

## Business intelligence? Æh, desculpe ? nÃo entendi (parte II de III)

Business Intelligence (BI) ...hummm. Se colocarmos esta frase &agrave; entrada, de uma m&aaacute;quina de tradu&ccedil;&atilde;o &laquo;(Ingl&ecirc;s) para (Portugu&ecirc;s)&raquo; e clicarmos no bot&atilde;o que diz "GO", &nbsp; o resultado ser&aaacute; algo do g&eaacute;nero : "Intelig&ecirc;ncia nos Neg&oaacute;cios".

De que estou a falar concretamente aqui ? parece algo relacionado com uma organiza&ccedil;&atilde;o ultrasecreta, que trabalha para o governo, mas n&atilde;o, nada disso. Trata-se sim, de um t&oaacute;pico muito cativante e interessante, no mundo da informa&ccedil;&atilde;o, mundo este casado (com sucesso) com o mundo empresarial. Se n&atilde;o leu a parte 1 deste artigo, pare!

N&atilde;o vale a pena continuar, volte atr&aaacute;s por favor e come&ccedil;e pela numera&ccedil;&atilde;o correcta... Se est&aaacute; a ler esta linha ent&atilde;o vou assumir que j&aaacute; leu a parte 1 e que est&aaacute; em condi&ccedil;&otilde;es de prosseguir leitura. Excelente.

Nesta parte, eu e o leitor, iremos abordar os t&oaacute;picos abordados na leitura anterior : - (F)erramentas

- (R)ecolha

- (A)n&aaacute;lise Vamos ent&atilde;o analisar, eu e o leitor, em conjunto : (F)erramentas : J&aaacute; temos a no&ccedil;&atilde;o do que trata o BI. Mas, para implementar o pretendido, s&atilde;o necess&aaacute;rias ferramentas de software. S&oaacute; assim, poder&atilde;o ser representados fisicamente os dados, s&oaacute; assim poder&atilde;o ser processados,

manipulados e apresentados. Por exemplo, no mercado, existem ferramentas de software que permitem registar os dados fisicamente sob a forma de bases de dados (por exemplo produtos como SQL Server,

Oracle), existem ferramentas que permitem movimentar e manipular os dados de ponto a ponto (designam-se normalmente de ETL - Extraction (extrac&ccedil;&atilde;o), Transformation (transforma&ccedil;&atilde;o) e Loading (carregamento), algo muito importante no mundo de BI; por exemplo, a Microsoft tem o produto SQL Server Integration Services, a IBM tem o produto Cognos Data Manager), existem ferramentas &nbsp;que permitem efectuar an&aaacute;lise sobre os dados em termos de agrega&ccedil;&otilde;es e opera&ccedil;&otilde;es de conjunto (por exemplo a Microsoft tem o produto SQL Server Analysys Services), &nbsp;existem ferramentas que permitem exibir os resultados depois de tratados, j&aaacute; numa forma mais intelig&iacute;vel para os decisores de neg&oaacute;cio (por exemplo a Microsoft tem o Excel, SQL Server Reporting) e mesmo ferramentas de programa&ccedil;&atilde;o, que possam ser necess&aaacute;rias para complementar a solu&ccedil;&atilde;o de BI, com outros programas inform&aaacute;ticos desenvolvidos &agrave; medida. Sem as ferramentas n&atilde;o haveria BI, ficaria todo no papel apenas! (R)ecolha : Os dados precisam de ser registados num suporte f&iacute;sico adequado - bases de dados. E j&aaacute; falei delas neste artigo.

As bases de dados t&ecirc;m o conhecimento, a informa&ccedil;&atilde;o, na sua forma bruta, muito importante para os decisores de neg&oaacute;cio. Existem v&aaacute;rios poss&iacute;veis formatos de bases de dados : umas est&atilde;o simplesmente em ficheiros de texto, formatados de determinada maneira; outras est&atilde;o no formato excel ou

xml por exemplo; outras est&atilde;o em definidas em servidores de bases de dados (SQL Server, Oracle por exemplo); outras est&atilde;o embebidas em sistemas propriet&aaacute;rios, etc. Como se pode constatar, existem v&aaacute;rios formatos e localiza&ccedil;&otilde;es poss&iacute;veis. Este passo de Recolha permite realizar isso mesmo : recolher os dados necess&aaacute;rios, dos v&aaacute;rios pontos onde est&atilde;o presentes, com poss&iacute;veis manipula&ccedil;&otilde;es e transforma&ccedil;&otilde;es a serem aplicadas ainda, para um ponto especial, interm&eaacute;dio (no sentido em que ainda n&atilde;o &eaacute; acess&iacute;vel pelos decisores de neg&oaacute;cio), onde possam ser tratados e processados.

Este processo, j&aaacute; falado em cima, designa-se de ETL - (E)xtremam-se os dados, (T)ransformam-se os dados e (C)arregam-se os dados, agora sim, num ponto j&aaacute; acess&iacute;vel pelos decisores de neg&oaacute;cio. A extrac&ccedil;&atilde;o &eaacute; necess&aaacute;ria porqu&ecirc; ? &eaacute; necess&aaacute;ria porque &eaacute; preciso ir aos s&iacute;tios onde os dados est&atilde;o e copi&aaacute;-los (apenas os necess&aaacute;rios). A transforma&ccedil;&atilde;o &eaacute; necess&aaacute;ria porqu&ecirc; ? &eaacute; necess&aaacute;ria porque &eaacute; preciso corrigir eventuais falhas nos dados origem, ou criar novos dados a partir dos existentes e tamb&eaacute;m porque &eaacute; necess&aaacute;rio formatar e organizar os dados, de acordo com padr&otilde;es pretendidos. O carregamento &eaacute; necess&aaacute;rio porqu&ecirc; ? &eaacute; necess&aaacute;ria porque sen&atilde;o os decisores de neg&oaacute;cio nunca veriam os dados no seu formato final, intelig&iacute;vel para eles. Esta fase encarrega-se de movimentar os dados, do ponto interm&eaacute;dio de ETL, para o ponto final, onde j&aaacute; podem ser apresentados sob a forma de relat&oaacute;rios aos decisores de neg&oaacute;cio. (A)n&aaacute;lise : Ora bem, os dados est&atilde;o passaram pelo ETL. E agora? agora ter&aaacute; que haver uma an&aaacute;lise sobre eles. Ou n&atilde;o... Ou n&atilde;o? - pensa voc&ecirc;... Est&aaacute; confuso, eu explico :

Os dados, resultantes da aplica&ccedil;&atilde;o do processo de ETL, podem apenas servir para serem consultados directamente por relat&oaacute;rios de dados. N&atilde;o &eaacute; necess&aaacute;rio e imperativo haver uma fase de an&aaacute;lise.

Ent&atilde;o... o que significa &laquo;An&aaacute;lise&raquo; ? Diz respeito a um tratamento especial, efectuado sobre esses dados resultantes do ETL, e que envolve agrega&ccedil;&otilde;es de dados, defini&ccedil;&otilde;es de hierarquias entre eles e opera&ccedil;&otilde;es de conjunto (somat&oaacute;rios, totais, m&eaacute;dias, m&aaacute;ximos, m&iacute;nimos, etc), podendo depois serem exibidos pelos decisores de neg&oaacute;cio,

atrav&eaacute;s da navega&ccedil;&atilde;o pela informa&ccedil;&atilde;o, juntando m&eaacute;tricas, dimens&otilde;es e atributos... "&Aacute;h ?! o qu&ecirc;?! m&eaacute;tricas, dimens&otilde;es e atributos ?! o que &eaacute; isso???" - est&aaacute; de novo a pensar. Mais uma vez, eu explico : Os

dados resultantes do processo de ETL, normalmente, são organizados sob a forma de métricas e dimensões de análise :  
- &laquo;Métricas&raquo; : são os factos que os decisores de negócio querem medir e conhecer. Por exemplo: número de vendas, total de custos, total de receitas, total de empregados, etc. - &laquo;Dimensões&raquo; : definem as várias maneiras como os factos podem ser analisados. Por exemplo, "listar &laquo;total de vendas&raquo; POR &laquo;região&raquo; POR &laquo;empregado&raquo; POR &laquo;mercado&raquo;" significa que se quer saber o total de vendas (facto) discriminado por : 1-região (dimensão)

2-empregado (dimensão)

3-mercado (dimensão) o que dá uma perspectiva de conhecimento aos decisores de negócio mais completa e abrangente! Assim, eles conseguem perceber onde estão concentradas mais (ou menos) as vendas, quais os empregados que vendem mais e em que mercado do ano vendem mais. Isto é espectacular! E podemos definir mais dimensões para além destas. Depende do projecto, do que os decisores de negócio querem exactamente. Estabelecendo uma analogia, trata-se como se fosse de uma equação matemática :  $Y = X_1 + X_2 + X_3 + X_4$ , onde  $X_n$  seriam as dimensões e  $Y$  o facto a ser medido... E as hierarquias ? É uma maneira de estabelecer hierarquias de informação, de um nível mais abrangente, para outro mais atómico. Por exemplo, a dimensão &laquo;região&raquo; é um exemplo adequado : um país tem regiões, uma região tem cidades, uma cidade tem concelhos, etc... isto é uma hierarquia, no topo o país, no fundo o concelho.

No fundo, é uma maneira mais inteligente e auxiliar para o decisor de negócio navegar e seleccionar a informação que pretende. E os atributos ? são campos informativos definidos nas dimensões e que ajudam na filtragem e selecção de informação. Por exemplo, na dimensão &laquo;empregado&raquo; podem-se

definir atributos como a idade, sexo, categoria, etc. Por exemplo, um decisor de negócio, pode querer ver se são os empregados masculinos que vendem mais ou não. Bastaria então, discriminar a informação dos factos, pela dimensão &laquo;empregado&raquo; usando o atributo "sexo"... Fiel ! A fase de análise envolve estes conceitos todos. E existem também ferramentas para fazer este tipo de análise.

No fundo, os dados resultantes do ETL, num dado formato, são guardados em outro formato numa base de dados multidimensional (capacidade de poder de analisar informação numa perspectiva dimensional). Isto é fantástico! Mas eu sou suspeito por gostar muito desta realidade... Neste momento, já vamos com setenta e um por cento (71%) da nossa missão cumprida.

Na parte III (e última), na sequência da nossa missão conjunta, iremos analisar e aprender sobre : (P)rocessamento e (A)presentação. &nbsp; Até lá, e obrigado pela sua atenção. Os meus cumprimentos,

Pedro Mesquita ( <http://pedromesquita.pt.to> )

## Sobre o Autor

(ver site pessoal)

Source: <http://www.artigopt.com>