

Documentação e migração do processo de *data mining* do CIGarne (Centro de Inteligência da Carne Bovina)

Primeiro autor: Miguel Dias Montibeller de Oliveira
Demais autores: Oliveira, M. D. M.¹; Carromeu, C.²; Dias, F. R. T.³

Resumo

O CIGarne é um observatório com foco em consolidar e monitorar análises e tendências do mercado de inovações da pecuária bovina de corte. Dentre seus diversos projetos, atua na vertente de mineração de dados (do inglês, *data mining*) sobre análises da qualidade da carne usando como base informações de abate. O objetivo deste trabalho é apresentar o processo de documentação e migração deste para os servidores da Embrapa Gado de Corte do CIGarne, *data mining* que vem sendo realizado em parceria com a startup BOVBI. Para a realização da migração, foi clonado o GIT (um repositório online, que controla as versões dos arquivos). O processo de *data mining* do CIGarne é realizado utilizando o Pentaho, uma plataforma de software, composta por várias ferramentas. Foram executadas algumas rotinas, que criaram o data warehouse (ou depósito de dados digitais) em servidor da Embrapa Gado de Corte e usando o PDI (*Pentaho Data Integration*), foi realizado o ETL (*Extract, Transform and Load*), processo onde a ferramenta executa os *jobs* e *transformations*, resultando na extração de dados úteis. Com os dados no *data warehouse* pôde-se disponibilizá-los em um endereço de acesso local, utilizando o *software Schema Workbench*, onde cria-se o *schema* de como

(1) Graduando da Universidade Federal de Mato Grosso do Sul – UFMS, mig.beller@live.com
(2) Analista de TI da Embrapa Gado de Corte (3) Pesquisador da Embrapa Gado de Corte. * Autor correspondente.

essas informações serão organizadas. Posteriormente, com esse *schema*, foram criados os dashboards (painéis que mostram métricas importantes) utilizando o CTOOLS, um *plugin* com ferramentas que permite a criação de dashboards. Na documentação, foi utilizado o Jekyll (um gerador de sites que utiliza a linguagem de marcação *Markdown*), gerando uma página web com informações técnicas. Este trabalho possibilitou que o processo de *data mining* do CICarne fosse documentado e executado na Embrapa Gado de Corte, possibilitando expandi-lo futuramente para outras bases de dados da pecuária, e aos colaboradores do CICarne um melhor entendimento do processo como um todo.

Parceria / Apoio financeiro

Embrapa Gado de Corte.