

Análise de Eficiência Alimentar de Gado Leiteiro a partir da Integração de Bases Heterogêneas e Ontologias

Heitor Magaldi¹, Regina Braga¹, Wagner Arbex^{1,2}, Mariana Magalhães Campos², Carlos Cristiano Hasenclever Borges¹, José Maria N. David¹, Fernanda Campos¹, Victor Stroele¹

¹ Programa de Pós-Graduação em Ciência da Computação, Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil

² Embrapa Gado de Leite, Empresa Brasileira de Pesquisa Agropecuária – Embrapa, Juiz de Fora, Minas Gerais, Brasil

heitor.magaldi@gmail.com.br, {regina.braga, wagner.arbex, jose.david, carlos.borges, fernanda.campos, victor.stroele}@ufjf.edu.br, mariana.campos@embrapa.br

***Abstract.** With today's increasingly competitive market, dairy farmers need to cut costs and make their herds competitive. In this sense, the computational support has provided alternatives to the identification of more efficient animals and, consequently, providing economic and environmental gains. This paper presents an architecture to support food efficiency research developed at Embrapa Gado de Leite, with the aim of discovering new knowledge and new relationships in a large experiment dataset, using ontologies and data analysis visualization techniques. The preliminary evaluation results showed to be promising. Therefore, we consider that the use of ontologies and visualization techniques can contribute to the advancement of research in feed efficiency.*

***Resumo.** Com o mercado atual cada vez mais competitivo, os produtores de leite precisam reduzir custos e tornar seus rebanhos competitivos. Nesse sentido, o suporte computacional tem proporcionado alternativas para a identificação de animais mais eficientes e, conseqüentemente, proporcionando ganhos econômicos e ambientais. Este artigo apresenta uma arquitetura de apoio à pesquisa de eficiência alimentar desenvolvida na Embrapa Gado de Leite, com o objetivo de descobrir novos conhecimentos e novas relações em um grande conjunto de dados experimentais, utilizando de técnicas de visualização, ontologias e análise de dados. Os resultados preliminares da avaliação mostraram-se promissores. Portanto, consideramos que o uso de ontologias e técnicas de visualização podem contribuir para o avanço da pesquisa em eficiência alimentar.*

1. Introdução

Atualmente, a pecuária leiteira brasileira vem lidando com novos desafios. O aumento dos custos de produção, da preocupação dos consumidores com a qualidade e segurança alimentar, do bem-estar animal e impactos ambientais da agropecuária ocasionam margens de lucro menores aos produtores [Campos et al. 2015]. Campos et al. (2015) destacam que os gastos com a alimentação representam o principal custo da atividade

pecuária. Além disso, com a eficiência alimentar, além de impactos econômicos, impactos ambientais são observados, pois animais eficientes produzem menor desperdício de nutrientes e excreções. Um animal é classificado como eficiente quando atinge o mesmo nível de produção consumindo menos alimento que os demais.

Hoje, os pesquisadores analisam esses dados de forma manual, utilizando de ferramentas estatísticas para seleção dos animais eficientes e, a partir disso, analisarem em que os animais eficientes diferem dos não eficientes. Essas análises são realizadas a partir de planilhas eletrônicas. O cruzamento de dados entre animais, considerando diferentes índices de eficiência alimentar e considerando os diversos experimentos é uma atividade que demanda tempo e expertise computacional e por conta disso, nem sempre é realizada. Além disso, a possibilidade de uso de ferramenta adequada para estas análises em diversos contextos, seja *in loco* onde os animais vivem, sejam em laboratórios específicos, é um requisito importante. Dessa forma, uma arquitetura para o apoio das pesquisas em eficiência alimentar, capaz de processar análises sobre os dados considerando descoberta de informações implícitas, permitindo avaliar os dados sobre diferentes perspectivas, é de grande importância para os avanços das pesquisas nessa área. Além disso, o uso de uma arquitetura orientada a serviços, considerando uma abordagem de *Software as a Service*, é importante neste contexto, permitindo que diferentes aplicações façam uso dessas análises e os resultados possam ser reutilizados em diferentes contextos.

O artigo está organizado em quatro seções. Na seção 2, apresenta-se alguns conceitos importantes e trabalhos relacionados. A seção 3 apresenta a arquitetura proposta. A seção 4 traz as discussões sobre os resultados; e, por fim, a seção 5, as conclusões.

2. Trabalhos Relacionados

Gruber (1993) definiu ontologia como uma “especificação explícita de um conceito”. Guarino, em 1998, por sua vez, descreveu que a ontologia pode ser descrita nas mais variadas formas, dependendo de seu campo de aplicação. Ontologias computacionais são meios para modelar formalmente a estrutura de um sistema, isto é, as entidades relevantes e suas relações que emergem a partir da sua observação e que são úteis para os objetivos desejados.

Em 1999, Baker et al. destacam o crescimento do uso de ontologias no domínio de engenharia de software e aplicações web, com o intuito de promover a integração, interoperabilidade e visualização dos dados. Posteriormente, Miah, Gammack e Kerr (2007) desenvolveram um modelo ontológico capaz de centralizar o acesso aos dados de diversas bases, com objetivo de facilitar a consulta e descoberta de informações de forma mais simplificada aos usuários. Em 2015, Verhoosel, Bekkum e Evert propuseram uma abordagem com uso de ontologias com o propósito de unir bases distintas, proporcionando a análise de um grande volume de dados. Os autores desenvolveram uma ontologia composta de 28 conceitos relacionados a eficiência alimentar.

Tomic et al. (2015), com a arquitetura agriOpenLink, utilizam ontologia para centralizar as consultas e acesso aos dados produzidos nos diversos serviços relacionados a agropecuária, tais como dados climáticos, eólicos, pragas, raças, entre outros. Observa-se, pelos autores citados, a utilização da ontologia para fins exclusivos

de mapeamento. Vale frisar, até o presente momento, nas buscas realizadas em bases indexadas da área da computação, não houveram retornos similares a abordagem proposta neste artigo. Dessa forma, o presente trabalho busca apresentar uma arquitetura capaz de apoiar as pesquisas em eficiência alimentar, a partir do uso de ontologias e análise de dados, utilizando bases de experimentos coletados em campo.

3. Arquitetura *FeedEfficiencyService*

Considerando a necessidade de interoperabilidade entre os dados dos diversos experimentos relacionados a eficiência alimentar e a dificuldade na realização de análises precisas acerca dos dados, incluindo a descoberta de informações implícitas que poderiam apoiar a evolução dos experimentos, a arquitetura *FeedEfficiencyService* foi proposta no contexto das pesquisas em eficiência alimentar da EMBRAPA-Gado de Leite.

Inicialmente, devido aos diversos experimentos já conduzidos estarem armazenados em bases de dados heterogêneas, foi desenvolvida uma camada tradutora (*wrapper*) genérica que permitiu a interoperabilidade dos experimentos já conduzidos para o modelo de dados da arquitetura, conforme apresentado na Figura 1. Esta camada foi especializada para considerar dados específicos dos experimentos analisados. No entanto, caso novos experimentos devam ser analisados, a camada genérica pode ser facilmente especializada para um novo contexto.

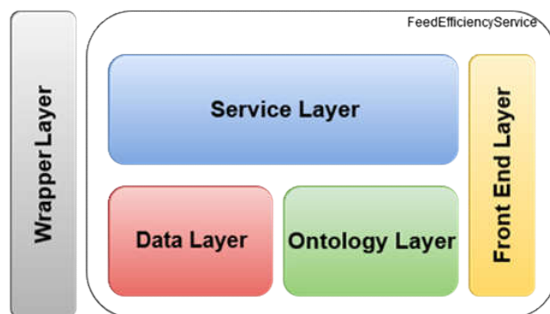


Figura 1. Arquitetura *FeedEfficiencyService*.

Para o armazenamento dos experimentos, um modelo de dados foi especificado, de forma englobar facilmente dados dos novos experimentos. Para apoio a integração e análises dos dados dos experimentos, foi especificada uma ontologia, denominada *Feed Efficiency Ontology* (FEO) (Figura 2). A ontologia permite a integração semântica entre os experimentos relacionados a eficiência alimentar, permitindo aos pesquisadores a classificação dos animais nos experimentos e a interoperabilidade entre os dados, com vistas a realizar análises cruzadas e descoberta de novas conexões entre experimentos.

A linguagem adotada para implementação da ontologia foi o OWL (*Web Ontology Language*), recomendada pelo W3C [Bechhofer et al. 2004; Hitzler, Krotzsch e Rudolph 2009]. Além disso, a ontologia foi especificada para a classificação dos índices de eficiência alimentar. Assim, foram criadas classes específicas e regras específicas para realizar a classificação, considerando os três possíveis níveis: eficiente, intermediário e ineficiente (Figura. 2).

A classificação dos animais se dá através do cálculo de seus respectivos índices, cujas informações foram obtidas junto aos pesquisadores da EMBRAPA Gado de Leite.

Para os cálculos, a arquitetura necessita que o pesquisador insira os dados alimentares do experimento, Ganho de Peso Diário (GPDkg), Ingestão de Matéria Seca (IMS) e Peso Metabólico Médio (PMmediokg) de cada animal participante do experimento. Considerando que a obtenção desses dados é por meio de análises químicas e que essas variam conforme o alimento fornecido, não foi possível a automatização dessa etapa.

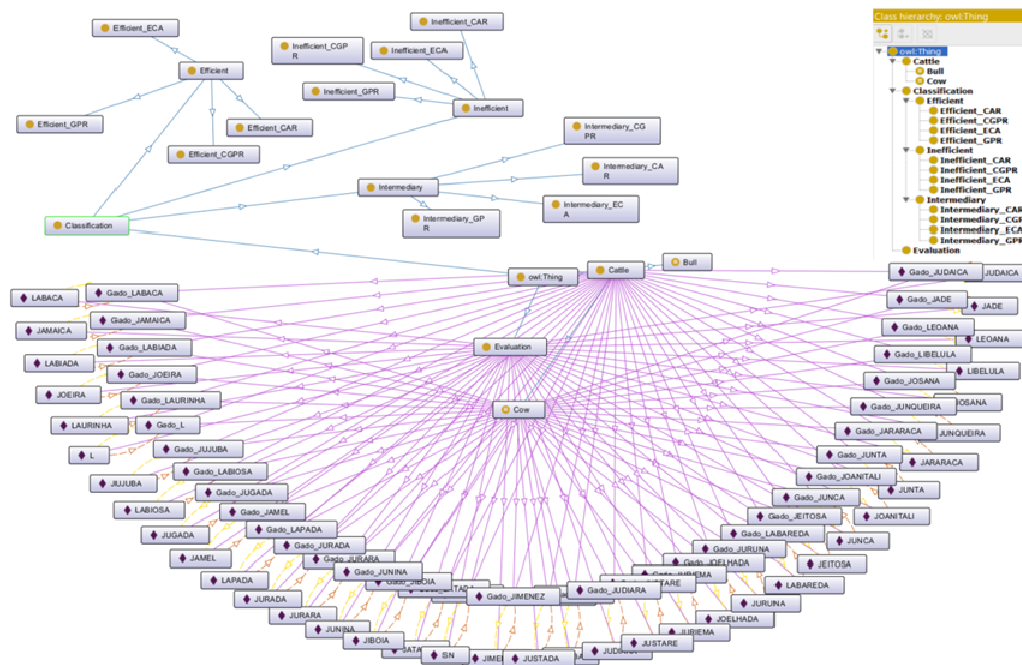


Figura 2. Feed Efficiency Ontology (FEO).

Os valores de IMS Observado e GMD Observado representados na Tabela 1 referem-se aos valores reais coletados durante todo o experimento. Os três índices adotados para o experimento, apesar de fórmulas distintas, baseiam-se nos dados alimentares GPDkg, IMS e PMmediokg para suas apurações, apresentado na Tabela 1.

Tabela 1. Fórmulas para apuração dos índices de eficiência alimentar.

Sigla	Descrição	Fórmula
CAR	Consumo Alimentar Residual	IMS Observado – IMS Esperado
ECA	Eficiência de Conversão Alimentar	IMS/GMD
GPR	Ganho de Peso Residual	GMD Observado – GMD Esperado
CGPR	Consumo e Ganho de Peso Residual	GPD + (CAR * (-1))

A apuração dos valores IMS esperado e GMD esperado é obtida através de regressões lineares múltiplas. Para a classificação automática dos animais segundo os índices e posterior análise dos resultados, foram criadas regras lógicas na ontologia, que permitem a classificação e a identificação de novas relações ontológicas. Para isso, regras especificadas em SWRL (*Semantic Web Rule Language*) foram criadas, considerando os índices CAR, GPR e ECA (Figura 3). As regras SWRL (Figura 3) permitem a composição de associações em busca de novo conhecimento, além da redução da complexidade da consulta à ontologia através do SPARQL. Por exemplo, a regra S1 (Figura 4) tem o papel de classificar os animais eficientes sobre o índice CAR.

Para tal, ela utiliza informações previamente conhecidas, tais como: ser uma instância de *Cattle*, possuir uma avaliação no índice CAR e ter uma avaliação inferior a -0.13. Assim, um animal que possua essas combinações é classificado como CAR eficiente.

Name	Rule
S1	Cattle(?cattle) ^ isEvaluationOf(?cattle, ?evaluation) ^ Experiment CAR(?evaluation, ?EvaluationCAR) ^ swrlb:lessThan(?EvaluationCAR, -0.13) -> Efficient CAR(?cattle)
S2	Cattle(?c) ^ isEvaluationOf(?c, ?v) ^ Experiment CAR(?v, ?EvaluationCAR) ^ swrlb:greaterThan(?EvaluationCAR, 0.13) -> Inefficient CAR(?c)
S3	Cattle(?c) ^ isEvaluationOf(?c, ?v) ^ Experiment CAR(?v, ?EvaluationCAR) ^ swrlb:lessThanOrEqual(?EvaluationCAR, 0.13) ^ swrlb:greaterThanOrEqual(?EvaluationCAR, -0.13) -> Intermediary CAR(?c)
S4	Cattle(?cattle) ^ isEvaluationOf(?cattle, ?evaluation) ^ Experiment GPR(?evaluation, ?EvaluationGPR) ^ swrlb:greaterThan(?EvaluationGPR, 0.0422) -> Efficient GPR(?cattle)
S5	Cattle(?cattle) ^ isEvaluationOf(?cattle, ?evaluation) ^ Experiment GPR(?evaluation, ?EvaluationGPR) ^ swrlb:lessThan(?EvaluationGPR, -0.0422) -> Inefficient GPR(?cattle)
S6	Cattle(?c) ^ isEvaluationOf(?c, ?v) ^ Experiment GPR(?v, ?EvaluationGPR) ^ swrlb:lessThanOrEqual(?EvaluationGPR, 0.0422) ^ swrlb:greaterThanOrEqual(?EvaluationGPR, -0.0422) -> Intermediary GPR(?c)
S7	Cattle(?cattle) ^ isEvaluationOf(?cattle, ?evaluation) ^ Experiment ECA(?evaluation, ?EvaluationECA) ^ swrlb:lessThan(?EvaluationECA, -0.3685) -> Efficient ECA(?cattle)
S8	Cattle(?c) ^ isEvaluationOf(?c, ?v) ^ Experiment ECA(?v, ?EvaluationECA) ^ swrlb:greaterThan(?EvaluationECA, 0.3685) -> Inefficient ECA(?c)
S9	Cattle(?c) ^ isEvaluationOf(?c, ?v) ^ Experiment ECA(?v, ?EvaluationECA) ^ swrlb:lessThanOrEqual(?EvaluationECA, 0.3685) ^ swrlb:greaterThanOrEqual(?EvaluationECA, -0.3685) -> Intermediary ECA(?c)

Figura 3. Regras SWRL.

```
Cattle(?cattle) ^ isEvaluationOf(?cattle, ?evaluation) ^ Experiment CAR(?evaluation, ?EvaluationCAR)
^ swrlb:lessThan(?EvaluationCAR, -0.13) -> Efficient CAR(?cattle)
```

Figura 4. Regra SWRL – S1 : CAR Eficiente .

Foi implementado um serviço web RESTful em JAVA, responsável por disponibilizar serviços para o armazenamento, gerência e consulta aos dados. E, através desse permitir a interoperabilidade com outras aplicações e serviços. Com objetivo de ilustrar a interação entre as tecnologias, um cenário de uso foi elaborado para um melhor entendimento dessa. Assim, a Figura 5 ilustra os passos de uma requisição do pesquisador até o retorno do serviço web com a visualização desejada.

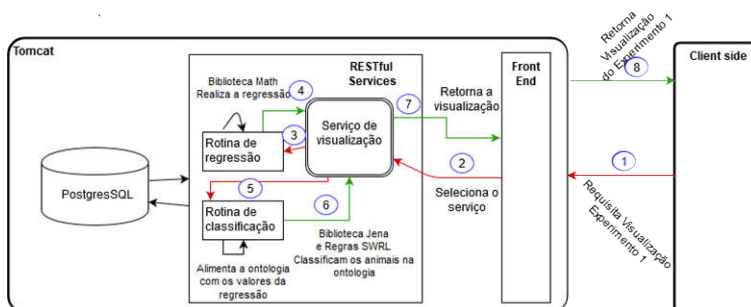


Figura 5. Exemplo de interação entre tecnologias, a fim de retornar a classificação dos animais no experimento.

Foi também elaborado um diagrama de sequência com os passos anteriores (sem uso da arquitetura *FeedEfficiencyService*) (Figura 6) e com o uso da arquitetura (Figura 7), para a classificação dos animais em um experimento. A constante necessidade da intervenção do pesquisador e a inexistência de automação para a transcrição e a análise dos dados produzidos estão presentes no modelo anterior. Nesse sentido, um comparativo entre os experimentos, os animais e os índices era uma tarefa complexa, pois os comparativos ocorriam em planilhas e as análises ficavam a cargo do pesquisador. Assim, quanto maior o volume dos experimentos, animais e índices, mais inviável ficavam as análises.

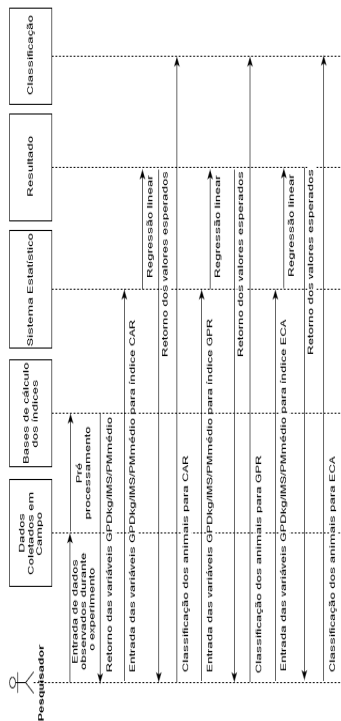


Figura 6. Diagrama de sequência para classificação dos animais, anterior ao uso da arquitetura.

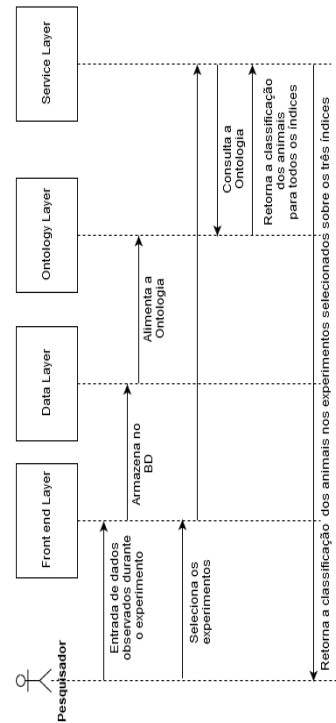


Figura 7. Diagrama de sequência para classificação dos animais utilizando a arquitetura.

Através da utilização da arquitetura, o pesquisador deixa de lado a utilização de inúmeras ferramentas de terceiros e passa a utilizar apenas a arquitetura. Essa tem o papel de centralizar todas as informações dos experimentos e permitir o acesso a todos os pesquisadores da instituição. Para o acesso às classificações, o pesquisador escolhe apenas o experimento ou os experimentos desejados e repassa à arquitetura, ela irá classificar os animais e trazer as análises dos dados

Na próxima seção, são apresentados alguns resultados gerados a partir do uso da arquitetura, considerando um conjunto de dados reais, utilizados em experimentos na Embrapa Gado de Leite.

5. Resultados e Discussões

Como prova de conceito da viabilidade da arquitetura, utilizou-se dados de experimentos relacionados com eficiência alimentar. Mais especificamente, dados do experimento fase 1 – Aleitamento com 30 dias, 56 dias e 80 dias, com animais entre 0 e 3 meses, conforme especificado na Tabela 1. O experimento foi conduzido com 37 novilhas F1 Girolando. Para isso, os serviços da *FeedEfficiencyService* foram utilizados a partir do *Front End Layer*. Os dados do experimento foram carregados, segundo o modelo de dados, e instanciados na ontologia. A partir do uso do *reasoner* e regras lógicas definidas, estes dados relativos a animais do experimento foram classificados

como: eficientes, intermediárias e ineficientes, considerando os índices de eficiência alimentar. A partir dos dados classificados, foi possível produzir saídas de acordo com as necessidades da pesquisa. Por exemplo, a figura 8 apresenta dados do experimento 1, sob o ponto de vista de cada um dos índices de eficiência alimentar. Isso se dá através da utilização visualizações interativas, que quando adotadas permitem a observação de um grande volume de dados. Outro ponto que deve ser destacado é a inclusão na visualização de uma paleta de três cores, sendo elas: verde para eficiente, amarelo para intermediário e vermelho para ineficiente. As cores foram utilizadas na classificação dos animais, bem como na coloração das arestas que relacionavam os animais aos experimentos, como pode ser visto na Figura 8.

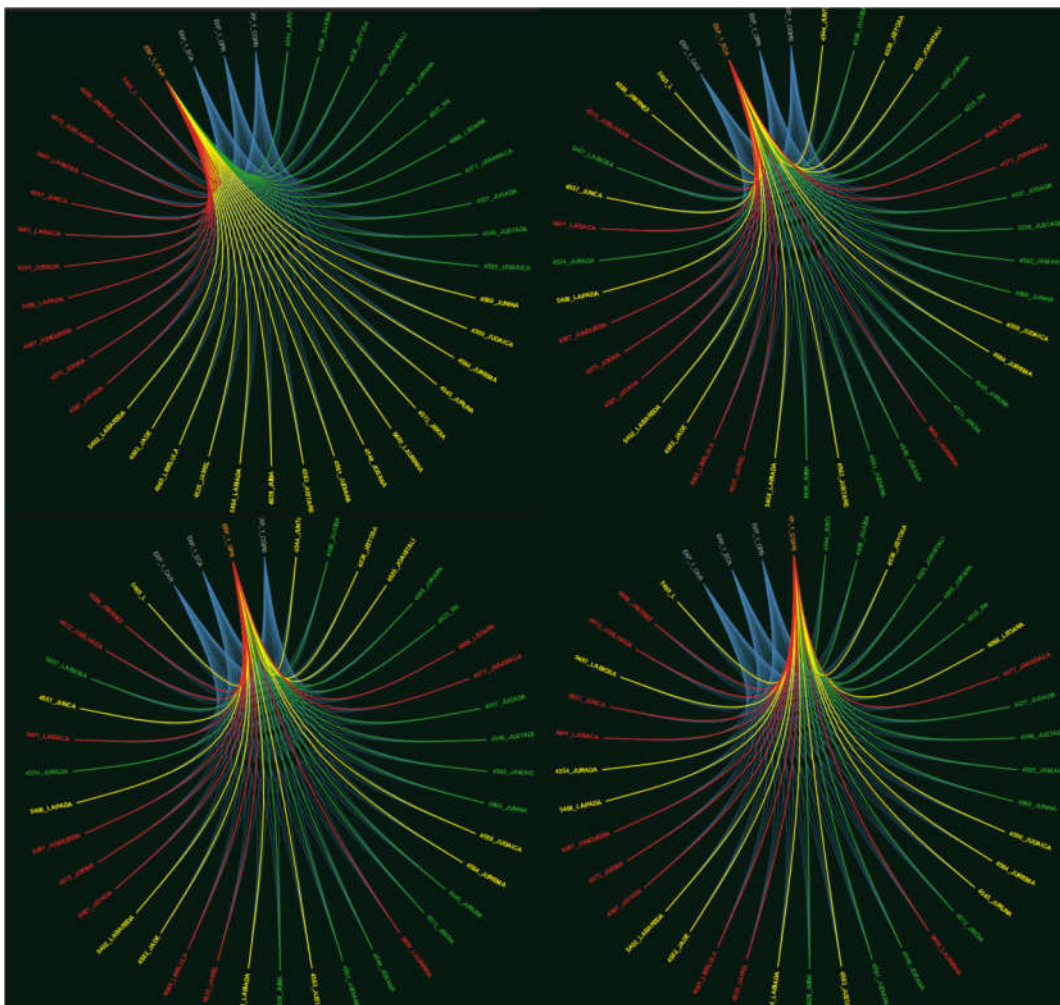


Figura 8. Visualização de classificação sob o ponto de vista do experimento 1 e dos índices de CAR, ECA, GPR e CGPR.

5 Conclusões

A grande competitividade do mercado atual exige que os produtores leiteiros busquem meios para reduzir seus custos e tornar seus rebanhos competitivos. Nesse sentido, o

apoio computacional vem fornecendo alternativas a identificação de animais mais eficientes e, por consequência, proporcionando ganhos econômicos e ambientais.

Observa-se, pelos autores citados, a utilização da ontologia para fins exclusivos de mapeamento. Vale frisar, até o presente momento, nas buscas realizadas em bases indexadas da área da computação, não houveram retornos similares a abordagem descrita. Dessa forma, o presente trabalho buscou apresentar uma arquitetura capaz de apoiar as pesquisas em eficiência alimentar, a partir do uso de ontologias e análise de dados, utilizando bases de experimentos coletados em campo.

A arquitetura visa promover maior poder de análise, de acesso aos experimentos, aumentando a confiabilidade e o reuso desses dados por outros pesquisadores. A partir de uma prova de conceito, pode-se observar indícios de que o uso de ontologias podem contribuir para o avanço das pesquisas em eficiência alimentar na Embrapa Gado de Leite. No entanto, experimentos formais devem ser conduzidos, de forma a comprovar estes indícios.

Como trabalho futuro, pretende-se desenvolver uma rede de ontologias, relacionando todos os experimentos, com o objetivo de observar a manutenção das classificações de eficiência nas fases de avaliação. Outro ponto a ser pesquisado é a busca por padrões nos dados brutos, através de algoritmos de classificação e agrupamentos.

Referências

- Baker, Patricia G.. et al. An ontology for bioinformatics applications. *Bioinformatics*, v. 15, n. 6, p. 510-520, 1999.
- Bechhofer, Sean. OWL: Web ontology language. In: *Encyclopedia of Database Systems*. Springer US p. 2008-2009 (2009)
- Campos, Mariana .M., Leao, Juliana .M., Lima, Juliana .A.M., Machado, Fernanda .S. Tecnologias de precisão na avaliação de eficiência alimentar. *Cadernos Técnicos de Veterinária e Zootecnia*. n79 p. 73-85 (2015)
- Gruber, Thomas R. et al. A translation approach to portable ontology specifications. *Knowledge acquisition*, v. 5, n. 2, p. 199-220 (1993)
- Guarino, Nicola et al. Formal ontology and information systems. In: *Proceedings of FOIS*. p. 81-97 (1998)
- Hitzler, Pascal; Krotzsch, Markus; Rudolph, Sebastian. *Foundations of semantic web technologies*. CRC Press (2009)
- Miah, Shah J.; Gammack, John; Kerr, Don. Ontology development for context-sensitive decision support. In: *Semantics, Knowledge and Grid, Third International Conference on*. IEEE. p. 475-478 (2007).
- Tomic, Dana, et al. "Experiences with Creating a Precision Dairy Farming Ontology (DFO) and a Knowledge Graph for the Data Integration Platform in agriOpenLink." *Journal of Agricultural Informatics* 6.4 (2015).
- Verhoosel, Jack PC; Van Bekkum, Michael; Van Evert, Frits. Ontology matching for big data applications in the smart dairy farming domain. In: *OM*. 2015. p. 55-59.