

Desenvolvimento de um protótipo de SAD para analisar e classificar os processos dentro da procuradoria geral do Estado do Tocantins

Wllynilson Pereira Cardoso Carneiro¹, Ewerton Ferreira Santiago¹, Marcus Henrique de Sousa Oliveira¹, Fernando Luiz de Oliveira²

¹Sistemas de Informação - Centro Universitário Luterano de Palmas (CEULP/ULBRA) – Avenida Teotônio Segurado, CEP 77.019-900 – Palmas, TO – Brasil

²Mestre em Ciência da Computação

{wllynilson, ewertonferreirasantiago, lmarcushenrique, nandoluiz.br}@gmail.com

Resumo. *O presente estudo aborda um levantamento realizado sobre um conjunto de informações de processos disponibilizados pela Procuradoria Geral do Estado do Tocantins, sendo estes distribuídos dentro do Tribunal de Justiça do Estado do Tocantins, gerenciados por meio da ferramenta E-proc. São informações públicas que podem ser consultadas por qualquer cidadão. A partir dessa base de dados foi desenvolvido um protótipo funcional de um sistema de apoio à decisão que passou pela criação de um Data Warehouse e pela utilização de uma ferramenta OLAP para realizar análises. Este trabalho aborda soluções de apoio à gestão estratégica, como parte integrada, envolvendo ações que tramitam em todo o estado e seus diversos processos de decisão, visando garantir melhor análise e rapidez na solução dos processos gerado.*

1. Introdução

O número de dados gerados por ações governamentais cresce exponencialmente, levando a uma grande dificuldade na análise e manipulação desses dados. É válido ressaltar que, para funcionários de órgãos que necessitam gerar informações precisas, em um contexto de grande volume de dados e, além de tudo, desorganizados, este cenário é bastante desafiador.

Muitos órgãos enfrentam grandes desafios, conforme apresentado anteriormente, quando se trata de informações para tomada de decisão. Isso se dá por causa da falta de ferramentas e modelos de processos adequados para auxiliar a transformação dos dados em informação e que ajude a administração na tomada de decisões.

Há uma necessidade da Procuradoria Geral do Estado do Tocantins (PGE-TO) por ferramentas que auxiliam o gerenciamento e análise de processos. Diariamente é encaminhado para o PGE-TO uma grande quantidade de processos, das diversas áreas governamentais, gerando dificuldades no gerenciamento dos mesmos.

O presente trabalho tem por objetivo apresentar uma solução para o problema relatado. Para tanto, foi desenvolvido um protótipo inicial de um Sistema de Apoio a Decisão (SAD), baseado no desenvolvimento de um Data Warehouse (DW) e utilização de uma ferramenta OLAP para proporcionar à PGE-TO a possibilidade de realizar consultas analíticas aos processos.

1.1. Sistema de Apoio à Decisão

Em grandes e pequenas corporações, independente de qual área profissional, está cada vez mais presente o processo de tomada de decisão. É de grande importância para essas corporações chegar na melhor decisão, mas isso nem sempre acontece. Devido essa realidade, existem sistemas que auxiliam esses profissionais na tomada de decisão, que é o caso SAD.

Segundo os autores Silva, Melo, Ribeiro e Borges (2013), o SAD é um grande avanço dos Sistemas de Informação (SI), que colaboram auxiliando nesse processo.

1.2. Data Warehouse

Diferente de um banco de dados comum, Data Warehouse (DW), é utilizado para definir um conjunto de novos conceitos e ferramentas que evolui para uma tecnologia permitindo atacar o problema de oferecer a todas as pessoas-chave da empresa acesso a qualquer nível de informação necessária para que a organização possa sobreviver e prosperar em um mundo cada vez mais competitivo (Date, 2004). A informação deve ser tanto bem trabalhada quanto bem apresentada.

Segundo Machado(2010) estatísticas mundiais mostram que o tempo gasto com a obtenção e análise dos dados para a tomada de decisão é significativamente menor quando da existência de aplicações e estratégias de DW, permitindo que a tomada de decisão seja mais calculada.

Para que os dados sejam melhores estruturados em subconceitos bem definidos de acordo com Inmon (2005), que foi o primeiro a cunhar o termo Data Warehouse (ou armazém de dados) para repositórios analíticos corporativos, sua definição clássica é a de "uma coleção de dados orientada por assunto, integrada, variante e não volátil, que tem por objetivo dar suporte aos processos de tomada de decisão". Estas características serão detalhadas abaixo:

- Orientado por assunto: toda a modelagem de um Data Warehouse baseia-se nas características e principais assuntos da empresa.
- Integrado: característica mais importante de um Data Warehouse, pois trata da integração do ambiente operacional com as aplicações do Data Warehouse.
- Variável com o tempo: diferente dos sistemas Online Transaction Processing ou Processamento de Transações em Tempo Real (OLTP), que guardam históricos por um tempo curto, um Data Warehouse guardam históricos por muito tempo, possibilitando que análises de tendências ao longo do tempo sejam feitas com seus dados.
- Não volátil: após as etapas de inclusão, alteração e exclusão de dados nos sistemas transacionais os dados no ambiente do Data Warehouse são filtrados gerando informação. Após esta etapa os dados estão disponíveis para operações de consulta e exclusão, sem que possam ser alterados.

Ainda para o autor, o subconjunto do DW é o Data Mart (DM) que "desempenha o papel de um DW (departamental, regional ou funcional), podendo-se construir uma série deles ao longo do tempo e eventualmente vinculá-los através de um DW lógico empresa-inteira". É de fundamental importância ressaltar que o uso de DM é opcional e vai de acordo com a necessidade de cada organização.

A seguir, a Figura 1 apresenta a estrutura e o processo de geração de um DW. A partir de diferentes bases de dados, sejam elas um banco operacional, planilhas e/ou arquivos estruturadas dentre outros, é realizado o processo ETL, que será melhor detalhado ao longo do artigo, com a finalidade de melhorar as informações e alimentando os DM que por sua vez alimenta o DW.

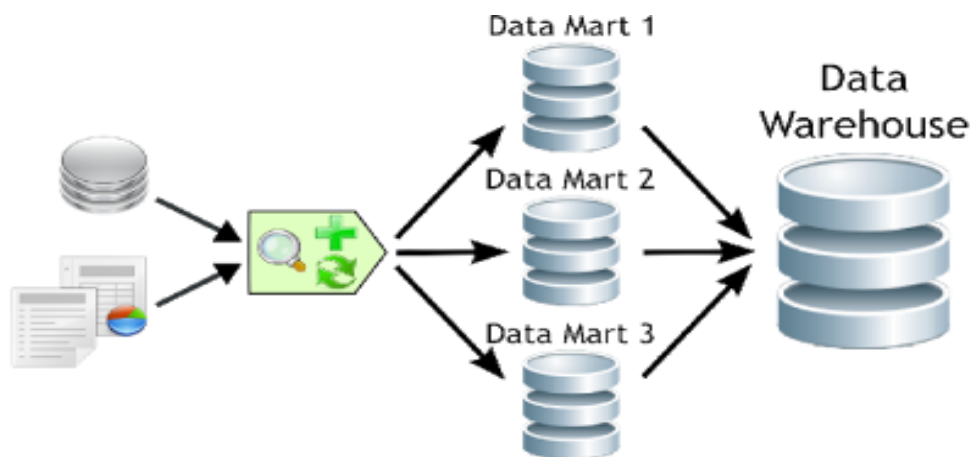


Figura 1. Estrutura e processo de geração de uma base de dados DW

Fonte: <http://sandrorgguimaraes.blogspot.com.br/2012/03/datawarehouse-dw-data-mart-dm.html>

Construir um ambiente de dados de acordo com o conceito DW não é tão simples quanto a criação de um banco de dados relacional. A arquitetura muda em relação a um banco de dados operacional, pois no DW é necessário a extração de dados de mais de uma base operacional. Para esta atividade faz-se necessário o processo ETL, foco da próxima seção.

1.2.1. ETL

O processo ETL é o mais penoso no desenvolvimento de um DW. Segundo Abreu(2010), ETL (Extract, Transform and Load) destina-se à extração, transformação e carga dos dados de uma ou mais bases de dados de origem para uma ou mais bases de dados de destino (Data Warehouse). De acordo com GONÇALVES (2003), um ponto importante que deve ser analisado pelas empresas no momento de adquirir a ferramenta ETL é se certificar de que a mesma consegue integrar as diferentes fontes de dados existentes na organização (exemplo: sistemas ERP, CRM, SAP, VSAM, People Soft, Siebel, OLTP Relacionais, etc.). Na figura 2, apresentada abaixo, mostra uma visão do processo ETL.

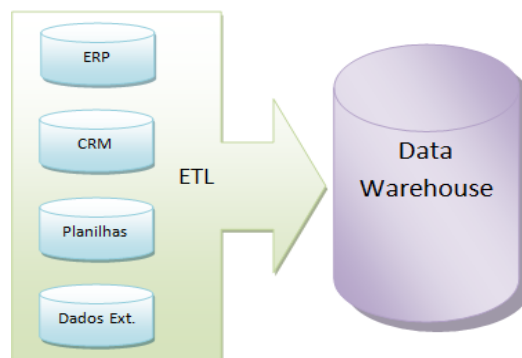


Figura 2. Visão do ETL

Fonte: <https://www.devmedia.com.br/extract-transformation-and-load-etl-ferramentas-bi/24408>.

1.2.2. Olap

A lista acima apresenta apenas alguns tipos de itens. A lista completa, com sua respectiva É uma ferramenta muito importante no contexto gerencial, ajudando a analisar de forma mais eficiente, a quantidade de dados crescente armazenada pelas organizações, transformados-os em informação útil (THOMSEN, 2002).

O termo processamento analítico online (OLAP), segundo Turban e Aronson (2001), refere a uma variedade de atividades usualmente executadas pelos usuários finais em sistemas

on-line. Laudon e Laudon (2004) utilizam a expressão análise multidimensional de dados para se referir a esta ferramenta de apoio à decisão. Turban et al. (2005, p. 86) interpretam o OLAP como “uma categoria ampla de aplicações e técnicas para coletar, armazenar, analisar, fornecer acesso aos dados e ajudar os usuários da empresa a fazerem melhores negócios e tomarem melhores decisões estratégicas”.

2. Metodologia

Para melhor compreensão acerca da metodologia utilizada no respectivo trabalho foi elaborada a figura abaixo.



Figura 3. Metodologia utilizada para o desenvolvimento do presente trabalho

Para o desenvolvimento do projeto foi necessário primeiramente a aquisição de conhecimento teórico sobre DW, ETL e OLAP por meio de artigos científicos, livros e teses, bem como o conhecimento dos procedimentos processuais da PGE-TO. Após a aquisição de conhecimento foram realizadas análises de ferramentas que melhor se adequassem ao contexto de aplicação do projeto.

Dentre as ferramentas analisadas foram selecionadas três ferramentas que proporcionam um melhor desempenho e plataforma de trabalho amigável:

- SQL Server Integration Services (SSIS): Ferramenta responsável pelo processo ETL.
- SQL Server: Ferramenta de modelagem e gerenciamento de banco de dados.
- MicroStrategy: Ferramenta OLAP utilizada para geração dos gráficos para análise.

Após a conclusão da etapa acima realizou-se a extração dos dados cedidos pela PGE-TO. As etapas foram assim dispostas: obtenção dos dados, criação do DW, realização do processo ETL e apresentação dos dados obtidos através do processo OLAP.

3. Desenvolvimento

O desenvolvimento do trabalho se deu nas seguintes etapas: obtenção dos dados, criação do DW, realização do processo ETL e apresentação dos dados obtidos através do processo OLAP.

3.1. Obtenção dos dados

Para a construção e análise dos processos, foram obtidos dados de uma base de dados no formato CSV (planilha Excel), conforme apresentado na figura 4, coletado junto a PGE-TO. Nesta base de dados contém informações de processos dos meses de outubro de 2013 a setembro de 2017 de 1º grau e dos meses de outubro de 2013 a setembro de 2017 de 2º grau.

	A	D	E	F	G	I	J
1	Localidade Judicial	Número Processo	Classe da Ação	Assunto Principal	Réu Preso	Mês	Ano
2	Itacajá	0000646-06.2016.827.2723	Procedimento Comum	Obrigação de Fazer / Não Fazer	Não	10	2016
3	Aurora do Tocantins	0000830-95.2016.827.2711	Embargos à Execução Fiscal	ICMS / Incidência Sobre o Ativo Fixo	Não	10	2016
4	Paraná	0000909-11.2016.827.2732	Procedimento Comum	Erro Médico	Não	10	2016
5	Arapoema	0000941-88.2016.827.2708	Procedimento Comum	Direito de Imagem	Não	10	2016
6	Ananás	0000954-05.2016.827.2703	Procedimento Comum	Fornecimento de Medicamentos	Não	10	2016
7	Ananás	0000955-87.2016.827.2703	Procedimento Comum	Fornecimento de Medicamentos	Não	10	2016
8	Palmeirópolis	0000967-20.2016.827.2730	Inventário	Inventário e Partilha	Não	10	2016
9	Almas	0000973-17.2016.827.2701	Usucapião	Usucapião Extraordinária	Não	10	2016
10	Almas	0000984-46.2016.827.2701	Procedimento Comum	Direitos da Personalidade	Não	10	2016
11	Almas	0000988-83.2016.827.2701	Procedimento Comum	Obrigação de Fazer / Não Fazer	Não	10	2016
12	Arraias	0001001-58.2016.827.2709	Inventário	Inventário e Partilha	Não	10	2016
13	Arapoema	0001018-97.2016.827.2708	Tutela Cautelar Antecedente	Obrigação de Fazer / Não Fazer	Não	10	2016
14	Palmeirópolis	0001024-38.2016.827.2730	Procedimento Comum	Obrigação de Fazer / Não Fazer	Não	10	2016
15	Wanderlândia	0001030-12.2016.827.2741	Sobrepilha	Inventário e Partilha	Não	10	2016
16	Palmeirópolis	0001035-67.2016.827.2730	Inventário	Inventário e Partilha	Não	10	2016
17	Goiatins	0001037-67.2016.827.2720	Procedimento Comum	Multas e demais Sanções	Não	10	2016
18	Arraias	0001060-46.2016.827.2709	Inventário	Inventário e Partilha	Não	10	2016

Figura 4. Relatório dos processos da PGE-TO

3.2. Criação do DW

Com a análise das informações contidas nos relatórios gerados, a modelagem dimensional adotada foi o esquema estrela. A tabela de fatos, contém dados quantitativos sobre os processos. Nas tabelas de dimensão, contém dados que descrevem as medidas contidas na tabela de fatos. A figura 5, a seguir, apresenta a modelagem dimensional do DW baseado no esquema estrela.

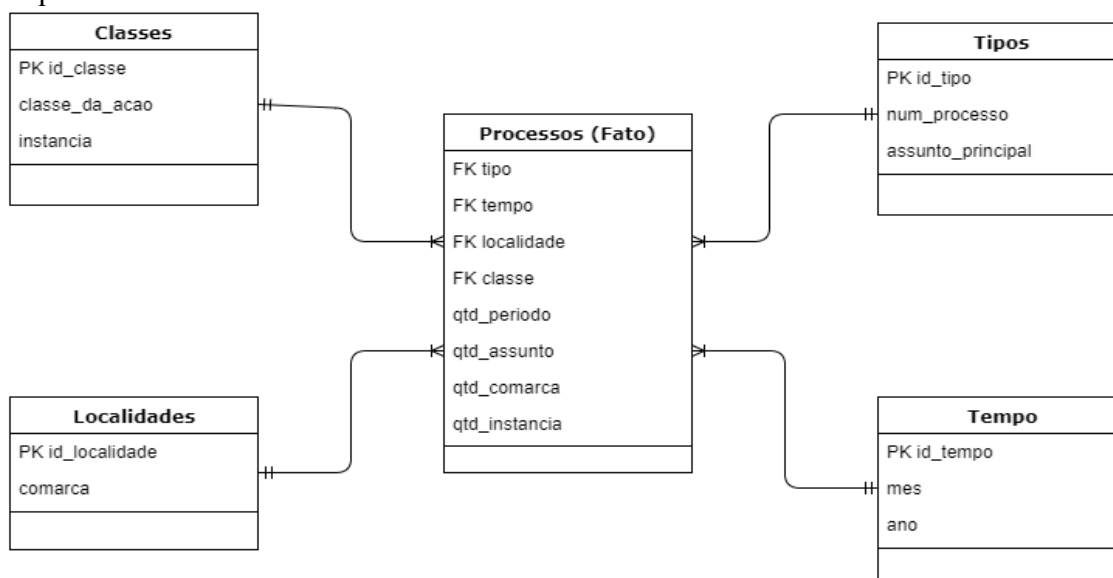


Figura 5. Esquema Estrela

O esquema estrela apresentado acima contém uma tabela de fato e 4 tabelas de dimensão, conforme descritas a seguir:

- Tabela de fato Processo: contém os atributos qtd_perodo, qtd_assunto, qtd_comarca e qtd_instancias. A chave primária consiste nas quatro chaves estrangeiras: tipos, tempo, localidades e classes.
- A tabela de dimensão Tipos é composta pela chave primária o id de tipo e pelos atributos: número do processo e assunto principal.
- A tabela de dimensão Tempo tem como chave primária o id de tempo e abriga mais dois atributos: mês e ano.
- A tabela de dimensão Localidades contém sua chave primária que é o id de localidade o atributo comarca.

- A tabela de dimensão Classes possui sua chave primária que é o id de classe, e os atributos: classe da ação e instância.

Para esse trabalho foi adotado nível alto de granularidade, visto que os dados disponibilizados não apresentam grande nível de detalhe, devido restrições de segurança da informação por ter que preservar dados pessoais dos envolvidos.

Após a coleta dos dados e a definição do modelo dimensional, foi criado um banco de dados utilizando o SQL Server para receber as dimensões preenchidas com as informações dos arquivos CSVs.

3.3. Processo ETL

Após a criação do DW foi realizado o processo ETL. A ferramenta SSIS está presente no ambiente de desenvolvimento SQL Server Data Tools, que pode ser encontrado após a instalação em: Iniciar - Todos os programas - Microsoft SQL Server - SQL Server Data Tools.

Essa ferramenta proporciona realizar a extração de dados de outros bancos relacionais ou planilhas, arquivos de texto, e inseri-los em um fluxo de trabalho, possibilitando o carregamento desses dados ao seu destino. Através da criação do fluxo de dados é possível selecionar os arquivos criando o componente Flat File Source. Na figura 6, a seguir, é apresentado o mapeamento das colunas do arquivo com as colunas no banco.

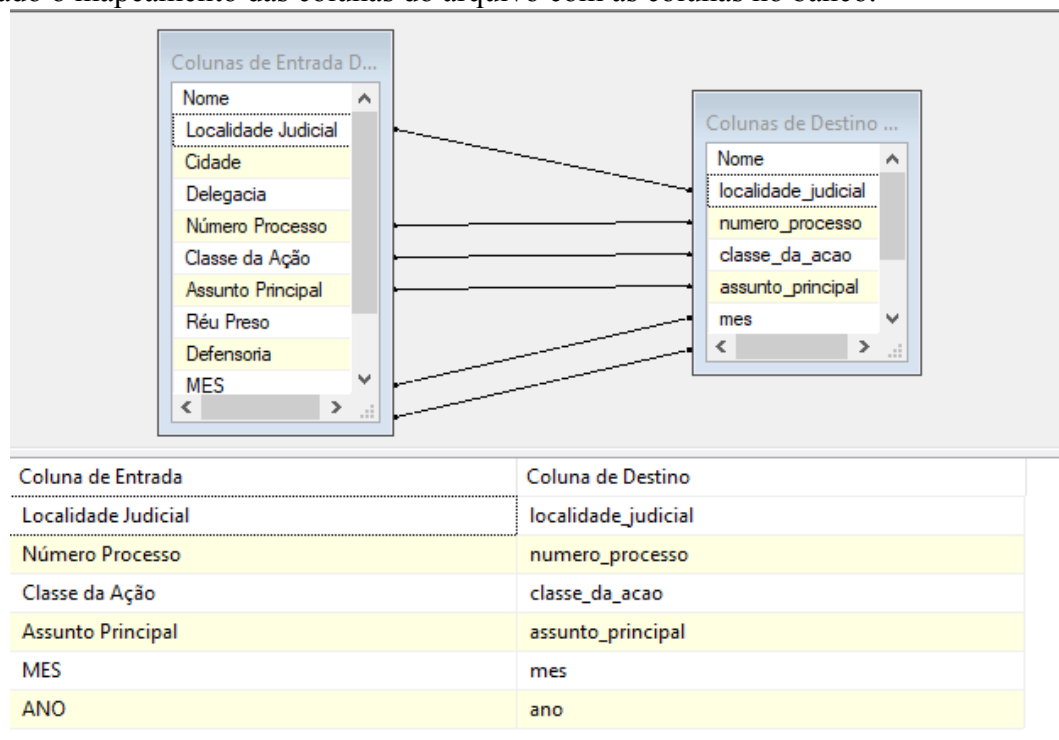


Figura 6. Mapeamento das colunas do arquivo com as colunas do banco de dados

No quadro esquerdo é apresentado as colunas do arquivo, e a do lado direito as do banco de dados. Caso o nome das colunas do arquivo e do banco sejam os mesmos, a própria ferramenta faz a ligação entre eles de forma automática, caso não seja, abaixo dos quadros é apresentado as opções de fazer a ligação.

Na última etapa, com todo o mapeamento feito e os componentes ligados foi executado o processo de carga dos dados no banco de dados em suas respectivas tabelas dimensão.

3.4. Processo OLAP

Com os dados carregados no DW, foi realizado o processo OLAP através da ferramenta MicroStrategy. Nessa etapa foi necessário criar a ligação da ferramenta com o DW. No MicroStrategy o procedimento para ligação é bem simples, devendo apenas, o usuário, selecionar qual o tipo do banco de dados, informar o número da porta para conexão e selecionar a base de dados.

A partir da conexão estabelecida é possível que o usuário tenha acesso a todos os dados existentes no DW. e assim possa realizar todas análises necessárias. Os dados apresentados na figura 7, a seguir, apresenta a quantidade de processos gerados por localidade judicial (comarca) no mês de setembro de 2015, que estão em 1ª instância.

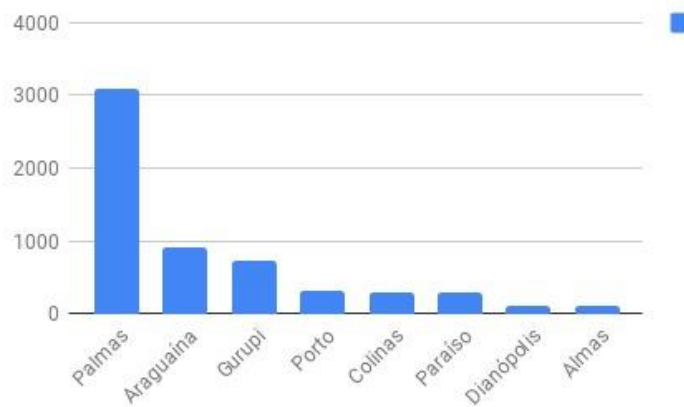


Figura 7. Quantidade de processos por Localidade Judicial (Comarca)

Nota-se que as maiores cidades do estado são as que geraram mais processos. Na figura 8, a seguir, é apresentado o gráfico de processos por assunto que estavam em 2ª instância no mesmo período, setembro de 2015.



Figura 8. Quantidade de processos por assunto (2ª instância)

Não foi possível apresentar os gráficos da 2ª instância por localidade judicial, devido ser a mesma localidade informada nos relatórios: “2ª Instância do Tribunal de Justiça do Tocantins”.

Analisando os dois relatórios gerados (Figura 7 e Figura 8), percebe-se que a cidade de Palmas foi a número 1 em processos gerados, isso se dá por ser capital e por concentrar o maior quadro de funcionalismo público. Na segunda instância o assunto que mais gerou processos foi relacionado a “Férias”.

4. Considerações Finais

O trabalho apresentado teve como proposta o estudo e levantamento da necessidade de uma ferramenta para auxiliar na análise de processos da PGE-TO. Devido a grande quantidade de processos gerados e tramitados diariamente, é fundamental a existência de uma ferramenta de apoio a tomada de decisão. O conceito de DW favorece as ferramentas OLAP a trabalharem os dados em diversas situações de pesquisas para auxiliar a identificar situações que possam gerar lentidão no andamento dos processos, retorno e finalização na instância de origem. As valiosas informações de apoio à decisão podem representar novas melhorias e opções estatísticas antes não observadas. Para tal implantação é necessário conhecimento de conceitos e técnicas para a realização do problema proposto.

Na realização do processo de pesquisa dos dados, foi identificado que a PGE-TO não possui ferramentas que facilitam a busca de processos que tiveram tramitação em mais de uma instância, isso gera dificuldades em uma pesquisa a fim de identificar se tais processos pertenciam anteriormente a 1ª instância no momento que é dada entrada em 2º instância, pois quando é passado para a 2ª instância gera-se um novo número de processo, enquanto o anterior fica tramitando na 1ª instância, aguardando ajuizado sem programação ou estimativa de conclusão. Atualmente só é possível obter este resultado/associação acessando cada processo manualmente para averiguar se o mesmo passou pela primeira e pertence ao 2º grau, fazendo com que o processo seja moroso e exaustivo, inviabilizando estabelecer dados estatísticos.

Para que haja bom proveito das funcionalidades da ferramenta OLAP, é importante que se tenha na base de dados do DW muitas colunas ou várias tabelas com informações. Devido o zelo, responsabilidade e sigilo de pessoas físicas e jurídicas, não foi concedido por meio da PGE-TO para fins de pesquisa acadêmica, os atributos que identificam as partes do processo. Esta medida restringiu a busca de pessoas, sendo que no atual sistema é a única forma de acessar o relacionamento entre instâncias.

Uma possível solução para identificar se um determinado processo continua tramitando na primeira instância ou progrediu para as demais instâncias, é relacionar os registros da coluna “Classe da ação” que esteja definido como “Apelação”. Essa informação é que pré-determina se o processo coube recurso da primeira instância para segunda instância. Com esse cruzamento de informações é possível chegar a um quantitativo de processos próximo do esperado que podem pertencer às duas instâncias.

Referências

- DATE, C. J. Introdução a Sistemas de Bancos de Dados. 8ª Ed., Rio de Janeiro: Campus, 2004.
- FERREIRA, João; MIRANDA, Miguel; ABELHA, Antônio; MACHADO, José. O Processo ETL em Sistemas Data Warehouse. INForum 2010 - II Simpósio de Informática. Setembro, 2010.
- INMON, W. H. Building the Data Warehouse, Fourth Edition. Indianápolis EUA: Wiley Publishing, 2005.
- SILVA, Breno Rodrigues; MELO, Mahelder Carvalho de; RIBEIRO, Michel David de Assis; BORGES, Leandro. Sistemas de Apoio à Decisão Médica (SADM). Revista Eletrônica de Sistemas de Informação e Gestão Tecnológica. Centro Universitário de Franca (Unifacef). 2013.
- THOMSEN, E. Construindo Sistemas de Informações Multidimensionais. 2ªed. São Paulo: Campus, 2002.
- TURBAN, E. & ARONSON, J. E. Decision support systems and intelligent systems. 6ª ed. Upper Saddle.