Business Intelligence tools' comparison
MS SQL Server Vs Pentaho Open Source

Muhammad Adnan Sidiqui and Mr. Shabir Mukhi
SZABIST, Karachi.

Abstract— *This paper deals with the Business Intelligence (BI), which is one of the most widely demanding business software systems.*

This Independent study deals with the comparison among the available BI tools. This Independent study will give the insight view of leading BI tools and propose BI solution as ease of use along with following consideration.

- *BI deployment challenges in an organization.*
  Complexity of BI tools and interfaces
  Cost of BI software and per-user licenses
  Difficulty accessing relevant, timely, or reliable data.
  &More

This study will provide the better understanding of BI tools with respect to Industry requirement.

Index Terms— BI tools, Business Intelligence tools.

1. INTRODUCTION

Business intelligence tool is application software. It is used to analysis, visualize data and report. This tool has been applied on the stored data. It is not essentially required data mart or data ware house. [1]

Business Intelligence (BI) is a software terminology. It is a combination of technologies, processes and tools. It helps to achieve more profit by significantly improving the efficiency, productivity and sales of an organization. With the support of BI techniques, the organization data can be organized, analyzed in a better way and then changed into an useful knowledge. It needs to initiate a profitable business action. Turning a raw, collected data into an intelligent information by analyzing and re-arranging the data according to the relationships between the data items by knowing what data to gather and manage and in what situation. [14]

Types of BI Tools
BI tools are divided into main 9 different types. There is Reporting and query software. It is our focal point There is a list of BI tools available worldwide. Theses BI tools can be categorized as open source, commercial, proprietary,

proprietary free products. These following products are Open source and commercial products but some of their editions are free of cost.

Open source and commercial products (Some of the versions of same product are free of cost)		
1	**Pentaho**	**Reporting, analysis, dashboard, data mining and workflow capabilities**
2	Palo (OLAP database)	OLAP Server, Worksheet Server and ETL Server

Proprietary BI products	
1	IBM Cognos
2	SQL Server Reporting Services
3	SQL Server Analysis Services
4	Hyperion Solutions Corporation
5	Oracle Business Intelligence Suite Enterprise Edition
6	Pentaho
7	SAP Business Objects

1.1Why do most organisations use BI?

There are some concrete reasons that organisations select to use Business Intelligence are to:

a. Centralized automated reporting improve data accuracy and minimise errors

Manual reports take more time and efforts Offline work to produce (decentralized system) spreadsheets and reports, have more errors chances that shake results.BI online work to produce centralized automated reports may reduce error chances and time. Users may access across the organization.

b. Enable fact-based decision-making

With accurate organisation information at their fingertips, employees at all levels can make more informed decisions to maximize performance.

c. Proactively detect issues and opportunities at the earliest possible stage

When objectives aren't met, a tailored BI solution alerts and allows you to analyse, understand and report on the underlying problems.

d. Provide self-service information

Managers and executives can access information anytime, anywhere. No longer do they need to make successive requests through others who may have alternate priorities or limited resources at peak times.

e. Provide a shared understanding across the organisation

Employees need a clear, consistent view of the organisation to make decisions that will improve performance. Your data is one of your most valuable assets, and it makes sense to ensure it is visible and usable in day-to-day operations.

f. Focus your business and ensure plans are executed

BI allows the Executives to define business objectives that are visible and measurable across the whole organisation. A central system means that plans can be reviewed hourly, daily, weekly – not just once a month or once a quarter.

1.2 Who uses Business Intelligence?

BI users may be from Novice computer users to Expert computer users in an organization. They may belong to different functional areas or dept. and different levels like in an educational institute e.g.

Executive Management
Operational Management
Research and Development
Student Administration
Finance
Marketing
HR

2. Chapter – Literature Review

The collection of topic related information was very difficult. There was no BI tools comparison related paper published on ACM,IEEE during 2009-2010 that provided the leading vendors tools comparison e.g. SAP's Business objects, Oracle's OBIEE, IBM's cognos. There was only source the leading consulting firms' BI tools related surveys. These latest surveys were very expensive to obtain.

We studied previous years' BI tools market surveys. These surveys provided us the insight view of the BI Industry and BI tools' market trends. These surveys were very helpful for selecting the BI tool for our research.

We selected two tools for comparison.

1. Microsoft BI Suite
2. Pentaho

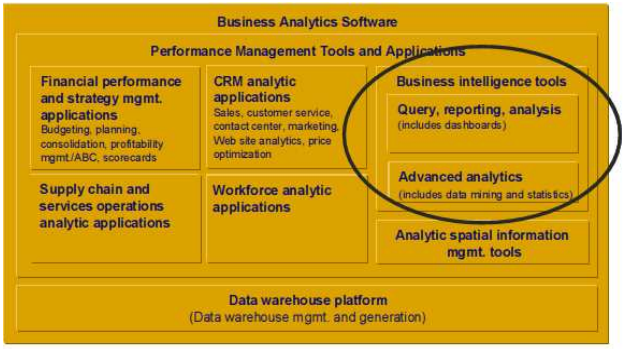
1. Selected BI tool "MS BI Suite" from proprietary BI Tool list. This was one of the top ranking tool as mentioned in different surveys. See surveys

2. Second selected BI tool was "Pentaho BI Suite" from open source BI tools list.
   In Gertner's Jan2010 report, Pentaho declared one of the top emerging open source BI Tool.

During research, we found BI related videos very useful to understand difficult BI related terminologies.

2.1 Business Intelligence Tools Market Surveys

In different surveys, BI tools market was being divided into two parts (Segment/ categories).



Source: IDC, 2007

a. Query, reporting , and analysis
b. Advanced analytics.

a. Query, reporting, and analysis (QRA)

QRA BI tools includes

A .Ad hoc query. (Query and reporting tools) and
B. Multidimensional analysis tools (OLAP servers and client-side analysis tools) as well as dashboards and production reporting tools.

Excluded : Application development tools that may be used for building reports but are not specifically designed for that purpose.

b. Advanced analytics.

Adv. analytics

1. data mining and
2. statistical software (previously called technical data analysis).

In short, BI tools are available in single package that covers above all areas. BI tools market is the combination of Standalone BI Suite vendors like Pentaho and Jaspersoft and embedded BI suite supplied by leading Industry vendors like MS, SAP and oracle .

Here we included some surveys' tables to understand Worldwide BI Regional Markets, BI Growth, BI vendors, BI Competitors, BI leaders during 2003 - 2007.
Recent surveys' are available on the net but they are costly.

TABLE 1

Worldwide Business Intelligence Tools Revenue by Segment, 2004–2006

	Revenue ($M)			Share (%)				
	2004	2005	2006	2004	2005	2006	2004–2005 Growth (%)	2005–2006 Growth (%)
Query, reporting, and analysis	4,004.9	4,487.6	5,008.5	79.5	80.0	80.1	12.1	11.6
Advanced analytics	1,031.9	1,118.6	1,244.6	20.5	20.0	19.9	8.4	11.3
Total	5,036.7	5,606.2	6,253.0	100.0	100.0	100.0	11.3	11.5

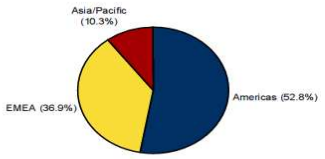
Source: IDC, June 2007

Worldwide BI Tools revenue share by region

Year: 2005

FIGURE 2

Worldwide Business Intelligence Tools Revenue Share by Region, 2006

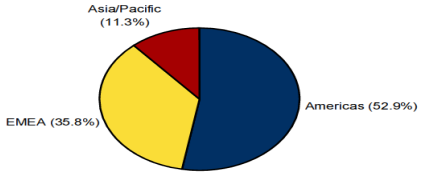


Total = $6.25B

Source: IDC, June 2007

Year: 2006

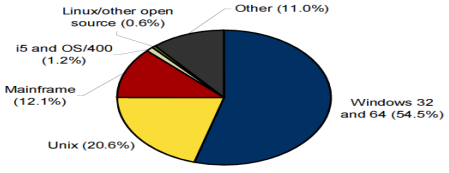
Worldwide Business Intelligence Tools Revenue Share by Region, 2005



Total = $5.725B

Source: IDC, July 2006

Worldwide Business Intelligence Tools Revenue Share by Operating Environment, 2005



Total = $5.725B

Source: IDC, July 2006

Worldwide Business Intelligence Revenue by Vendor, 2003–2005

	Revenue ($M)			Share (%)				
	2003	2004	2005	2003	2004	2005	2003–2004 Growth (%)	2004–2005 Growth (%)
Business Objects	649.9	712.6	795.3	14.3	13.9	13.9	9.6	11.6
SAS Institute	473.4	514.5	582.4	10.4	10.0	10.2	8.7	13.2
Cognos	415.5	511.5	567.2	9.1	9.9	9.9	23.1	10.9
Microsoft	125.0	281.4	353.1	2.7	5.5	6.2	125.1	25.5
Hyperion Solutions	262.8	258.6	287.1	5.8	5.0	5.0	-1.6	11.0
Oracle	195.2	214.0	247.7	4.3	4.2	4.3	9.6	15.7
MicroStrategy	142.2	185.0	212.3	3.1	3.6	3.7	30.1	14.8
SAP AG	123.9	152.2	181.8	2.7	3.0	3.2	22.8	19.4
SPSS	140.9	158.5	176.2	3.1	3.1	3.1	12.5	11.2
Information Builders	112.0	140.0	170.0	2.5	2.7	3.0	25.0	21.4
IBM	163.4	131.2	139.1	3.6	2.6	2.4	-19.7	6.0
Actuate	88.6	88.6	89.9	1.9	1.7	1.6	0.0	1.5
Hummingbird	38.2	41.4	42.9	0.8	0.8	0.7	8.4	3.6
Visual Numerics	36.8	37.4	38.7	0.8	0.7	0.7	1.8	3.5
Fujitsu Ltd.	31.2	36.5	38.1	0.7	0.7	0.7	17.1	4.4
Lawson Software	31.0	31.0	35.0	0.7	0.6	0.6	0.0	12.9
CA	27.4	29.0	29.3	0.6	0.6	0.5	5.8	0.9
arcplan	20.0	22.0	23.7	0.4	0.4	0.4	10.0	7.7
QlikTech	7.2	12.9	22.1	0.2	0.3	0.4	79.1	71.7
ProClarity	13.0	17.7	21.5	0.3	0.3	0.4	36.2	21.5
Teradata (NCR)	16.3	18.4	20.6	0.4	0.4	0.4	13.0	12.0
Other	1,444.0	1,548.7	1,660.8	31.7	30.1	29.0	7.3	7.2
Total	4,557.9	5,143.2	5,734.8	100.0	100.0	100.0	12.8	11.5

Source: IDC, July 2006

Year : 2007

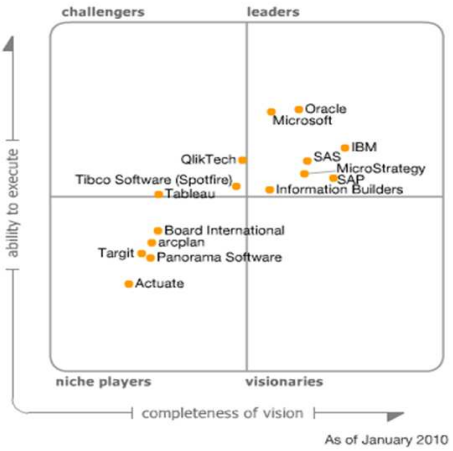
TABLE 4

Worldwide Advanced Analytics Tools Revenue by Vendor, 2004–2006

	Revenue ($M)			Share (%)				
	2004	2005	2006	2004	2005	2006	2004–2005 Growth (%)	2005–2006 Growth (%)
SAS	322.7	340.8	381.7	31.3	30.5	30.7	5.6	12.0
SPSS	130.4	151.5	174.0	12.6	13.5	14.0	16.2	14.8
Visual Numerics Inc.	37.4	38.7	41.7	3.6	3.5	3.3	3.5	7.5
Oracle	16.9	20.0	22.0	1.6	1.8	1.8	18.3	10.0
Teradata	16.3	18.4	20.4	1.6	1.6	1.6	13.0	10.5
Microsoft	9.5	13.1	18.3	0.9	1.2	1.5	37.9	40.0
Insightful Corp.	14.5	16.0	17.3	1.4	1.4	1.4	10.3	8.1
IBM	8.0	9.0	10.3	0.8	0.8	0.8	12.5	14.0
Fair Isaac	6.6	8.2	8.7	0.6	0.7	0.7	24.2	6.4
Unica Corp.	8.5	5.7	7.1	0.8	0.5	0.6	-33.0	23.7
Hitachi	8.8	9.1	6.0	0.9	0.8	0.5	2.7	-33.6
Fujitsu	6.2	6.4	5.8	0.6	0.6	0.5	2.2	-8.3
Silicon Graphics	9.1	7.2	5.4	0.9	0.6	0.4	-20.3	-25.7
ANGOSS Software International Ltd.	4.0	4.0	4.5	0.4	0.4	0.4	1.7	11.7
Subtotal	598.9	648.1	723.1	58.0	57.9	58.1	8.2	11.6
Other	432.9	470.5	521.5	42.0	42.1	41.9	8.7	10.8
Total	1,031.9	1,118.6	1,244.6	100.0	100.0	100.0	8.4	11.3

Source: IDC, June 2007

In Gartner's Jan2010 report, It mentioned that SAP, Oracle, IBM and Microsoft are leading BI players.



Pentaho and Jaspersoft have emerged as possible players in the Business Intelligence market. These two open-source players provide total solution for BI.

Both vendors' strategy is to forge OEM relationships with commercial independent software vendors (ISVs) looking to easily embed BI functionality at a low price point.

2.2 BI Future

Old Chinese saying, "To accomplish a goal, make sure the proper tools are selected." This is very true when the goal is to manage right business intelligence tool.

During 2008, the business intelligence (BI) tools market reached $7.8 billion in terms of software license and maintenance revenue. IDC predicted the BI market growth of 10.6% in 2008 which surpassed. Organisations spent and focused on BI projects that help them to reduce operational costs or retain clients. There is growing proof that more pervasive BI have a direct effect on competitiveness. It support to take better decision to consumption of resources when resources become restricted during a recession period, so BI projects will essential [2]

"The Microsoft's BI Tools annual revenue in 2005 was more than 25 percent, growing at more than twice the rate of the overall market. MS's impact on the BI tools market cannot be overemphasized. This impact will mark an evolutionary change that has been put into motion by the database vendors overall and will reshape the BI tools market over the next 15 years." [2]

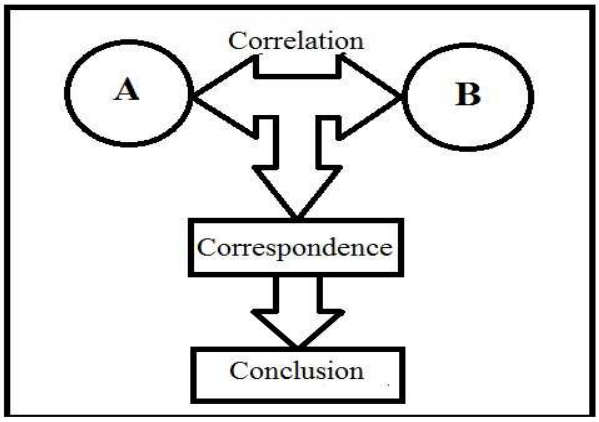
BI Future – HR demands

"There is a BI skills shortage, particularly at senior levels, but also at a technical level. In the meantime, demand and monetarily investment in BI continue in flow. Now we need to focus or bridge the gap between demand and supply for appropriately skilled professionals to design, implement and manage BI projects and, the limited supply." [8]

3. Chapter – Research Methodology

3.1 Our Research Methodology

The foundation for all our research solutions involves a thorough and proven, step methodology besides Review Articles published in magazines and on web and discus professionals.



Our topic is comparison between two products.

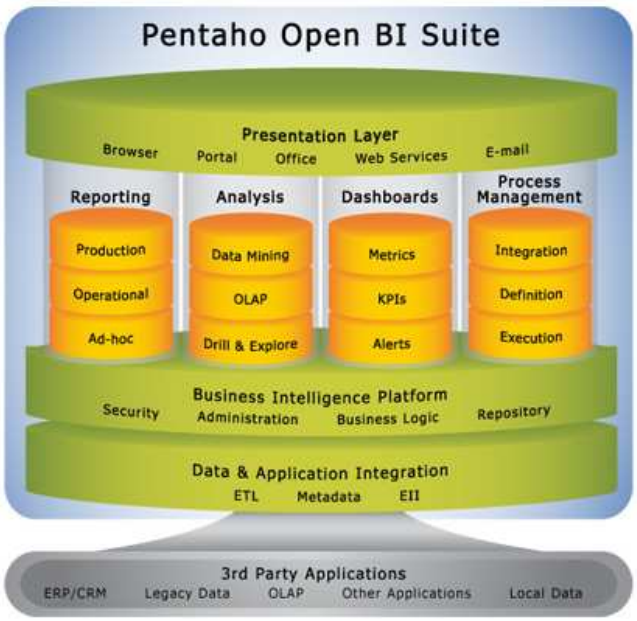
Step 1: Correlation

"A connection between two or more things" [9]

We selected two products for comparison. We take BI suite as connection to make easy for apple to apple comparison.

These products were MS SQL Server 2008 and Open source BI Pentaho.

Step 2: Correspondence



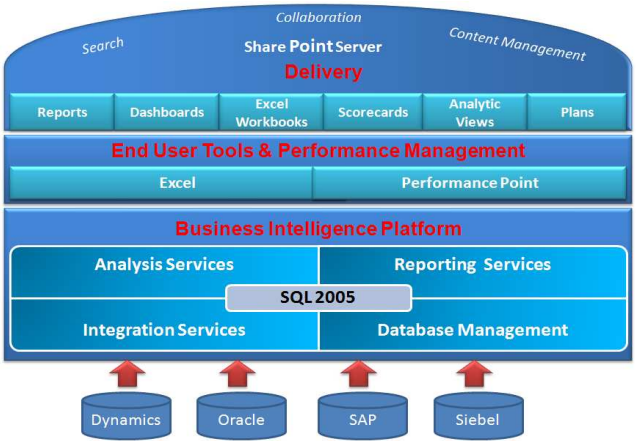
"A connection between two things" [10]

In this step, we reviewed the similarities, differences and non- compared features of the selected BI products besides we shall perform installation of both suite for our better understanding of technical knowledge.

Step 3: Conclusion
In this step, we will conclude our study with some useful suggestions after BI products' analysis.

4. Chapter – Correlation-BI Suite

Product A- MS SQL



The Business Intelligence Platform (Back-end)
MS SQL Server is a combination of 4 technologies:

1. **Relational Database:** a robust, scalable and enterprise-ready database platform.
2. **Integration Service**s: ETL component used to move, cleanse, transform and integrate data between sources and destinations.
3. **Analysis Services:** Unified Dimensional Model, OLAP technology and advanced analytical features like Key Performance Indicators and data mining.
4. **Reporting Services:** reporting authoring platform and report publishing server.

The End-User tools and Performance Management (Front End)
The End-User tools and Performance Management applications are used to access and analyze the data stored on the Business Intelligence platform.[12]

Product B- Pentaho

Pentaho BI Suite

Components Included

This release includes the complete Pentaho BI Suite, the BI server, and all client tools as described in the list below:

Component Name	Description
Pentaho BI Server	The BI Server is an enterprise-class Business Intelligence (BI) platform that supports Pentaho's end-user reporting, analysis, and dashboard capabilities with back-end security, integration, scheduling, and workflow capabilities.
Pentaho Metadata Editor (client tool)	The Metadata Editor (PME) is a tool that builds Pentaho metadata domains and models. A Pentaho metadata model maps the physical structure of your database into a logical business model.
Report Designer (client tool)	The Report Designer is the primary tool for creating and publishing Pentaho Reports. It provides a graphical interface allowing users to connect to their data, design and preview reports, and publish the reports to the Pentaho BI Platform.
Pentaho Schema Workbench (client tool)	The Schema Workbench is the primary tool for designing, editing, and publishing Pentaho Analysis(Mondrian) OLAP schemas.
Pentaho Aggregation Designer (client tool)	Pentaho Aggregation Designer is a graphical environment used to increase query performance of a Mondrian OLAP schema through the creation of aggregate tables.
Pentaho Data Integration (client tool)	Pentaho Data Integration is a graphical, dragand-drop design environment for delivering Extraction, Transformation and Loading (ETL) capabilities using a metadata-driven approach. The extensible standards based architecture

	avoids the adoption of proprietary methodologies into your ETL solution.
Pentaho Analyzer (client tool)	Pentaho Analyzer is an interactive analysis tool, (an OLAP viewer), that allows non-technical business users to create meaningful, attractive, and interactive Webbased reports and charts quickly and easily. Pentaho Analyzer puts intuitive, analytical capabilities in to the hands of knowledge workers without the usual complexities of **traditional Business Intelligence applications and with little to no training.**

[13]

5. Chapter – BI Tools Installation

Appendix "A" Installation Snapshot of Microsoft and Pentaho

6. Chapter – Analysis

(Correspondence in BI tools- Similarities, differences and noncompared features)

BI Tools Comparison Matrix (Partial)

S.#	Features	MS SQL	Pentaho
Functional Requirements			
1	Reporting Capabilities	Y	Y
2	Formatting	Y	Y
3	SQL Functionality for Report Creation	Y	Y
4	Data Visualisation Tools	Y	Y
	OLAP reports		
5	Slice and Dice	Y	Y
	Dimension hierarchy		
6	Graph and Charts	Y	Y
	Bar charts, Stacked bar, etc		
7	Calculation Functions	Y	Y
8	Dashboard Capabilities	Y	Y
9	Alerting and Notification	Y	Y
Technical Requirements			
1	Performance and Scalability	Y	Y
2	Platforms		
	Server End		
	Microsoft Windows	Y	Y
	Unix	**N**	Y
	Linux	**N**	Y
3	**Integration Requirements**	Y	Y
4	**Supported Interfaces**	Y	Y
5	**Audit and Security**	Y	Y
6	**Support Capabilities**	Y	**N**
7	**Documentation**	Y	Y
8	**Training**	Y	**N***
9	**License's fee**	Y	N*
	Enterprise	**23,50,000 ***	Free/ Nominal cost

BI Tools Analysis

- MS SQL Server BI Suite and Pentaho BI Suite both are excellent BI tools.
- Both support OLAP
- MS SQL Server doesn't support platform- Server end Unix and Linux.
- Pentaho support facilities are now available but there is no comparison with Microsoft's best support.
- Pentaho provides training videos and documentation. These documentation does not fulfill the user requirements. On the other hand Microsoft has complete road map for trainings, proper MS certified centres with MS certified trainers.

- Pentaho has no license issue, it is open source that is the big difference over Micro Soft's product MS SQL BI Suite.
- Estimated cost of MS BI Suite is Rs 23,50,000 plus where Pentaho is free of cost.

Conclusion

These Business Intelligence tools installations were performed in leading textile industry. It was quite interesting experience. During this study, I visited different software houses for support but Pentaho practical knowledge found only one person.It is interested to note that Every BI developer reach this position through "trial and error". Majority of BI developers don't like to see out of the box.

We proposed that although MS SQL server is more productive and best BI suite for any good financial conditioned organization but the best complete alternative cheapest solution in this recession time is Pentaho open source.

Still we think there is need for further study on this topic.

As we see there is a huge gap worldwide between demand and supply.

We assure you a great future ahead.

References

[1] Elliot Timo, Clear Intelligence Future: Simple, Seamless, Social, and Strategic SAP Business Intelligence Forum 2010 "The Butterfly Event" – SAP Business Objects.

[2] Ying Yan, Wen-Syan Li, Correlation aware synchronisation for near real time decision support systems. – ACM 2010.

[3] Malu Castellanous, Song Wang, Umeshwar Dayal, Chetan Gupta, SIE-OBI: a streaming information extraction platform for operational business intelligence.-ACM 2010

[4]Moez Essaidi, ODBIS: towards a platform for on-demand business intelligence services. –ACM 2010.

[5] David Lapp, Jayant Sharma. Integrating maps and location analysis with Oracle business intelligence. ACM 2010

[6] Cindy Jutrus,ERP and BI: When 1 + 1 = 3. Aberdeen Group Report Nov 2009.

[7] http://www.osgard.co.za

[8]http://dictionary.cambridge.org/dictionary/british/correlation

[9]http://dictionary.cambridge.org/dictionary/british/correspondence_2

[10] Microsoft SQL Server 2008 New Features
By Michael Otey

[11]http://www.element61.be/e/resourc
detail.asp?ResourceId=59

[12] Penatho.com

[13] http://www.learnbi.com/bi.htm

[14] www.microsoft.com

[15] Dr. Kirsti Ala-Mutka, EUROPEAN E-SKILLS 2009 CONFERENCE FOSTERING ICT PROFESSIONALISM

20 NOVEMBER 2009, BRUSSELS CONFERENCE REPORT-20 NOVEMBER 2009.