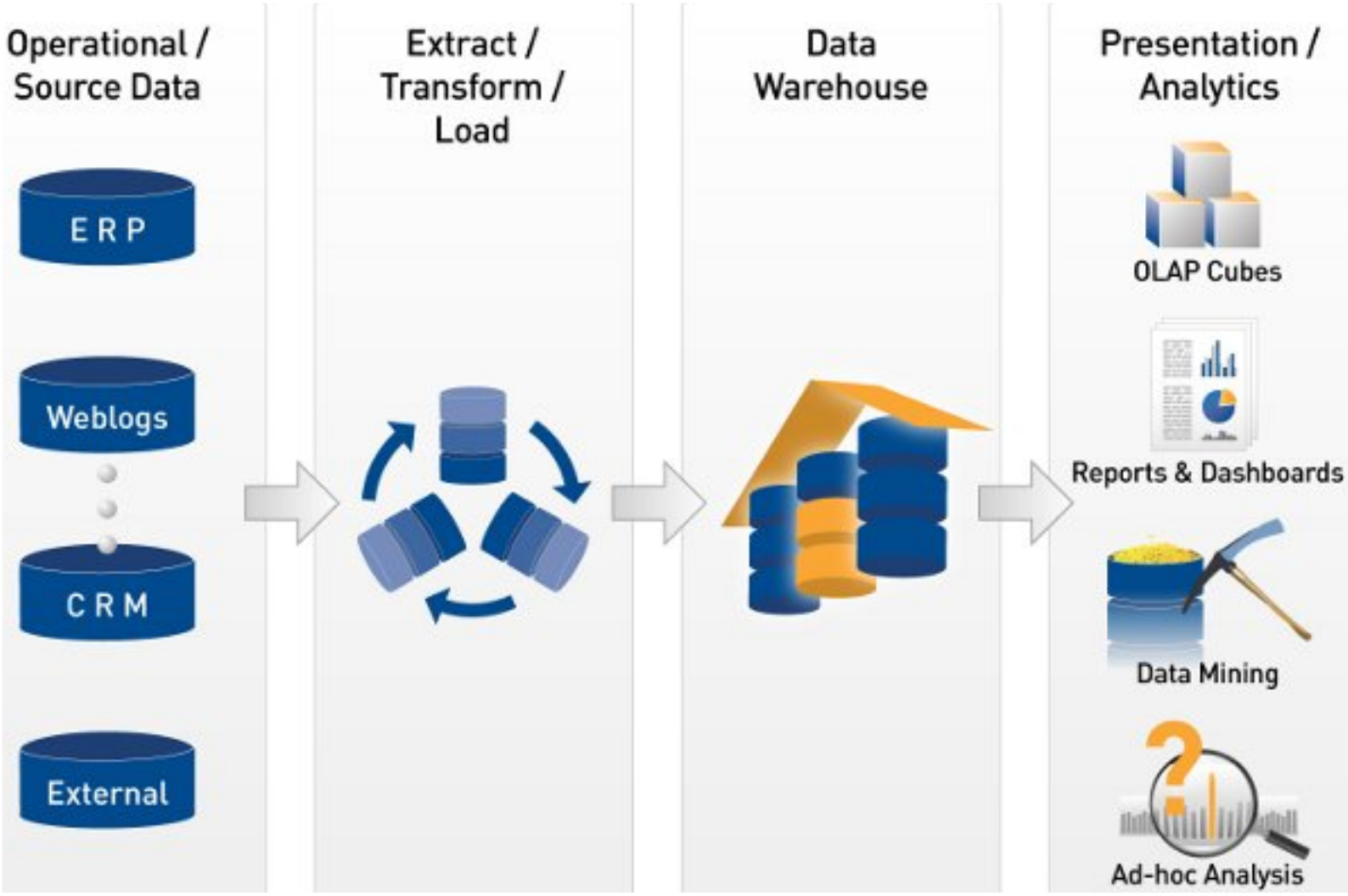
1. Data Storage

Organizations typically store the extracted and transformed data in a separate database distinct from the production systems. This insulates the production systems from any 'run away' queries that could be generated by users. It also allows the organization to control updates to the data since there is often a desire to ensure that everyone reports from the same time series.

2. Extract/Transform/Load

Extract/Transform/Load (ETL, http://en.wikipedia.org/wiki/Etl) describes the process and tools whereby data is extracted from source or transactional systems.

Figure 1: Components of a a Business Intelligence Solution



Typically, data is extracted and changed or transformed prior to loading into the reporting database which is more commonly referred to as a Data Warehouse or Data Mart. The transformation is meant to accomplish the following:

Provide data in a consistent format. The same data may be represented differently depending upon the source system. Gender may be represented as Male/Female in one system, M/F in another, and even 1/2 in a third system. The ETL process supports the transformation of data to a consistent format.

Transform data to a format easily understood by the business. Transactional systems are designed to process information as quickly as possible. As a result, the relational data models employed are often very complex and difficult to understand, even for seasoned database programmers. A Dimensional Data Model (http://en.wikipedia.org/wiki/Data_warehouse#Normalized_versus_dimensional_approach_for_storage_of_data) or Star Schema Model (http://en.wikipedia.org/wiki/Star_schema) is typically employed in order to represent data in a way that is both easily understood by the business and that expedites database access.

Data may be enriched or corrected when transformed. For example, integration of multiple customer records to a single customer record may be undertaken, and rules may be applied to ensure that the data is correct. Customer records may be segregated by region or area based on the zip or postal code. This correction step is now being eschewed in favour of correcting data at the source level, as the corrective activities themselves can build in incorrect assumptions regarding the data.

3. Reporting and Analysis Tools

There is a plethora of BI tools available for the surfacing and analysis of data. Most organizations will utilize multiple tools as there isn't a 'one size fits all' BI tool. To describe the tool, we segregate their users into categories:

1. Consumers: typically use delivered reports or dashboards that provide summarized information presented on a web page. If further analysis is required, another member of the organization undertakes the work. Consumers are often senior executives within an organization who don't have the cycles to do the analysis and, more importantly, have individuals who can undertake the analysis. Open source software (OSS) offerings include BIRT (http://eclipse.org/birt), Pentaho (http://pentaho.com/) and Jaspersoft (http://www.jaspersoft.com/). Micro Strategy (http://.www.microstrategy.com), Business Objects (http://www.sap.com/solutions/sapbusinessobjects) and Cognos (http://ibm.com/software/data/cognos) provide commercial BI software.

2. Analysts: often utilize tools that allow a structured review of the data. These include configurable reports and queries or the use of data cubes, often referred to as Online Analytical Processing (OLAP, http://en.wikipedia.org/wiki/Olap), to slice and dice the data and to drill-through to details. In addition to the commercial leaders listed above, Wabit (http://sqlpower.ca/page/wabit), Pentaho, and Jaspersoft provide OSS products.

3. Explorers: undertake detailed analysis of data. They often use ad-hoc reporting tools or program their own queries using Structured Query Language (SQL, http://en.wikipedia.org/wiki/Sql). Additionally, some users may utilize data mining tools (http://en.wikipedia.org/wiki/Data_mining) to undertake advanced statistical analysis of data to drive out patterns and trends. IBM, Pentaho's Weka and R products. Base SAS and Enterprise Mine from SAS, Clementine from SPSS (http://spss.com), are among the commercial leaders in this area.

The Evolution of BI Solutions

Originally, specialized offerings were developed by organizations to provide point specific BI functionality. Over time, larger commercial software vendors acquired these offerings in order to provide fully integrated 'across the board' solutions. Examples include SAP's acquisition of Business Objects, IBM's acquisition of Cognos, and Oracle's acquisition of BRIO and Hyperion. Each of these organizations in turn swallowed up other best of breed solutions prior to being acquired.

This consolidation has meant that organizations looking to implement BI solutions are faced with the choice of purchasing a very expensive consolidated solution from a large vendor rather than acquiring best of breed tools. Open source solutions offer organizations the opportunity to implement best of breed BI Solutions that offer everything the commercial products have in a cost effective and as-needed manner.

Why? Because the freely available nature of open source code means that the community develops the interoperability to published standards, and the OSS model eliminates lock-in from one vendor.

John Kemp is a Principal Consultant with SQL Power Group. John is responsible for the delivery of Business Intelligence and Data Warehousing projects at SQL Power. A 20-year information technology veteran, John previously provided Data Warehouse and Business Intelligence consulting services with KPMG, SAS Institute, Infobright Inc., and as the lead of his own consulting firm.

Benjamin Dietz is a Business Intelligence Consultant at SQL Power Group. Benjamin wrote his Master's thesis on Open Source Business Intelligence and graduated from University of Applied Science Karlsruhe, Germany, with a Master of Business Information Systems.

Recommended Resources

Information Management
http://information-management.com

The BeyeNetwork
http://www.b-eye-network.com

The Data Warehouse Institute
http://www.tdwi.org

"Do not repeat the tactics which have gained you one victory, but let your methods be regulated by the infinite variety of circumstances."

Sun Tzu

A business intelligence (BI) project has to be managed with as much discipline as any other information technology (IT) project in order to be successful. There are a few items that need special consideration given the nature of BI solutions, regardless of the specific methodology or technology involved. This article will discuss how to extract maximum value from an investment in BI software.

BI Challenges

A BI project presents some unique challenges to an organization of any size. The software sits on the cusp of business operations and technology, which demands tight collaboration between end-users and technologists. Without input from the business users, who stand to benefit from better information to help them with their work, the project will have limited value. To obtain useful input from the business users, they must be educated about a complex technology and what is possible with their data. This interdependence rarely exists in other projects. Many IT organizations and consultants have focused on projects that are transparent to users, such as upgrading and enhancing databases and other back office applications, which leave them without experience in dealing with business operations.

Another factor is the often elusive return on investment (ROI) calculation for a BI implementation. A good contrarian example is server virtualization: If virtualization software allows two servers to do the work previously done by ten, then it is easy to calculate the value gained.

No such simple formula exists for calculating the value of BI software. Non-financial gains, such as easier and broader access to timely information, less time spent sorting through spreadsheets, and increased confidence in the numbers, often outweigh the financial gains, but the payback is longer and difficult to measure. This in turn can undermine executive support for the project.

Later in the project, when business users start testing the solution, they will sometimes discover that the numbers do not match their intuitive feelings for the business, or that the data contains surprises. Previously hidden data quality issues or sparsely populated dimensions are common discoveries. This can create a great deal of tension between the business users and the technical team, as data quality is not usually within the scope of the project. The business users and executive sponsors start to worry that the project will not deliver what is required.

There is no silver bullet technology or project management methodology to overcome these unique challenges. However, by giving special attention to the following aspects of the project, the chances of a successful BI implementation can be greatly increased.

Requirements Gathering

There are two types of requirements: the wish list and the necessity. Unfortunately, these often overlap and conflict. An approach suited to a BI project is to map out the business processes that can be enabled or enhanced with better reports and metrics. Think about what decisions are made during the course of doing business, then what would be required in order to make better decisions. Brainstorm what can help grow revenue, increase profits or save lives.

When are managers and staffers relying on gut instinct? What information would tell management if they are on track to achieve their business objectives? What numbers do you wish you and your staff had on hand? For example, is it the top 10 customers or how many customers are served per hour?

Taking Inventory

This is a two part step. The first step is to find the source of the required data that will serve as the raw material for the BI system. Typically, at least part of the data will be sourced from existing enterprise software, such as a customer relationship management (CRM, http://en.wikipedia.org/wiki/Customer_relationship_management) or enterprise resource planning (ERP, http://en.wikipedia.org/wiki/Enterprise_resource_planning) system. Then, the search may have to go further afield to home-grown systems, software outside the firewall such as web analytics, and flat files exported from numerous other sources where database connectivity is impossible.

The second part of taking inventory is acknowledging the existing BI in place. Is the organization heavily reliant on user-generated spreadsheets and email? What database software and reporting tools are currently being used? Many software packages come with reporting tools embedded. Knowing which alternate BI systems are currently in place will allow better decisions about the future of these tools in an organization. Can, or should, they be replaced?

Investigating Software Vendors

There are hundreds, if not thousands, of products that provide reports, dashboards, data visualization, analytics, and extract/transform/load (ETL, http://en.wikipedia.org/wiki/Etl) operations.

There are niche products specialized for specific industries and horizontal BI development platforms. Licensing varies from commercial to open source to dual licenses. Reliance on standards runs the gamut as well, although the recent trend is toward more open standards, even on proprietary platforms. The only way to sort through the myriad of vendors is to define some criteria. But a feature and functionality list is not enough. Consider also the following questions: What software packages meet the requirements and fit into your existing infrastructure? How much customization will be necessary? Are the licenses cost effective and flexible enough for future requirements? Does the provider's roadmap or vision make sense to you?

Creating a Technical Roadmap

Once a short list of vendors has been created, an overall architecture needs to be worked out. The end to end BI system needs to be considered, from the source data to the business user client interface. The architecture will be influenced by the choice of software vendor, but it should be considered somewhat independently as well. Consider the following questions: If you had to change software or hardware vendors, would your architecture have to be completely overhauled and would that type of change be even possible? What infrastructure do you need in place to support the requirements? What is the most cost effective way to grow to meet future demands?

Defining Programs, Projects, and Timelines

A program is a series of projects. Defining a program helps sustain project momentum and executive interest. Without a program, there is a temptation to pack too much into a single project, and expectations rise considerably.

Reduce the risk by defining a program which contains a series of related projects. Each project is created with well defined goals and timelines, such as developing a sales pipeline dashboard for the sales executives by end of quarter. The overall BI program can have a vague ending date, with strategic objectives closely tied to the business, such as increased revenue per customer or better profit margins per transaction. This is an effective way to reduce the project risk, as even if one project is deemed a failure, the program itself can continue to be a success. It also keeps the team focused on the overall business objectives.

Evaluating Staff and Consultants

BI skills are unique. The project management team must be able to make an objective assessment of the human resources available for the tasks at hand. How much technical training will be required before work even begins? Does it makes sense to bring in consultants with product expertise or specific design skills for all or part of the project? What pieces of work can be outsourced to save time and money?

Executing a Change Management Program

An effective change management program not only trains and engages the business users, it also serves to set expectations. Engaging the business users and their management early and often builds trust and diffuses the potential for the "big bust" rollout. The goal is maximum user adoption, without which the BI program will fizzle out as the software turns into shelfware.

Measuring the Benefits

The system has to roll out at an opportune time when the users are not too preoccupied with other issues. Early user interest is critical to success, so avoid conflicting with business deadlines or busy periods, such as end of quarter or end of fiscal year. Collect statistics around usage and benefits to calculate an accurate ROI. It is important to also assess and value the non-tangible benefits.

Automation

Once a project is in production, the technical team needs to start work on any automation that can reduce the administrative and maintenance overhead. Reports, cubes (http://en.wikipedia.org/wiki/OLAP_cube) and other objects should have "push button" scripts to migrate them from development to testing to the production servers. User provisioning, temporary file cleanup, and other regular tasks need to be documented and made as simple as possible. This is especially true if the ongoing support and maintenance will be handed off to another team or eventually outsourced to another organization.

Support

Have a great support team in place, as nothing reduces credibility faster with business users than an unreliable system. With BI there are two categories of support to consider. The first is infrastructure support, which includes ensuring connectivity to source databases, making sure that the reports are published in a timely fashion, and maintaining reliable client connectivity. The second set of issues is unique to BI. If the numbers in the reports or on the dashboards appear to be wrong, then the business users need to have confidence in the support team's ability to investigate the issue and quickly come to an unambiguous conclusion. Either the numbers are wrong because of a technical or data issue, and the problem will be fixed, or the numbers are correct, and the business users must have enough faith in the system to be able to accept that conclusion. Are the current IT practices sufficient to deal with both types of issues? Do extra measures need to be taken in order to build trust between the BI consumers and the technical team? Does the support team have enough understanding of the business to be able to troubleshoot problems with business logic and calculations?

Support also includes good documentation. The vendor's documentation will never be sufficient. At a minimum, start writing a frequently asked questions (FAQs) document. Documentation needs to be specific to the data. Screen captures and videos (screen casts) are ideal for answering how-to questions.

Congratulating the Team

Every project needs some type of closing ritual that fits with the organization's culture. It is an opportunity to take pride in the finished product and to talk about the lessons learned. Time needs to be allotted for knowledge transfer, especially from outside consultants who will be leaving.

Iteration

An effective BI system will continue to evolve and change with the business. Solicit user feedback. Start work on enhancements to improve the system. Deliver more functionality to more staff. Get feedback on what is helping and what is just noise.

Conclusion

Independent of methodology, technology, project scope, or size of the organization, a BI project is different than most other IT projects. Tight collaboration and cooperation is required between technology and business teams. With extra attention paid to some key areas, the additional risks can be mitigated, and success will be easier to come by.

David Currie is the founder of Clearview Informatics (http://www.clearviewinformatics.com), a software startup focused on BI solutions and analytics for customer data. He has worked as an independent BI consultant, providing technical and project management advice to both private sector and government organizations. Previously, he was employed by IBM Cognos, advising strategic customers and troubleshooting enterprise BI implementations gone bad.

Recommended Resources

Business Intelligence ROI example
http://www.bettermanagement.com/images/library/presentations/10/BI_roi/sld001.htm

Scott Berkun's Essays and Blog
http://www.scottberkun.com/

Herding Cats blog
http://herdingcats.typepad.com/my_weblog/

Ralph Kimball's Data Warehousing Articles
http://www.ralphkimball.com/html/articles.html

Tod Means Fox
http://blog.todmeansfox.com/

"...free has become the default, and pay walls the route to obscurity."

Chris Anderson
Free: The Future of a Radical Price

The emergence of open source software (OSS) has changed the buying patterns and approaches for the evaluation and selection of business intelligence (BI) tools. This article discusses how the buying patterns have changed and what it means for businesses looking at open source BI software.

Introduction

The adoption of OSS by businesses has been gaining traction for a number of years now, starting originally with the Linux operating system. OSS has not only shaken up the commercial software vendors, but it has also had a profound effect on the way business evaluates and purchases software.

The evaluation of commercial enterprise-use software tends to be a long and time-consuming process. While we concentrate on the evaluation of BI solutions, the process is similar for other product offerings and typically includes the following stages:

1. Identify the players: either by Internet search, referrals, or using the services of a research and advisory company such as Gartner (http://www.gartner.com) or Forrester Research (http://www.forrester.com). Once identified, the websites of the preferred software vendors have to be reviewed in order to contact each vendor for further information. Reaching out often involves the purchaser to deal directly with the vendor's sales team.

2. Negotiate a trial license: in order to evaluate the product. This often involves, usually at the insistence of the vendor, an onsite demonstration and sales pitch.

3. Compare the capabilities: of the products selected for evaluation against a list of requirements.

4. Determine the preferred vendor: depending upon the size of the software purchase, this may require vendor debrief sessions with the unsuccessful vendors.

5. Purchase negotiations: this can be the most time consuming part of the selection process, depending upon how the vendor prices their offering. Pricing can vary by site, cpu, named user, user count, or other restrictions. Typically, the annual maintenance fee is non-negotiable. However, the upfront license fee is often discounted.

The evaluation process for OSS is similar, but differs in several respects:

With OSS, the software vendors and organizations that have viable offerings still need to be identified. Finding viable OSS BI vendors can be challenging. OSS community sites like http://Sourceforge.net are massive and difficult to research. However, OSS BI solutions are garnering attention from analyst firms. While not on the latest Gartner Magic Quadrant, both Pentaho (http://www.pentaho.com) and Jaspersoft (http://jaspersoft.com) feature prominently in Gartner's discussions. We usually suggest that organizations look to the partner pages of prominent OSS database vendors like MySQL (http://www.mysql.com) and PostreSQL (http://www.postgresql.org) as they typically only partner with BI firms that have a solid product.

Once the software to evaluate is identified, no trial licenses need to be negotatiated. OSS is freely available, and may typically be found at the publisher's site or an organization like Sourceforge.net.

Typically, some instructions on installing and configuring the software are included. Support may be had via the open forums for the product.

Most OSS BI offerings are available under the GNU licensing schema. This does have some impact upon how you can use the software if you plan on embedding it into something you redistribute. If you intend to use the software as is, and rely exclusively on the community for support, the GNU license will meet your needs. If you intend to obtain third party support, or if the sponsoring vendor offers a commercial version of their software, a review of the licensing agreement will be required.

Once downloaded, the OSS can be installed and evaluated on multiple test systems, possibly including a production or pseudo-production environment. BI tools are usually run against production environments since they are end-user tools used to query existing data.

Once a tool has been selected as meeting the necessary requirements, it can be extended into production. Typically, no debriefing is necessary for projects or vendors representing the unsuccessful OSS offerings.

After determining the level of support that is required for the software, a support vendor can be contacted to arrange for a support agreement. In our experience, the enterprise ready BI tool vendors and sponsors all provide support agreements.

Purchasing Open Source

While the software itself is usually free, there are still costs associated with using an OSS solution. What should be considered when comparing these costs with a proprietary solution?

1. Don't expect to discount the price of support. With a commercial vendor, the vendor has already sunk their effort into developing the product and any discounts they negotiate will be with licensing. However, succesfully negotiating reductions in the cost of the annual maintenance or support is nearly impossible. Expect the same from an open source support provider.

2. Determine what you need for licensing. Typically, with a commercial provider you need to negotiate development and test environments in addition to any production sites. With an OSS provider, you need to identify which environments need 'rapid support' or indemnification and which environments are suitable for using the open source community's resources as the support model.

3. Is the open source solution 'enough'? Many vendors in the OSS BI world, such as Pentaho, JasperSoft, and BIRT/Actuate (http://www.eclipse.org/birt), are turning to a two tier software model. In this model, the open source version is available to all, but a commercial offering has enhanced functionality that makes it easier to use the product. In this case, one should evaluate the open source offering, since the core features will be there. Once a preferred vendor has been identified, move forward on a trial of the commercial offering.

4. Validate the publisher. Some open source solutions are truly community-only efforts. While this can be cost beneficial, there is no commercial support and you are at the peril of the community or your own development team for enhancements to the offering. Community direction is often driven by those who participate most: where money doesn't count, input into the product does. However, for most organizations resources to participate are limited to non-existent.

5. Go without support at your peril. In a production environment, commercial support provides a number of things, including rapid response to issues through a dedicated support desk and indemnification against IP infringement. If you have users or systems that require high availability, support is a must have.

6. With OSS you can go at your own pace. Unlike commercial trials, where a time limit often finds organizations racing against the clock, OSS evaluations are without vendor interaction or time limits.

7. OSS offers an opportunity for the evaluation of BI tools. Since these tools are freely available, they can easily be installed for end user evaluation. You may be able to make a business case for trying tools and keeping the ones that users actually use.

8. Mix and match tools: in the commercial paradigm, there is strong pressure on organizations to select the full suite from a vendor in order to negotiate license discounts. With no upfront licensing fees, OSS has removed the promise of discounts as a driver for single vendor purchasing. Organizations may select the tools of their choice and can utilize best of breed tools rather than a commercial vendors' self-proclaimed 'best practice'.

9. Extend the solution to address missing/unique functionality. If you have a specific need, you can tailor the OSS offering to address it since you have access to the source code. You can also participate actively in the community and get the community to develop the required functionality.

A Methodology for Evaluating and Selecting Open Source

Given all of these options, what approach works best for organizations looking at evaluating OSS BI tools?

Unless there are airtight controls over what employees are allowed to download and install, it is quite likely that some open source tools have already made their way into your organization. Find out what people are using and why. This feedback will point out the tools that have already been unofficially prequalified by the organization. Follow up this information by researching what tools are being used in other organizations.

Ensure that all stakeholders have a chance to review the software. This includes business users as well as the technology team that is called upon to support the toolsets and possibly integrate them into other company systems.

If you find one or more tools that meet your needs, investigate the community support and availability of additional features that might be offered as part of the subscription by a vendor. You may find that community support is fine, especially if you can tolerate slow response to support issues or compromised functionality. However, you might also find that the supported version provides what you need. At this point, and only at this point, will you need to contact the vendor for subscription pricing.

Do take the time to compare pricing to similar commercial offerings. The cost savings offered by OSS can help sway an organization that is reticent about using OSS.

Not everyone will find an open source solution that meets their needs. In this increasing rare situation, the option of utilizing a more traditional commercial offering is still available.

John Kemp is a Principal Consultant with SQL Power Group. John is responsible for the delivery of Business Intelligence and Data Warehousing projects at SQL Power. A 20 year information technology veteran, John previously provided Data Warehouse and Business Intelligence consulting services with KPMG, SAS Institute, Infobright Inc., and as the lead of his own consulting firm.

"The difficulty lies not so much in developing new ideas as in escaping from old ones."

John Maynard Keynes

The following survey provides a list of open source software (OSS) tools used in business intelligence (BI) and data warehousing systems. The tool selection criteria was based on the frequency and currency of the releases and on whether the product has released a stable build which could be used in a production environment. We only present those solutions which have had updates within the past two years. Our study looked at BI tools in the following categories: i) databases; ii) extract/transform/load (ETL, http://en.wikipedia.org/wiki/Etl); iii) master data management; iv) BI reporting tools; and v) data mining. In the case of an open source software bundle that overlaps categories, we divide the software bundle into its separate parts for ease of categorization.

Survey Criteria

In the following tables, we define "Fully Open Source" to mean software solutions that offer all of their source code to the public under an open source license. "Y" is used in this column to indicate the software falls into this category. Solutions which provide software versions containing additional features or functionality without offering that source code freely to the public are listed as "-" in this column. Any features mentioned in the tables refer only to the fully open source version of the software. Organizations that do not provide commercial support for their product, but who point their customers to third-party support solutions, receive a "Y*" in the Commercial Support column of the tables, to indicate that commercial support is available. Our analysis was based on a review of the OSS tools market as of July, 2009.

Database Management Systems

There are a wide variety of database management systems (DBMS, http://en.wikipedia.org/wiki/Dbms) available for a data warehouse solution. Standard relational databases, such as MySQL (http://www.mysql.com) and PostgreSQL (http://www.postgresql), have dominated the open source database market for some time. More recently, new open source database systems have emerged that target the unique requirements of data warehousing. Column-oriented databases such as Infobright (http:/www./infobright.com), LucidDB (http://www.luciddb.org), and MonetDB (http://monetdb.cwi.nl/) increase data reading performance by storing data in columns rather than rows. This feature is important for a data warehouse where read-optimization takes precedence over write-optimization, and where typical queries look at the attributes of a column for a set of records rather than the components of a row.

There have been several developments in the area of distributed-computing such as cloud computing. The distributed database solutions presented below are built on top of the open source Java framework for distributed systems: Hadoop (http://hadoop.apache.org). Hadoop-based databases are best suited for data warehouses which number in the hundreds of millions of rows or greater. These distributed data warehouse systems have typically been used for web analytics of high-traffic websites. Finally, although there are several open source embeddable database solutions, none of them are scalable to the extent required by most data warehouse systems, and for that reason, were excluded from this study.

Table 1: Open Source Database Management

DBMS						
Software	**Company / Sponsor**	**Website**	**DB Type**	**Fully Open Source**	**Commercial Support**	**Last Version Release Date**
Cloudbase	Business.com	cloudbase.sourceforge.net	MapReduce, distributed	Y	-	June 2009
Firebird	Firebird Foundation / Borland	www.firebirdsql.org	Relational	Y	Y*	April 2009
HBase	Apache	hadoop.apache.org/hbase	Column-oriented, distributed	Y	-	May 2009
Hive	Apache	hadoop.apache.org/hive	MapReduce, distributed	Y	Y*	April 2009
Hypertable	Doug Judd / Zvents	www.hypertable.org	Column-oriented, distributed	Y	-	May 2009
Infobright	Infobright	www.infobright.com	Column-oriented	-	Y	June 2009
Ingres	Ingres	www.ingres.com	Relational	-	Y	November 2008
LucidDB	LucidEra / Eigenbase	www.luciddb.org	Column-oriented	Y	-	May 2009
MonetDB	Centrum Wiskunde & Informatica	monetdb.cwi.nl	Column-oriented	Y	Y	July 2009
MySQL	Sun / Oracle	www.mysql.com	Relational	-	Y	April 2009
Palo OLAP Server	Jedox	www.jedox.com	Multidimensional	-	Y	March 2009
PostgreSQL	PostgreSQL	www.postgresql.org	Relational	Y	Y*	June 2009

Table 1 summarizes the database types, commercial support availability, and last release version of open source database management systems.

Extract, Transform, and Load Tools

ETL is the most time consuming component of the data warehouse development lifecycle. Typically for this reason alone, it is important to have a good tool which: i) allows the user to interact with as many different source systems as possible; ii) maintains an acceptable level of usability and performance; and iii) provides metadata regarding the transformations undertaken against the data. There are several open source ETL tools available today that provide good performance, good error handling, and metadata management. These tools are summarized in Table 2.

Master Data Management

The datasets in a BI environment are often sourced from many different source systems ranging from spreadsheets to enterprise resource planning (ERP, http://en.wikipedia.org/wiki/Enterprise_resource_planning) applications to third party data. The data may have the same key dimension data repeated in more than one source. Master data management (MDM, http://en.wikipedia.org/wiki/Master_Data_Management) tools are used to identify duplicate records and consolidate them into one unified record which is then stored in the data warehouse.

The same customer could be viewed as a potential lead by the marketing department and as an existing client by the billing department.

Table 2: Open Source ETL Tools

ETL					
Software	**Company / Sponsor**	**Website**	**Fully Open Source**	**Commercial Support**	**Last Version Release Date**
Apatar	Apatar	www.apatar.com	Y	Y	April 2009
Clover.ETL	OpenSYS	www.cloveretl.com	-	Y	July 2009
KETL	Kinetic Networks	www.ketl.org	Y	Y*	October 2008
Open Data Integrator Project	Sun / Oracle	open-dm-di.dev.java.net	Y	-	September 2008
Palo ETL Server	Jedox	www.jedox.com	-	Y	December 2008
Pentaho Data Integration (Kettle)	Pentaho	kettle.pentaho.org	-	Y	May 2009
Talend Open Studio	Talend	www.talend.com	-	Y	May 2009

Table 3: Open Source MDM Tools

MDM					
Software	**Company / Sponsor**	**Website**	**Fully Open Source**	**Commercial Support**	**Last Version Release Date**
DataCleaner	Eobjects	datacleaner.eobjects.org	Y	Y*	April 2009
DQ Guru	SQL Power Group	www.sqlpower.ca/page/dqguru	Y	Y	April 2009
Mural	Sun / Oracle	mural.dev.java.net	Y	-	September 2008
Open Source Data Quality and Profiling project	Arrah	sourceforge.net/projects/dataquality	Y	-	July 2008

Potentially, this customer could receive service and marketing material to encourage the purchase of a service already being consumed. Data warehouse systems consolidate these two customer records into one through the use of MDM tools, ensuring that the customer information and status is consistent across the business system.

Table 3 summarizes the currently available open source MDM tools.

BI Reporting and Analytics Tools

BI reporting and analytics tools are what business users typically use to access data warehouse data, so it is imperative that the tool be intuitive and user-friendly. There are several open source solutions which deliver a variety of features targeting different groups of users such as report developers, decision makers, and non-specialized users. It is important to explore as many tools as possible before deciding on which software package is right for your organization.

Table 4: Open Source Reporting and Analytics Tools

BI Reporting & Analytics Tools									
Software	**Company / Sponsor**	**Website**	Standard Reporting	Ad Hoc Reporting	OLAP	Dash board	**Fully Open Source**	**Commercial Support**	**Last Version Release Date**
BIRT	Eclipse Foundation / Actuate	www.eclipse.org/birt	Y	Y	Y	-	-	Y	June 2009
DataVision	Jim Menard	datavision.sourceforge.net	Y	Y	-	-	Y	-	July 2008
Jaspersoft BI Suite	Jaspersoft	www.jasperforge.org	Y	Y	Y	-	-	Y	June 2009
OpenI	OpenI	www.openi.org	Y	Y	Y	Y	Y	Y	June 2009
OpenReports	Open Source Software Solutions	www.oreports.com	Y	Y	Y	-	-	Y	May 2009
OpenRPT	xTuple	www.xtuple.com/openrpt	Y	Y	-	-	-	Y	January 2009
Palo Worksheet	Jedox	www.jedox.com	Y	Y	Y	Y	-	Y	March 2009
Pentaho BI Platform	Pentaho	community.pentaho.com	Y	Y	Y	Y	-	Y	May 2009
Wabit	SQL Power Group	www.sqlpower.ca/page/wabit	Y	Y	Y	Y	Y	Y	July 2009

In Table 4, we focus on four common features found in BI reporting software: standard reporting, ad hoc reporting, OLAP reporting, and dashboards.

Standard reporting tools can address a variety of different business reporting requirements such as monthly sales by store or by product. Ad-hoc reporting allows non-technical business users to develop 'on-the-fly' reports without having any knowledge of structured query language (SQL, http://en.wikipedia.org/wiki/Sql) or the underlying database structure. OLAP reporting refers to online analytical processing, a technique specifically designed to improve performance of BI queries by pre-aggregating data and common analytical tasks, and by modeling the data multi-dimensionally. The benefit of having OLAP functionality is the ease with which users can drill-down and roll-up into different levels of data hierarchies such as by country, to by province, to by region, to by store.

Dashboards are user-defined visual summaries of key business performance indicators, allowing the user to interact with the charts and graphs to drill-down into specific business areas.

Table 4 summarizes open source reporting and analytics tools, including which features are available for each solution.

Data Mining Tools

Data mining techniques are used in BI systems to determine patterns and relationships amongst data attributes and potential outcomes. Data warehouses are the most common source of data for mining BI because they are typically the most comprehensive and extensive data source available within the company, and are well suited for time-series analysis. Data mining has the potential of discovering correlations that are hidden within the sea of data.

Table 5: Open Source Data Mining Tools

Data Mining Tools					
Software	**Company / Sponsor**	**Website**	**Fully Open Source**	**Commercial Support**	**Last Version Release Date**
Knime	University of Konstanz	www.knime.org	Y	Y	May 2009
Pentaho Data Mining (Weka)	Pentaho	weka.pentaho.org	-	Y	June 2009
R Project	R Development Core Team	www.r-project.org	Y	Y*	June 2009
RapidMiner (YALE)	Rapid-i	www.rapidminer.com	-	Y	March 2009

For example, by using data mining techniques, a company with sufficient customer data could determine which contacts, based on their attributes, such as age, location, and marital status, are more likely to respond to a direct mail marketing campaign. In turn, this could help the company reduce costs by targeting only those customers who are the most likely to respond to the advertisement.

Table 5 summarizes the currently available open source data mining tools.

Summary

Up until recently, the BI market has been dominated by large proprietary software vendors. The maturation period for open source BI solutions has passed, and many of these companies and communities are now directly competing for market shares in several BI categories.

BI consumers are now given a variety of enterprise-ready open source BI solutions to choose from to meet their business requirements, making their software decisions that much more compelling. If the present ubiquity of open source databases can be used as an indicator, then the future looks very bright for open source BI solution providers, and for BI users the world over.

Steve Holub is a Business Intelligence Consultant with SQL Power Group. Steve graduated from McMaster University with a Master's degree in Economics, and received Database Application Developer training from Seneca@York.

"Any fool can make things bigger, more complex, and more violent. It takes a touch of genius-and a lot of courage-to move in the opposite direction."

Albert Einstein

The main purpose of a business intelligence (BI) system is to provide access to information for decision makers. Reporting tools help to create, administer and distribute reports. Reports can be classified into three categories: i) standard reports containing information in table form; ii) cubes which are online analytical processing (OLAP, http://en.wikipedia.org/wiki/Olap) reports designed to analyze multidimensional data; and iii) dashboards which provide an overview of selected reports and key performance indicators (KPI, http://en.wikipedia.org/wiki/Key_performance_indicator). Each of these report types are often supplemented by graphical representations of the data.

There are many different reporting solutions available in the open source market. In this article, we compare four open source reporting solutions. They give a good insight of the current state of the market and are among the most popular solutions in use today.

Four Reporting Solutions

This section provides a brief overview of the four reporting solutions we compared.

BIRT (http://www.eclipse.org/birt) is a popular reporting solution hosted by the Eclipse Foundation. The BIRT software project started in 2004 when Actuate Corporation (http://actuate.com) joined the Eclipse Foundation and donated the software.

Based in Orlando, Florida, Pentaho (http://www.pentaho.com) is one of the most recognized open source BI vendors. Their flagship product, Open Source Business Intelligence Suite, includes reporting, analysis, dashboards and data mining features. This comparison focuses on the reporting capabilities of the open source version which has fewer features than the enterprise edition.

Founded in 2001, Jaspersoft (http://www.jaspersoft.com) is an established vendor of open source BI solutions. Jaspersoft's Business Intelligence Suite offers a feature set comparable to Pentaho's solution. While an enterprise addition is available, this review analyzes the features of their open source community edition.

Headquartered in Toronto, Canada, SQL Power Group (http://www.sqlpower.ca) open-sourced its reporting tool Wabit in 2008. Wabit offers reporting and OLAP analysis. Unlike most other open source BI reporting tools, Wabit comes with all existing features in its open source version.

Table 1 summarizes the four solutions and their most recent versions.

Comparison Criteria

The four solutions were compared using the following criteria:

1. Getting started: the software should be easy to install and to use. For the ease of installation criteria, we tested how many different files had to be downloaded and how many different programs were installed in order to create and view reports. We also determined whether sufficient examples are available, and whether prior knowledge of SQL is required to create reports.

Table 1: Open Source Reporting Solutions

Product	Supported by	Source	Latest version
BIRT	Actuate & Eclipse Foundation	www.eclipse.org/birt/phoenix	2.5 (June 2009)
Jaspersoft Business Intelligence Suite	Jaspersoft	www.jaspersoft.com	3.5 (April 2009)
Pentaho Business Intelligence Suite	Pentaho	http://www.pentaho.com	3.0 CE (May 2009)
Wabit	SQL Power Group	http://www.sqlpower.ca	0.9.7 (July 2009)

Table 2: Feature Comparison of Solutions

Product	BIRT	Jaspersoft Business Intelligence Suite	Pentaho Business Intelligence Suite	Wabit
Getting started				
Ease of Installation	★★★★☆	★★★☆☆	★★☆☆☆	★★★★★
Example reports available	Y	Y	Y	Y
SQL knowledge required	Y	-	-	-
Training offered	Y	Y	Y	-
Features				
Ease of use reporting	★★★☆☆	★★☆☆☆	★★☆☆☆	★★★★★
Data sources: JDBC / XML/A / Flat file	Y / - / Y	Y / Y / Y	Y / Y / -	Y / Y / -
Multiple data sources in one report	Y	Y	Y	Y
Output formats: PDF / HTML / CSV	Y / Y / Y	Y / Y / Y	Y / Y / Y	Y / Y / Y
Report templates	Y	Y	Y	-
Number of charts available	13	20	14	4
Images can be added	Y	Y	Y	Y
Report parameters	Y	Y	Y	-
Schedule reports	Y	Y	Y	-
Send reports by email	-	Y	Y	-
Client report application	Y	Y	Y	Y
Web based report creation	Y	Y	Y	-
OLAP reports	-	Y	Y	Y
Dashboards	-	-	Y	Y
Dashboard allows drill down to details	-	-	Y	-
User can create own dashboards	-	-	-	Y
Security	(Y)	Y	Y	-
Support				
Frequent posts in forum	Y	Y	Y	Y
Documentation	Y	Y	Y	Y
Paid support	Y	Y	Y	Y

2. Feature richness: compares the main product features in order to determine whether the software fulfills the most common reporting requirements. The ease of use criteria indicates how much effort is needed to create and view reports. It is also important to see which data sources can be used and which output formats are available.

3. Available support: an often heard prejudice against open source software is the lack of documentation and support. This category contains information on how well supported the reporting solution is based on its documentation and forum activity.

Table 2 summarizes the results of our survey. A star rating is used to rate each criteria as described above. "Y" indicates a feature is present, "(Y)" indicates a feature is partially present, and "-" indicates the feature is not present in the product.

Conclusions

All four tested reporting tools provide the features needed to create basic reports. BIRT is easy to get started with and is well supported, but is limited to relational reports since it does not support OLAP reports and dashboards. BIRT doesn't have its own security but depends on the security of its underlying application server. Nonetheless, BIRT has seen widespread use, especially embedded in other open source applications.

Although Jaspersoft's BI suite report development environment is more complex, it compensates for this fact with added features and functionality. Jaspersoft allows OLAP analysis but its current open source version doesn't support dashboards. Its report designer gives the user many options to create reports but the web interface lacks usability and functionality.

The most complex product in this comparison is Pentaho's BI suite which includes several client applications for different tasks. The effort users expend getting started pays itself back in a feature-rich BI solution with a state of the art web interface expected by medium to larger businesses.

Wabit takes a different approach than all of the other tested solutions. It provides all the functionality to create and view reports in one application, making it easy for the first time user. Its current version is missing some features such as a web interface, security and parameterized reports. But its ease of use makes this reporting solution a compelling alternative for small to medium size organizations. SQL Power has to add the missing features to compete with Pentaho and Jaspersoft.

Although none of the presented solutions can fulfill all criteria, the comparison shows that open source BI solutions are mature enough to be used in production environments.

Benjamin Dietz is a Business Intelligence Consultant at SQL Power Group. Benjamin wrote his masters thesis on Open Source Business Intelligence and graduated from University of Applied Science Karlsruhe, Germany with a Master of Business Information Systems.

Lily Singh is a Business Intelligence consultant at SQL Power. A graduate form the University of Waterloo, Lily has been a consultant with SQL Power for the past 2 years. In this time she has contributed to different projects ranging from Data Migration, Data Cleansing to building complete Data Warehouses.

"About eighty percent of all data stored in corporate databases has a spatial component."

Carl Franklin

Recently, interest in the huge potential of Geospatial BI has increased. It aims at combining geographic information system (GIS, http://en.wikipedia.org/wiki/Geographic_information_system) and business intelligence (BI) technologies. Geospatial BI combines spatial analysis and map visualization with proven BI tools in order to better support the corporate data analysis process and to help companies make more informed decisions.

BI is a business management term which refers to applications and technologies that are used to gather, provide access to, and analyze data and information about company operations. BI applications are usually used to better understand historical, current and future aspects of business operations. BI applications typically offer ways to mine database- and spreadsheet-centric data to produce graphical, table-based and other types of analytics regarding business operations. BI systems give companies a more comprehensive knowledge of the factors affecting their business, such as metrics on sales, production, and internal operations, in order to to make better business decisions.

This article provides a rapid introduction to some important BI concepts. It then highlights the need for geospatial BI software and deals with the integration of the spatial component in a BI software stack in order to consistently enable geo-analytical tools. We then present different works performed and tools designed by the GeoSOA research group.

A Rapid Introduction to BI

BI applications rely on a complex architecture of software that is usually composed of:

- An extract/transform/load (ETL, http://en.wikipedia.org/wiki/Etl) tool to extract data from different heterogeneous sources, provide integration and data cleansing according to a target schema or data structure, and load the data in a data warehouse.

- A data warehouse which stores the organization's historical data for analysis purposes.

- An online analytical processing (OLAP, http://en.wikipedia.org/wiki/Olap) server which enables the rapid and flexible exploration and analysis of the large amount of data stored in the data warehouse.

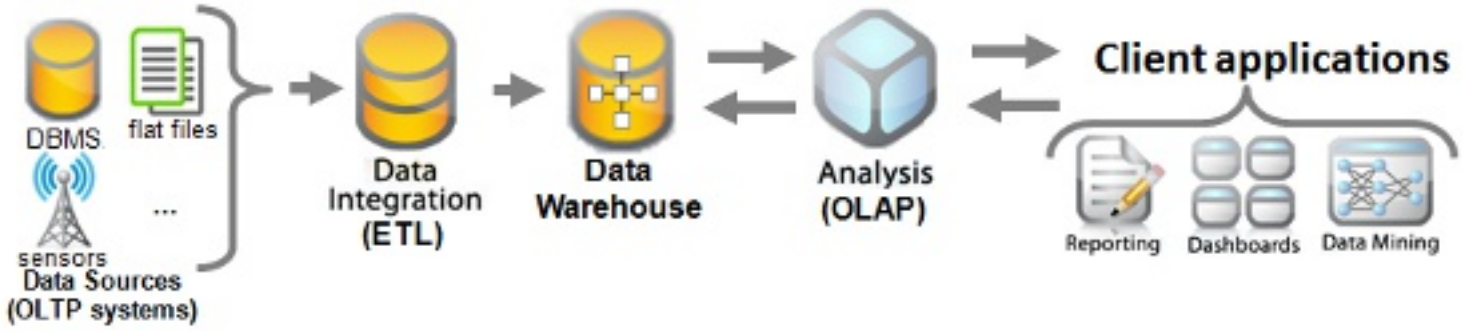
- On the client side, some reporting tools, dashboards and/or different OLAP clients to display information in a graphical and summarized form to decision makers and managers. These tools offer capabilities to explore data interactively and support the analysis process.

- Optionally, some data mining tools to automatically retrieve trends, patterns and phenomena in the data.

Figure 1 illustrates the typical infrastructure on which BI applications rely.

The data warehouse plays a central and crucial role in this architecture. It is the repository of an organization's historical data. It is separate from operational data sources but is often stored in relational database management systems.

Figure 1: Classical Architecture for Deploying BI



Data warehouses are optimized for handling large volumes of data, providing fast response during the analysis process, and handling complex analytical queries. They rely on de-normalized data schemas which introduce some redundancy to provide very fast replies to time consuming queries involved in analytical requests.

A data warehouse focuses more on the analysis and the correlation of large amounts of data than on retrieving or updating a precise set of data. This is fundamentally different to the functions of the transactional database systems used in the day to day activities of a company.

Contents of the data warehouse are often presented in a summarized form primarily for analysts and decision makers. Figure 2 (http://www.osbr.ca/ojs/september09/tbadard-figure_2.png) illustrates different tools from Pentaho (http://www.pentaho.com) used to present, explore and analyse data.

To query the data warehouse, these tools generally use the MultiDimensional eXpressions (MDX, http://en.wikipedia.org/wiki/Multidimensional_Expressions) query language implemented by the OLAP server. MDX is a de facto standard from Microsoft which is also implemented by other OLAP servers and clients.

MDX is for OLAP data cubes what the structured query language (SQL, http://en.wikipedia.org/wiki/Sql) is for relational databases. Queries are similar to SQL but rely on a model closer to the one used in spreadsheets.

OLAP client software propose alternate representation modes, such as pie charts and diagrams, and different tools to refine queries and to explore data. These tools are based on operators provided by the MDX query language and on a complex logic implemented in the client. The spatial component of data can be used to enhance the BI user experience with map displays and spatial analysis tools to better support the analysis and decision processes.

Merging BI and GIS Software

It is difficult for a decision maker to answer complex questions like: where are the urban spots that are more sensitive to heat waves, intense rain, flooding or droughts in a specific geographic area? How many people with cardiovascular, respiratory, neurological and psychological diseases will there be in 2025 and 2050 in a specific geographic area? How many people with low income live alone in a building requiring major repairs in a specific geographic area?

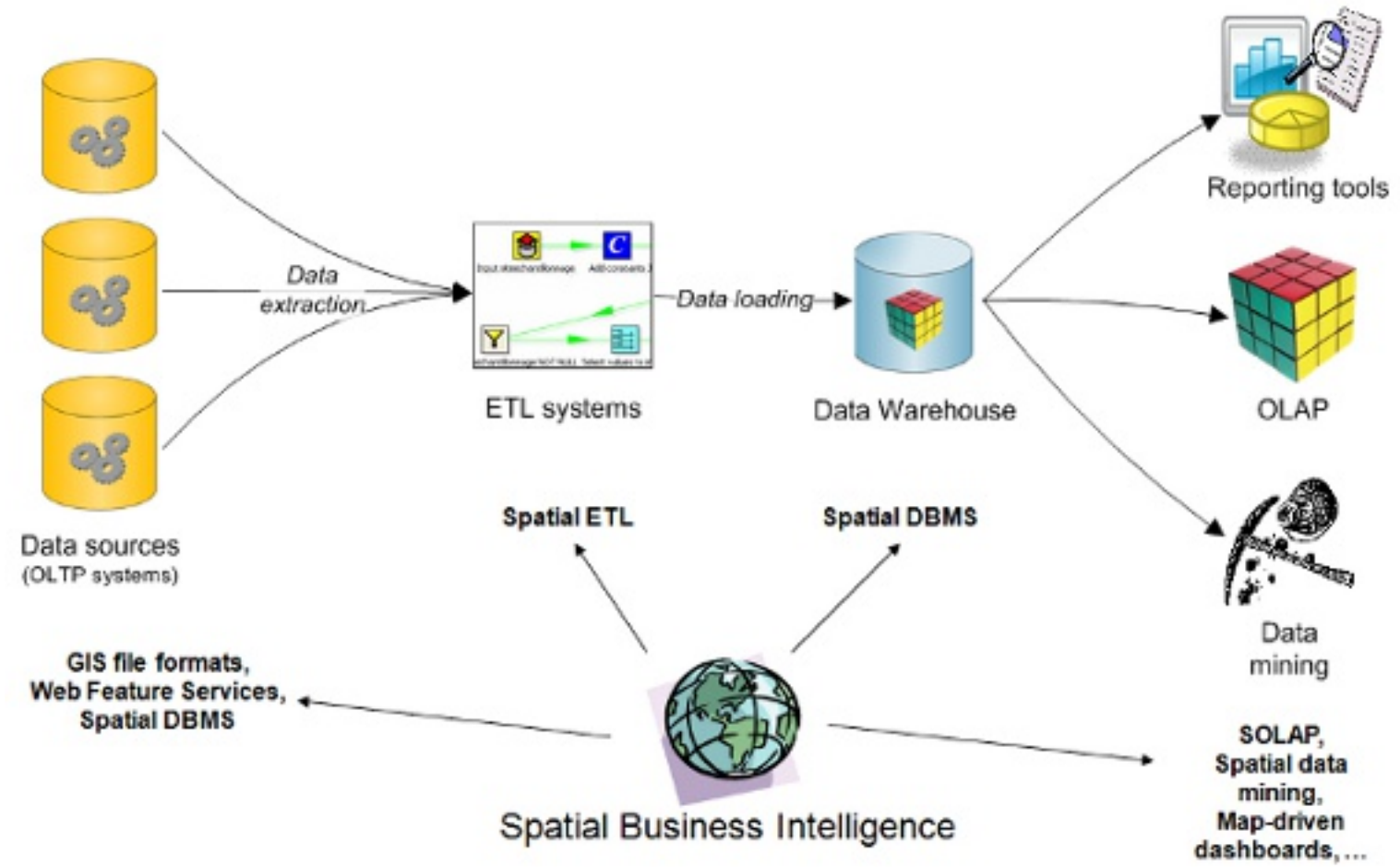
To answer such questions, you can use:

1. **GIS:** implies writing very complex SQL queries and dedicated human resources. Moreover, this job needs to be done anew every time data change or new analyses have to be achieved.

2. **Classical BI tools:** are often unable to handle the spatial dimension of data or only provide a very basic support. Some phenomena can only be adequately observed and interpreted by representing them on a map. This is especially true when you want to observe the spatial distribution of a phenomenon or its spatiotemporal evolution.

Geospatial BI has recently stirred marked interest for the huge potential of combining spatial analysis and map visualization with proven BI tools. Tools recently made available on the market rely on a loose coupling between existing GIS software and some proven BI components. They provide first solutions to display maps with summarized and aggregated information stemming from the BI infrastructure while GIS data have to be stored and managed in a separate and transactional database system or GIS data file. These solutions manage geospatial and corporate data in different systems which require additional efforts, resources and costs to consistently feed and maintain them. They also do not fully take advantage of the powerful analytical capabilities of a classical BI infrastructure and usually are not able to handle very large data volumes. This loose coupling often requires the development of dedicated applications each time a new analytical need emerges in the company.

Figure 3: Integrating Spatial Component into a Classical BI Infrastructure



The geometry data type on which geospatial data relies is not handled as any other data type in the BI infrastructure and connections with the GIS have to be carefully initiated and maintained. Drill down and roll-up capabilities in the analytical data to observe data at different levels of detail, time or scale are often not supported by the map display because they are not intrinsic operators available in GIS. This is mainly due to the transactional structure of geospatial data in the underlying GIS software. Dimensional data structures on which BI tools rely are more efficient to quickly reply to complex analytical queries which would have involved numerous time consuming join queries in a transactional system.

Consistently integrating the geospatial component in all parts of the BI architecture is required. Figure 3 illustrates that all components of the BI infrastructure have to be spatially-enabled.

Some spatial capabilities such as support for reading and writing GIS file formats, coordinate transformations, and spatial reference systems need to be injected into ETL tools. OLAP servers should be extended to become actual Spatial On-Line Analytical Processing (SOLAP) servers. SOLAP should bring the consistent handling of geospatial features, map displays and spatial analysis capabilities. SOLAP servers and clients should "allow a rapid and easy navigation within spatial data warehouses and offer many levels of information granularity, many themes, many epochs and many display modes of information that are synchronized or not: maps, tables and diagrams" (http://geosoa.scg.ulaval.ca/~badard/download.php?url=ogrs2009-towards_mobile_solap_infrastructure-tbadard_et_edube-final.pdf).

In this perspective and in order to not reinvent the wheel, the GeoSOA Research Group (http://geosoa.scg.ulaval.ca) at Laval University, Quebec, Canada started to consistently and completely integrate geospatial functionalities into an existing, mature, efficient and reputed open source BI software stack.

A complete open source BI software stack is offered by Pentaho (http://www.pentaho.com). It includes:

- an ETL tool to integrate data from heterogeneous sources to a data warehouse
- an OLAP server which provides multidimensional query facilities on top of the data warehouse
- reporting and dashboard tools, used to present data to analysts

The integration of the Pentaho software suite with open source GIS components has been investigated to create a complete spatially-enabled BI solution. This work has led to the implementation of GeoKettle (http://www.geokettle.org), GeoMondrian (http://geo-mondrian.org) and SOLAPLayers (http://www.solaplayers.org).

GeoKettle

GeoKettle is a spatially-enabled version of Pentaho Data Integration (PDI, http://kettle.pentaho.org), formerly known as Kettle. It is a powerful, metadata-driven spatial ETL tool dedicated to the integration of different spatial data sources for building and updating geospatial data warehouses. GeoKettle enables the transparent handling of the geometry data type as any other classical data type to all transformations available in Kettle. It is possible to access geometry objects in JavaScript and to define custom transformation steps. Topological predicates have all been implemented.

GeoKettle has been released under the LGPL (http://opensource.org/licenses/lgpl-2.1.php). Figure 4 (http://www.osbr.ca/ojs/september09/tbadard-figure_4.png) illustrates the GeoKettle user interface showing a basic geospatial data transformation.

At present, Oracle spatial, PostGIS (http://postgis.refractions.net/), and MySQL with ESRI shapefiles are natively supported in read and write modes. At present, Microsoft SQL Server 2008, Ingres (http://www.ingres.com), and IBM DB2 can be used with some modification. It is possible to build and feed complex and very large geospatial data warehouses with GeoKettle. Spatial reference systems management and coordinate transformations have been fully implemented. Native support for unsupported geospatial databases and raster and vector based data formats will be implemented in the near future as an active and growing community has federated around the project.

GeoKettle releases are aligned with PDI, allowing GeoKettle to benefit from all the new features provided by PDI. For instance, Kettle is natively designed to be deployed in cluster and web service environments. This makes GeoKettle suitable for deployment as a service in cloud computing environments. It enables the scalable, distributed and on demand processing of large and complex volumes of geospatial data in minutes for critical applications, without requiring a company to invest in an expensive infrastructure of servers, networks and software.

Upcoming features to be implemented in GeoKettle include:

- cartographic preview
- implementation of data matching steps to allow geometric data cleansing and comparison of geospatial datasets
- read/write support for other database, GIS file formats and geospatial web services
- native support for MS SQL Server 2008 and Ingres
- implementation of a spatial analysis step through a graphical interface

GeoMondrian

GeoMondrian is a spatially-enabled version of Pentaho Analysis Services (Mondrian, http://mondrian.pentaho.org). It has been released under the EPL (http://www.opensource.org/licenses/eclipse-1.0.php).

As far as we know, GeoMondrian is the first implementation of a true SOLAP server. It provides a consistent integration of spatial objects into the OLAP data cube structure, instead of fetching them from a separate spatial database, web service or GIS file. To make a simple analogy, GeoMondrian brings to the Mondrian OLAP server what PostGIS brings to the PostgreSQL database management system. It implements a native geometry data type and provides spatial extensions to the MDX query language, allowing embedding spatial analysis capabilities into analytical queries.

These geospatial extensions to the MDX query language provide many more possibilities, such as:

- in-line geometry constructors
- member filters based on topological predicates
- spatial calculated members and measures
- calculations based on scalar attributes derived from spatial features

At present, GeoMondrian only supports PostGIS based data warehouses but other databases should be supported soon.

SOLAPLayers

Formerly known as Spatialytics, SOLAP Layers is a lightweight web cartographic component which enables navigation in SOLAP data cubes. It aims to be integrated into existing dashboard frameworks in order to produce interactive geo-analytical dashboards. The first version of SOLAPLayers stems from a Google Summer of Code (GSoC) 2008 project performed under the umbrella of OSGeo (http://osgeo.org). The client is released under the BSD (http://opensource.org/licenses/bsd-license.php) license and the server under the EPL.

SOLAPLayers is based on the OpenLayers (http://openlayers.org) web mapping client and uses olap4j (http://olap4j.org) for connection to OLAP data sources. For now, it requires GeoMondrian to display members of a geospatial dimension on a map. SOLAPLayers allows the:

- connection with a spatial OLAP server such as GeoMondrian
- navigation in geospatial data cubes
- cartographic representation of some measures and members of a geospatial dimension as static or dynamic choropleth maps (http://en.wikipedia.org/wiki/Choropleth) and proportional symbols

A demo application is available online at http://geosoa.scg.ulaval.ca/Spatialytics. It demonstrates the interaction with GeoMondrian and how the cartographic navigation in the geospatial data cube is performed.

Upcoming features in the development for SOLAPLayers include:

- more map-driven OLAP navigation operators
- dimension member selection and navigation controls
- legend display
- new choropleth and graphics mapping styles
- styles for other geometry types
- multi maps

Conclusion

This article has highlighted the need for geospatial BI software and has emphasized that spatially-enabling a BI software stack requires the consistent integration of the spatial component and its functionalities into each component of the BI infrastructure. Works performed by the GeoSOA research group have led to the release of three open source building blocks of a consistent and powerful geo-BI software stack.

Based on these key software components, future works deal with the design of a geo-analytical dashboard framework. In order to easily design and deliver dashboards which embed some geospatial components and representations, a highly customisable and flexible geo-analytical dashboard framework is required. A first integration of SOLAPLayers with JasperServer (http://www.jaspersoft.com/jasperserver) and iReport (http://www.jaspersoft.com/ireport) has recently been performed in the GeoSOA research group.

The result of this integration allows displaying information in different ways and the synchronisation between the different representations when the user drills down or rolls up on the map or the charts.

More recently, some experiments dealing with the integration of SOLAPLayers into the Pentaho Community Dashboard Framework (CDF, http://code.google.com/p/pentaho-cdf) have been performed in the context of a GSoC 2009 (http://geosoa.scg.ulaval.ca/en/index.php?module=pagemaster&PAGE_user_op=view_page&PAGE_id=20), under the umbrella of OSGeo.

The integration work performed by the student during this period allows the display of the SOLAPLayers cartographic component together with a pivot table component in a CDF dashboard. Synchronisation between the map and the pivot table has been implemented. Further work is required in order to more properly and consistently integrate the SOLAPLayers component into CDF, but it represents a good and promising first step towards the design of a highly customisable and flexible geo-analytical dashboard framework. A live demo of the integration work performed by the student will be available shortly (http://geosoa.scg.ulaval.ca). The source code will also be available in the GSoC 2009 repository.

The reader is invited to consult the presentation about the research challenges dealing with the integration of the spatial component in BI tools and the design of intelligent mobile applications for better decision support (http://geosoa.scg.ulaval.ca/~badard/download.php?url=ogrs2009-towards_mobile_ solap_infrastructure-tbadard_et_edube-final.pdf). These research challenges are currently part of the research agenda of the GeoSOA research group.

This article is a short version of the original paper written for this special issue on Business Intelligence for the OSBR. The full version can be freely downloaded from http://geosoa.scg.ulaval.ca/~badard/download.php?url=article-tbadard-osbr_2009-long_version-enabling_geospatial_bi.pdf.

Dr. Thierry Badard is CTO of Spatialytics, a new company in geospatial BI. He is also a professor of geoinformatics at Laval University (Canada) where he heads the GeoSOA research group. He is a regular researcher of the CRG and of the GEOIDE NCE. He has more than 13 years of experience and he has been involved in national and international R&D projects of importance. He acts as a chair, editor and reviewer for several international journals and scientific conferences. Dr. Thierry Badard is also actively involved in the geospatial free and open source community. He is an OSGeo charter member and a member of the OSGeo conference committee. Member of the board of the OSGeo Francophone chapter, he is also a founding co-chairs the OSGeo Quebec local chapter. He is a founding co-chair of the ICA working group on open source geospatial technologies.

Etienne Dubé is a research assistant in the GeoSOA Research Group, Laval University. He holds a Masters degree in Geomatic Science and a Bachelor degree in Computer Engineering. He is the main developer in the GeoMondrian, SOLAPLayers and GeoKettle projects.

"Design is not just what it looks like and feels like. Design is how it works. "

Steve Jobs

The Business Intelligence and Reporting Tools (BIRT, http://eclipse.org/birt) Project is a highly successful open source business intelligence (BI) project developed and released as part of the Eclipse ecosystem. In this article, we look at the genesis of the BIRT project, why Actuate (http://www.actuate.com) chose the open source development model, and the benefits that this project brings to the BI user community.

Why a New Project?

While some projects are created for the sake of the technology, most projects are instituted and refined to solve a problem. In the case of BIRT, Actuate did not wake up one morning and say, "We want to develop an open source reporting tool." Prior to BIRT, Actuate provided many commercial BI tools. From its market research and prospective customer analysis over fifteen years, Actuate knows that its chief competitor is the home-grown solution developed by information technology staff. Such applications typically use Java Server Page (JSP) and other Java server-based technologies to connect to a database or warehouse to provide end users with data for decision making.

The home-grown development approach is not necessarily cheaper in the long run. It does generally require less initial capital than an investment in an off-the-shelf, vendor-supplied product. Another persuasive advantage to using this approach is the wealth of in-house developer expertise available to construct these internal, Java-based solutions.

While this type of application has many benefits, it is frequently inflexible, difficult to maintain, and even harder to extend. As user requirements evolve over the project lifetime, changes to the application frequently become time-consuming and expensive. Actuate recognized an opportunity to build a product with benefits similar to a home-grown solution that can also mitigate the disadvantages. The BIRT project was created to fit this market segment.

BIRT Overview

BIRT is based on the popular Eclipse integrated development environment (IDE) used to build and deploy BI applications in a Java/J2EE environment. The goal of the BIRT project is to create a set of tools for data analysis and reporting that are simple and easy to use, while supporting extensibility and modularization.

The main tool is a report designer built into the Eclipse platform that generates XML-based report designs. These designs support access to many data sources, which can be joined, cubed, or mined for data. The designer is easy to use, yet offers many advanced features, such as crosstabs, data cubes, and extensive charting capabilities. Figure 1 (http://osbr.ca/ojs/september09/features.PNG) provides screenshots of BIRT.

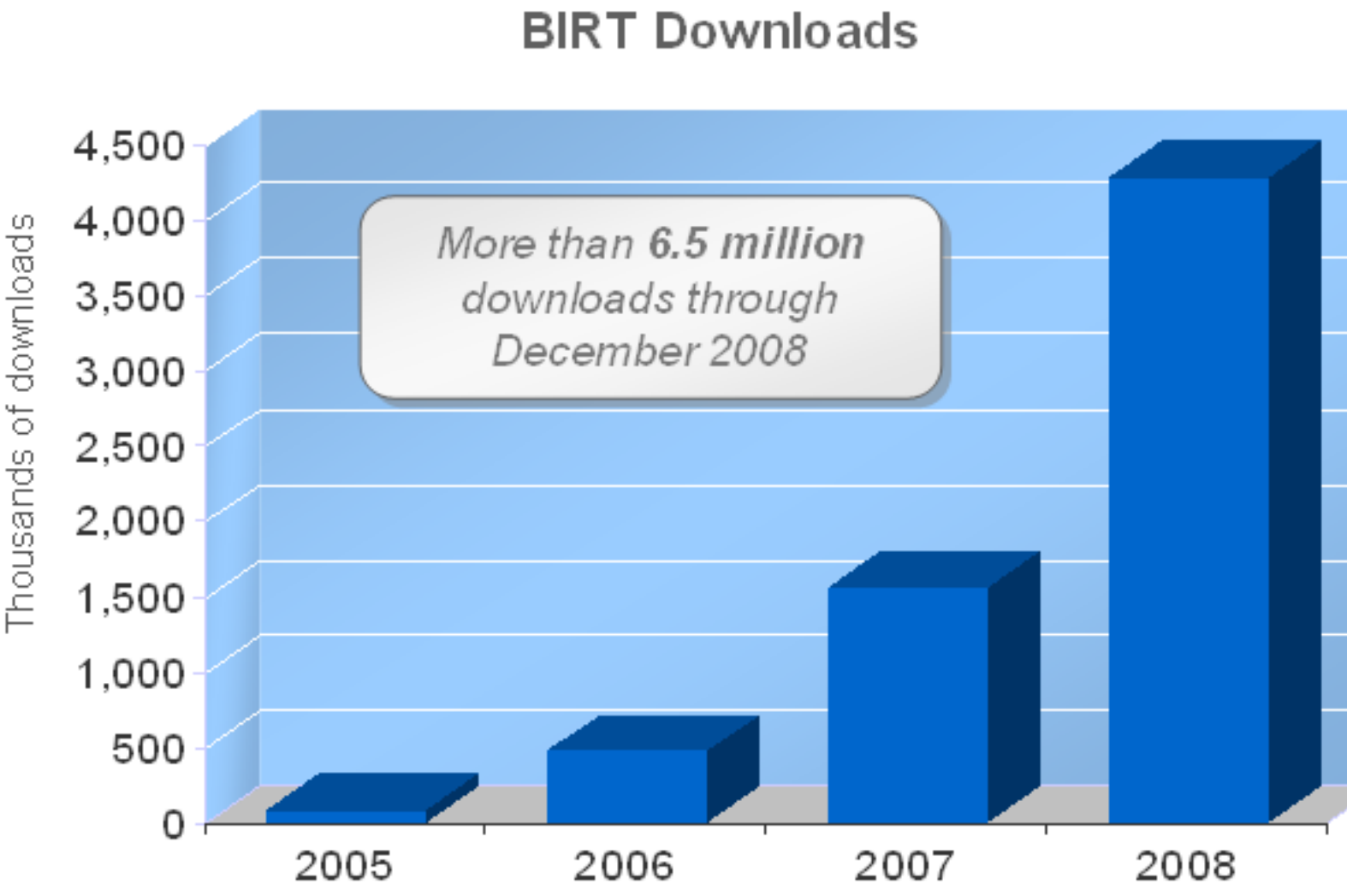
The designer was developed to support a diverse group of developers, from the business analyst to the hard-core Java developer. This rapid-development tool allows many reports to be built in hours instead of days, yet offers the capability to perform advanced customizations and integrations in the existing infrastructure. Upgrading a BIRT-based BI application can be abstracted to the level of an analyst developer.

Once a report is complete, it can be deployed using open source or commercial BIRT technology. The open source BIRT project includes a rich set of APIs and an AJAX-based Java servlet Viewer, either of which can be used for deployment. The servlet-based Viewer application provides compatibility with existing deployment environments and supports multiple integration options. BIRT also offers highly extensible features, providing many ways to expand and enhance the tool, such as adding new custom data sources, output formats, and chart types. These extensions are modular and can be deployed to a BIRT-based environment without changing any system infrastructure components.

The BIRT project had its first major release in the summer of 2005. By the end of 2008, BIRT had garnered over six and one-half million downloads. The BIRT website includes an introduction, tutorials, downloads, and examples of how to use BIRT. Figure 2 shows the exponential growth of BIRT downloads.

In addition to downloads, BIRT has one of the most active newsgroups in the Eclipse ecosystem, with approximately 500 posts per month. The BIRT Exchange (http://birt-exchange.com) website is a major hub for the BIRT community and has seen considerable growth in the two years since its launch. The site has over 19,000 members, receives over 6,000 daily views and more than 32,000 monthly visitors, and processes over 1,200 forum posts per month.

Figure 2: Number of BIRT Downloads



Why Open Source?

Actuate choose to develop BIRT using the open source model as the most effective way to generate awareness and to build a user community. In today's market, it is difficult to launch a new product and achieve visibility with potential users. Open source is a superb vehicle for generating product awareness. Users are much more likely to download and experiment with a readily available, customizable, open source tool than a proprietary product. Prospective customers are more inclined to select a product that they have already used and understand well, rather than one they have not.

Companies seldom produce software for altruistic reasons. Publicly-traded companies must show a profit to survive and fund the developers who create the technology. Selling products and services is the predominate way that a software company makes money. In the open source world, monetizing product development work is a complex task and varying strategies are employed.

Some companies provide an open source, free product that allows users to develop content in a development environment, but offer no deployment options or the deployment model is restricted in some way. This type of product is often perceived as a bait–and-switch tactic by developers, which often alienates them. On the other hand, if a company's monetization strategy depends on extended features, providing many free features can reduce the company's ability to earn financial compensation for a successful product.

Actuate's approach is to have open source BIRT support the core needs of application developers who want to incorporate data-driven content in an application.

A BIRT application must be fully functional, unrestricted, and deployable. Actuate also offers commercial products that deliver value-added improvements in the design and deployment environment to improve productivity, add scalability, and enhance the end-user experience. In this business model, the community is a valuable asset.

The communities that develop around open source and commercial projects can be quite diverse. A successful product attracts a vibrant community of experts, who can provide example content, participate in technical discussions, develop products, and provide services for users. An open source community offers increased product visibility and an extended pool of expertise to support peer-level code reviews, contribution testing, bug reporting, and provide other feedback and information on how to develop and extend the product. Products with low community involvement generally do not succeed or grow at a significantly slower pace.

Actuate understands this fact and has made a substantial investment of staff and other resources to facilitate community growth. Actuate has developed BIRT-Exchange.org to support community interaction, hired and consulted experts with experience in building open source communities, and committed substantial staff resources, including a full-time community manager, to assist users in learning about the project and getting the information and resources they need to be successful using BIRT.

Benefits of Open Source Development Model to Users

A successful, vibrant open source community with a broad knowledge base and deep pool of expertise offers substantial benefits to the users of the software.

Extending beyond the ability to download and alter the source, active community participation leads to improved quality, ease of use, greater extensibility, broader integration, and reduced turnaround time for feature requests.

This scenario has been very true for the BIRT project. With community sites and other projects incorporating the tool, the pool of expertise has grown dramatically in a short period of time, offering substantial benefits to the user community. Prospective users typically have a wealth of information and personnel at their disposal, which substantially reduces the learning curve and shortens development and deployment times.

With open source many developers are looking over the code, which reduces the amount of errors and produces a product with much better quality. Error fixes and feature requests are typically dealt with in a faster manner than a traditionally developed product. This process reduces turn-around time for required functionality, which increases the adoption rate and general satisfaction with the product.

Another benefit of the open source model is the way the product is developed. In order to allow multiple teams to work on a project, modularity and extensibility are not idealized goals, but are the tenets on which the project is built. When extensibility and modularity are implemented correctly, many disparate teams can work on various aspects of the product simultaneously. Not only are more features brought into the product, but end-users have a powerful extension framework for their own personal development and customization needs.

For these reasons, BIRT chose the Eclipse framework. The end result is that the user community has been actively involved in extending the feature set and the BIRT product has been incorporated in many additional products and infrastructures. All of this collective experience has vastly increased community awareness and expertise. Products like BIRT typically work in conjunction with other products and these integration scenarios have become very powerful, binding user communities together, and expanding knowledge bases through cooperative development.

As the community continues to grow, the number of feature requests and implementations submitted increases. The new features to add are typically decided by the community, ultimately producing a product that is much more in line with what prospective users want. Requests are usually directed at key aspects of the product, such as general usability, which when addressed, benefit the entire community. This process decides the product direction in both near and long-term perspectives.

Conclusion

Actuate choose to create the open source Eclipse BIRT project to meet an important, practical business need and has worked successfully to build a strong, active community around this project. While the project continues to evolve, both the community and Actuate have benefitted. These benefits continue to grow exponentially as the user community continues to develop and grow.

The pivotal factor for the success of the project has been the community of developers and users world-wide, who are constantly providing feedback, code, and expertise, which continues to improve and grow the product. Actuate knows that continuing to nurture this community will ensure the continued success of BIRT.

Recommended Resource

Breaking Market Barriers: The Open Source Business Strategy
http://www.eclipse.org/community/casestudies/Actuate_OS_Final.pdf

Tom Bondur is a Content Development Manager in the Developer Communications group of Actuate Engineering and a member of the extended BIRT development team. He has a background in both computer science and technical writing with many years experience in technical consulting, training, writing, and publishing about business intelligence tools and database technologies. He is a co-author of the Eclipse Series book, Integrating and Extending BIRT, published by Addison-Wesley. The book introduces programmers to BIRT architecture and the reporting framework. It is the second volume in a two-book series about business intelligence and reporting technology.

Jason Weathersby is the BIRT Evangelist at Actuate Corporation and a member of the Eclipse Business Intelligence and Reporting Tools (BIRT) Project Management Committee (PMC). Jason has over 15 years experience in the software development field, ranging from real time process control to business intelligence software. At Actuate, Jason is currently responsible for educating the Open Source community on BIRT and encouraging its adoption, and is responsible for managing the Eclipse BIRT newsgroup and website. Jason has co-authored two BIRT books and written many articles that discuss the BIRT technology.

Q. Oracle's acquisition of Sun is quite possibly the most strategically brilliant information technology (IT) move in more than 20 years. What will Oracle do with Sun's open source offerings?

A. Back in the mid 1980's, I worked for Oracle Corporation. Although much has changed over the years in this billion-dollar conglomerate, at least one thing remains the same: CEO and founder Larry Ellison is strategically brilliant.

Back in the 1980's, Ellison made sure that all of Oracle's products were functionally identical on all operating system platforms including Windows, Vax VMS and Unix. In order to penetrate larger accounts, he offered the personal computer (PC) version of Oracle's database and all of its tools for $299. This one low price removed the barriers to entry for smaller organizations and allowed IT groups in larger organizations to build fully functional online transaction processing (OLTP, http://en.wikipedia.org/wiki/OLTP) applications on the Oracle database platform. Soon companies wanted to deploy these prototypes and proof of concept applications, rolling them out to larger VMS and Unix production environments, and paying hundreds of thousands of dollars for the same functionality so that they could run them on these larger platforms.

This brilliant marketing move significantly grew Oracle's database market share and allowed Oracle to double their revenue for 13 straight quarters between 1987 and 1990, growing this once basement startup into one of the largest software companies in the world.

A lot has happened between 1985 and 2009: a couple of recessions, a dot com bust, many key acquisitions, and Oracle is left standing as one of the largest software players in the world.

Larry Ellison is close to pulling off the biggest coup in the last 20 years by acquiring Sun for a bargain basement price of approximately 5.6 billion dollars. This move is positioning Oracle as:

- the premiere open source database player with MySQL
- the owner of the Java development language
- a key operating system player with Solaris
- one of the premiere high-end hardware providers in the world
- a key Microsoft competitor
- a key IBM competitor with a full range of applications, services, and hardware
- a one-stop shop for all your application, database, hardware and development needs

Those companies that wrote off Oracle in the past have suddenly woken up to find themselves in bed with Oracle with their MySQL, Java, Open Office, Solaris and Sun installations.

For some, it's a dream come true: they finally get a large successful conglomerate to back their favourite open source software and to support their open standards. For others, who have been burned in the past by Oracle's somewhat greedy and forever changing pricing policies, it's quite the nightmare. But perhaps Oracle has learned from their past mistakes and will approach the Sun product line with a slightly less capitalistic view, in the hopes of restoring their reputation and to win over the loyal Sun customer-base and the open source community.

In any event, with Sun's acquisition, Oracle is now positioned to fully integrate and configure their applications, databases, operating system and hardware offerings. There is opportunity to provide customers with hassle-free, pre-configured appliances.

Many questions remain to be answered, such as:

- will Oracle invest in and promote MySQL as the open source database of choice?
- will Oracle invest in and promote Open Office as the Microsoft Office alternative?
- will Oracle expand their consulting division and take on IBM head to head?
- will they appeal to the free software community?
- what will they do with Sun hardware?
- what will they do with Java?

If history has taught us anything, it's that Larry Ellison will make the most of this acquisition and will take on the likes of IBM, Microsoft and SAP. He will once again live up to his motto "It is not sufficient that we succeed, everyone else must fail!"

Sam P. Selim is the Founder and Chief Technology Officer of SQL Power Group Inc. (http://www.sqlpower.ca), a leading Canadian Consulting firm specializing in Data Warehousing and Business Intelligence. Mr. Selim is the lead architect for many of SQL Power's open source data warehousing and business intelligence tools and consulting engagements. He has over 16 years experience designing and developing custom data warehousing solutions for SQL Power clients. Sam has been a presenter at several data warehouse conferences and tradeshows and has written extensively on these topics.

The Commercial Open Source Business Model

Copyright: Dirk Riehle

From the Abstract:

Commercial open source software projects are open source software projects that are owned by a single firm that derives a direct and significant revenue stream from the software. Commercial open source at first glance represents an economic paradox: How can a firm earn money if it is making its product available for free as open source? This paper presents the core properties of commercial open source business models and discusses how they work. Using a commercial open source approach, firms can get to market faster with a superior product at lower cost than possible for traditional competitors. The paper shows how these benefits accrue from an engaged and self-supporting user community. Lacking any prior comprehensive reference, this paper is based on an analysis of public statements by practitioners of commercial open source. It forges the various anecdotes into a coherent description of revenue generation strategies and relevant business functions.

http://dirkriehle.com/publications/2009/the-commercial-open-source-business-model/

Power in People's Hands: Learning from the World's Best Public Services

Copyright: UK Cabinet Office

From the Description:

This report presents the findings of a Strategy Unit study of leading edge innovations in worldwide public services, which involved interviewing 50 academics, public servants and other experts from around the world. The report highlights more than 30 case studies from 15 countries. It emphasises that innovation and productivity come from forging stronger relationships with citizens.

http://www.cabinetoffice.gov.uk/strategy/publications/world-class-public-services/html/contents.aspx

October 5-8

GTEC

Ottawa, ON

GTEC brings together leading public and private sector experts to collaborate on serving citizens better through innovation and technology. This year's theme, "Government 2.0 Service Mash-ups", applies the private sector concept of mash-ups to public sector services.

http://www.gtec.ca/

October 19-23

Open Access Week

Global

Open Access Week is an opportunity to broaden awareness and understanding of Open Access to research, including access policies from all types of research funders, within the international higher education community and the general public. The now-annual event has been expanded from a single day to accommodate widespread global interest in the movement toward open, public access to scholarly research results.

http://www.openaccessweek.org/

October 24

Ontario Linux Fest

Toronto, ON

Ontario Linux Fest is the conference, workshop and community meeting place that is organized and run by the community for the community. This full day of presentations, workshops, birds-of-a-feather and social networking is ready to go for the third year. Everybody with an interest in using, deploying and developing Free/Libre and Open Source Software will find Ontario Linux Fest well worth their while.

http://www.onlinux.ca

October 25, 28-30

Creative Places + Spaces

Toronto, ON

Creative Places + Spaces is a multi-media, interactive, art-infused experience designed to inspire, empower, and connect thinkers, policymakers and practitioners working to build vibrant, dynamic, sustainable and creative places. Delegates and speakers will have the opportunity to hear, see, exchange and practice global perspectives on collaboration and connect them to local opportunities for change.

http://www.creativeplacesandspaces.ca/conference

October 27-28

New Brunswick Innovation Forum

Saint John, NB

The New Brunswick Innovation Forum 2009 offers a venue for innovative companies, investors, scientists and researchers to find each other so that more New Brunswick and Canadian innovations can reach the marketplace. The themes for 2009 are: E-learning, Simulation, Animation and Gaming, Health ICT/Services, Mobile Technology/Intelligent Systems, and IT Systems and Services.

http://innovation2009.nrc-cnrc.gc.ca/main_e.html

October 29

Teaching Open Source Summit

Toronto, ON

The first of the Teaching Open Source Summits will be held as a pre-conference activity to the Free Software and Open Source Symposium.

http://teachingopensource.org/index.php/Teaching_Open_Source_Summit_2009-10-29

Toronto, ON

FSOSS

October 29-30

Open source, open content, and open formats are changing the way we work, play, and learn. From software to the web to television and the media, the open source movement is spreading. Come see and hear the future in person from some of the most important thinkers in open technologies.

http://fsoss.senecac.on.ca/2009/


MASSIVE INNOVATION
Next session starts November 3
Apply at www.LeadtoWin.ca
Lead to Win
Exciting new businesses that are in Lead to Win now!
selectstart
CheckInWiz
media
catwalk networks
SCTi
EclipseSource
Castel RESEARCH
zeebumobile
Videotelephony Inc.
NOVABRAIN TECHNOLOGIES Inc.
data bridge
EnTeraSec security
polynate
Accentway
EvenuSix
Triple Thought Sustainability
SiliconPby
cornerportal
SMART ROTOR SYSTEMS
MARKTEK
Clearview
NewEraHealth
MarkOff Security

The goal of the Open Source Business Resource is to provide quality and insightful content regarding the issues relevant to the development and commercialization of open source assets. We believe the best way to achieve this goal is through the contributions and feedback from experts within the business and open source communities.

OSBR readers are looking for practical ideas they can apply within their own organizations. They also appreciate a thorough exploration of the issues and emerging trends surrounding the business of open source. If you are considering contributing an article, start by asking yourself:

1. Does my research or experience provide any new insights or perspectives?

2. Do I often find myself having to explain this topic when I meet people as they are unaware of its relevance?

3. Do I believe that I could have saved myself time, money, and frustration if someone had explained to me the issues surrounding this topic?

4. Am I constantly correcting misconceptions regarding this topic?

5. Am I considered to be an expert in this field? For example, do I present my research or experience at conferences?

If your answer is "yes" to any of these questions, your topic is probably of interest to OSBR readers.

When writing your article, keep the following points in mind:

1. Thoroughly examine the topic; don't leave the reader wishing for more.

2. Know your central theme and stick to it.

3. Demonstrate your depth of understanding for the topic, and that you have considered its benefits, possible outcomes, and applicability.

4. Write in third-person formal style.

These guidelines should assist in the process of translating your expertise into a focused article which adds to the knowledgable resources available through the OSBR.

Upcoming Editorial Themes

October 2009:	Arts & Media Guest Editor: Anthony Whitehead
November 2009:	Co-Creation Guest Editor: Stoyan Tanev
December 2009:	Bootstrapping Startups Guest Editor: John Callahan

Formatting Guidelines:

All contributions are to be submitted in .txt or .rtf format.

Indicate if your submission has been previously published elsewhere.

Do not send articles shorter than 1500 words or longer than 3000 words.

Begin with a thought-provoking quotation that matches the spirit of the article. Research the source of your quotation in order to provide proper attribution.

Include a 2-3 paragraph abstract that provides the key messages you will be presenting in the article.

Any quotations or references within the article text need attribution. The URL to an online reference is preferred; where no online reference exists, include the name of the person and the full title of the article or book containing the referenced text. If the reference is from a personal communication, ensure that you have permission to use the quote and include a comment to that effect.

Provide a 2-3 paragraph conclusion that summarizes the article's main points and leaves the reader with the most important messages.

If this is your first article, include a 75-150 word biography.

If there are any additional texts that would be of interest to readers, include their full title and location URL.

Include 5 keywords for the article's metadata to assist search engines in finding your article.

The Talent First Network program is funded in part by the Government of Ontario.

The Technology Innovation Management (TIM) program is a master's program for experienced engineers. It is offered by Carleton University's Department of Systems and Computer Engineering. The TIM program offers both a thesis based degree (M.A.Sc.) and a project based degree (M.Eng.). The M.Eng is offered real-time worldwide. To apply, please go to http://www.carleton.ca/tim/sub/apply.html.

Coral CEA is a member-based company whose mission is to assist companies of all sizes with the commercialization of communications-enabled applications (CEA). We are creating and anchoring a business ecosystem that leverages a unique, technical platform that provides advanced ICT building blocks to members. Visit http://www.coralcea.ca to become a member.