UNIVERSIDADE DE SÃO PAULO
ESCOLA DE ENGENHARIA DE SÃO CARLOS
FACULDADE DE MEDICINA DE RIBEIRÃO PRETO
INSTITUTO DE QUÍMICA DE SÃO CARLOS

HILTON VICENTE CÉSAR

Mineração de Processos para Extração de Indicadores de Sistema de Informação
para Construção de Matriz de Saúde Mental

São Carlos

2019

HILTON VICENTE CÉSAR

Mineração de Processos para Extração de Indicadores de Sistema de Informação
para Construção de Matriz de Saúde Mental

Tese apresentada ao Programa de Pós-Graduação Interunidades em Bioengenharia da Escola de Engenharia de São Carlos – Faculdade de Medicina de Ribeirão Preto e Instituto de Química de São Carlos da Universidade de São Paulo, como requisito para a obtenção do Título de Doutor em Ciências.

Área de concentração: Bioengenharia

Orientador: Prof. Dr. Paulo Mazzoncini de Azevedo Marques

Coorientador: Prof. Dr. João Mazzoncini de Azevedo Marques

“VERSÃO CORRIGIDA”

São Carlos

2019

AUTORIZO A REPRODUÇÃO TOTAL OU PARCIAL DESTA TRABALHO,
POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS
DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Ficha catalográfica elaborada pela Biblioteca Prof. Dr. Sérgio Rodrigues Fontes da
EESC/USP com os dados inseridos pelo(a) autor(a).

V421m Vicente César, Hilton
Mineração de Processos para Extração de
Indicadores de Sistema de Informação para Construção de
Matriz de Saúde Mental / Hilton Vicente César;
orientador Paulo Mazzoncini Azevedo Marques;
coorientador João Mazzoncini Azevedo Marques. São
Carlos, 2019.

Tese (Doutorado) - Programa de Pós-Graduação
Interunidades em Bioengenharia e Área de Concentração
em Bioengenharia -- Escola de Engenharia de São Carlos;
Faculdade de Medicina de Ribeirão Preto; Instituto de
Química de São Carlos, da Universidade de São Paulo,
2019.

1. Mineração de Processos. 2. OpenEHR. 3. Big
data. 4. Matriz de Saúde Mental. I. Título.

Eduardo Graziosi Silva - CRB - 8/8907

FOLHA DE JULGAMENTO

Candidato(a): Hilton Vicente César

TÍTULO: "Mineração de processos para extração de indicadores de sistema de informação para construção de matriz de saúde mental"

Data da defesa: 28/08/2019

Comissão Julgadora	Assinatura	Resultado
Prof(a). Dr(a). Paulo Mazzoncini de Azevedo Marques - Orientador FMRP/USP		Não votante
Prof(a). Dr(a). Edilson Carlos Caritá UNAERP		APROVADO
Prof(a). Dr(a). Antonia Regina Ferreira Furegato EERP/USP		APROVADO
Prof(a). Dr(a). Alcion Sponholz Junior FAEPA/HCFMRP/USP		APROVADO
Prof(a). Dr(a). Janise Braga Barros Ferreira FMRP/USP		Aprovado
Prof(a). Dr(a). Newton Shydeo Brandão Miyoshi FAEPA/HCFMRP/USP		Aprovado

*A minha família pela compreensão,
carinho e apoio incansável.*

AGRADECIMENTOS

Agradeço a Deus pela inspiração e capacitação por atingir esta jornada de desafios, construção e amadurecimento.

Aos meus pais, Hélio e Líbia, meus irmãos por todo o amor, apoio e que compartilham da minha alegria.

Agradeço ao professor Paulo. Para mim, é uma imensa honra tê-lo como orientador. Levo comigo os seus ensinamentos, conselhos sempre centrados e oportunos.

Ao professor e coorientador João pela oportunidade de trabalho e aprendizado. Ao amigo e professor Domingos, pelo acolhimento e oportunidade em poder colaborar em seus projetos de pesquisa, afinal assim se deu a minha vinda a cidade de Ribeirão Preto.

Aos amigos do Centro de Ciências da Imagem e Física Médica: Rafael Reis, Lucas Phigreto, Rodrigo Pessini, Júlio Batista, José Raniery, Jaqueline, Saulo Cordeiro e Gerson Neto que compartilharam os momentos de dificuldades e alegrias do dia a dia.

A Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pelo auxílio financeiro que fez dessa jornada algo possível.

Para finalizar gostaria de agradecer ao programa de pós-graduação interunidades em Bioengenharia da Escola de Engenharia de São Carlos, Faculdade de Medicina de Ribeirão Preto e Instituto de Química de São Carlos da Universidade de São Paulo, em especial a Márcia, pela atenção e presteza, sempre auxiliando os alunos com muita boa vontade e, a Agência USP de Inovação pelo grande apoio.

EPÍGRAFE

“As invenções são, sobretudo, o resultado de
um trabalho teimoso.”

(Santos Dummond)

RESUMO

CÉSAR, H. V. **Mineração de processos para extração de indicadores de sistema de informação para construção de matriz de saúde.** 2019. 119 f. Tese (Doutorado) – Programa de Pós-Graduação Interunidades em Bioengenharia da Escola de Engenharia de São Carlos, Faculdade de Medicina de Ribeirão Preto, Instituto de Química de São Carlos, Universidade de São Paulo, São Carlos, 2019.

A implementação e o uso contínuo de Sistemas de Informação em Saúde (SIS) são hoje prioridades para se viabilizar um cuidado integral à saúde dos cidadãos. Processos informatizados são cada vez mais utilizados para garantir uma atenção à saúde com qualidade, nos diferentes níveis de complexidade. A gestão de grandes volumes de dados, chamados de *big data*, gerados em serviços clínicos, a possibilidade de seu acesso no momento e local necessários ao cuidado e o potencial de obtenção de indicadores de qualidade da atenção a partir da análise dos grandes repositórios de informação e seu uso no apoio à gestão, são algumas das inúmeras vantagens da informatização dos registros de atendimento. A disponibilidade da informação de saúde, no local em que a assistência é realizada, pode modificar o panorama de saúde seja em uma região, país ou, até mesmo, de forma globalizada. Nesse contexto, uma linha de investigação atual na área da informática médica é o uso da Mineração de Processos (MP) para extração de indicadores com vista a apoiar a tomada de decisão na gestão da saúde, por meio da estruturação de uma matriz de saúde. A hipótese é que o SIS, no caso particular da pesquisa aqui apresentada representado pelo Sistema de Informação em Saúde Mental (SISAM-13), possui os dados necessários para a construção da matriz de saúde mental. Porém, eles não estão organizados para esse fim e a MP se apresenta como uma boa opção para evitar a necessidade de retrabalho manual humano, extraíndo os indicadores necessários a partir dos logs que representam os processos realizados no SIS. Os métodos utilizados na investigação contemplam: o desenvolvimento de scripts de banco de dados para extração de registros do SISAM-13, utilização do aplicativo de MP open source DISCO para organização de logs, geração de *big data* de indicadores de saúde em conformidade com o modelo canônico openEHR e seleção de indicadores em conformidade com o modelo da matriz. Os eventos analisados estão compreendidos no período de 2013 a 2016. Os resultados obtidos indicam o potencial do uso da MP para apoio a estruturação da matriz de saúde mental.

Palavras-chave: Matriz de saúde mental. Mineração de processo. Big data. OpenEHR.

ABSTRACT

CÉSAR, H. V. Mining of processes for extraction of information system indicators for construction of health matrix. 2019. 119 f. Tese (Doutorado) – Programa de Pós-Graduação Interunidades em Bioengenharia da Escola de Engenharia de São Carlos, Faculdade de Medicina de Ribeirão Preto, Instituto de Química de São Carlos, Universidade de São Paulo, São Carlos, 2019.

The implementation and continuous use of health information systems (SIS) are today priorities to enable integral and integrated care for citizens' health. Computerized processes have been increasingly used to ensure quality health care at different levels of complexity. The management of data sets that are too large, called big data, generated in clinical services, the possibility of its access at the time and place necessary for care, and the potential to obtain quality of care indicators from the analysis of large information repositories and its use in management support, are some of the many advantages of the informatization of healthcare records. The availability of health information in the point of care can change the health landscape whether in a region, country or even globally. In this context, a current subject of research in the area of medical informatics is the use of process mining (PM) to extract indicators to support decision making in healthcare management, through the structuring of a mental health matrix. The hypothesis is that the SIS, in the particular case of the research presented here represented by the Mental Health Information System (SISAM-13), has the required data for the construction of the such matrix. However, they are not organized for this purpose and PM is a good option to avoid the need for human manual rework, extracting the necessary indicators from the logs that represent the processes performed into the SIS. The material and methods used in the research include: the development of SISAM-13 log extraction database scripts, use of the DISCO open source MP application for log organization, big data generation of health indicators in compliance with the openEHR canonical model and indicators selection in accordance with the matrix model. The period of events analyzed took into consideration the years 2013, 2014, 2015 and 2016. The results obtained indicate the potential use of PM to support the structuring of the mental health matrix.

Keywords: Mental health matrix. Process mining. Big data. OpenEHR.

LISTA DE ILUSTRAÇÕES

Figura 1 – Modelo de matriz de saúde mental adaptado à XIII Rede Regional de Atenção à Saúde (DRS XIII)	21
Figura 2 – Visão geral da MP.....	24
Figura 3 – MP em saúde.....	25
Figura 4 – Ciclo de vida de um projeto de MP.....	26
Figura 5 – Mapa mental geral de arquétipo	30
Figura 6 – Construção de Arquétipos por Especialistas Médicos.....	32
Figura 7 – Modelo EAV.....	36
Figura 8 – Banco de dados orientado a documentos.....	38
Figura 9 – Estrutura de armazenamento do MongoDB.....	38
Figura 10 – Etapas do ETL.....	40
Figura 11 – Etapas dos matérias e métodos utilizados	47
Figura 12 – Arquitetura conceitual do SISAM-13.....	48
Figura 13 – ETL para SISAM-13.....	49
Figura 14 – Arquétipo de entrada de dados: openEHR-EHR-CLUSTER.mental_health_matrix_inputs.v1.adl.....	51
Figura 15 – Arquétipo de processamento de dados: openEHR-EHR-CLUSTER.mental_health_matrix_process.v1.adl.....	52
Figura 16 – Arquétipo de saída de dados:openEHR-EHR-CLUSTER.mental_health_matrix_process.v1.adl.....	52
Figura 17 – Formulário clínico especializado para saúde mental.....	53
Figura 18 – <i>Template</i> para matriz de saúde mental.....	57
Figura 19 – Modelagem DATALAYER SISAM-13.....	60
Figura 20 – Exemplo de geração de indicadores no nível do paciente.....	63
Figura 21 – Exemplo de geração de indicadores no nível do paciente.....	64
Figura 22 – Exemplo de geração de indicadores no nível do paciente.....	65
Figura 23 – Exemplo de geração de indicadores no nível do profissional.....	66
Figura 24– Exemplo de geração de indicadores no nível do departamento regional de saúde.....	67
Figura 25 – Esquema de dados orientado a documentos utilizando MongoDB.....	69
Figura 26 – Documento openEHR, exemplo de episódio de internação, no repositório <i>big data</i> para saúde mental.....	70

Figura 27 – Exemplo de <i>Time line</i> do paciente.....	73
Figura 28 – Exemplo de <i>Time line</i> do paciente.....	75
Figura 29 – Exemplo de <i>Time line</i> do paciente.....	76
Figura 30 – Desenho dos processos clínicos.....	79
Figura 31 – Eventos dos cuidados de emergência.....	79
Figura 32 – Eventos dos cuidados de emergência.....	80
Figura 33 – Modelo de componentes de software para melhoria de processos.....	81
Figura 1 do Apêndice – Desenho dos processos clínicos.....	104
Figura 2 do Apêndice – Eventos dos cuidados de emergência.....	104
Figura 3 do Apêndice – Eventos dos cuidados de emergência.....	105
Figura 4 do Apêndice – Modelo de componentes de software para melhoria de processos...	105

LISTA DE TABELAS

Tabela 1 – Termos que são empregados em elementos dos bancos e dados relacionais e seus correspondentes no banco de dados MongoDB.....	39
Tabela 2 – Atributos demográficos do paciente.....	61
Tabela 3 – Atributos de episódios de internação.....	62
Tabela 4 – Atributos de comorbidades do paciente.....	62

LISTA DE ABREVIATURAS E SIGLAS

ADL	–	<i>Archetype Definition Language</i>
API	–	Interface de Programação de Aplicações
CAIS	–	Centro de Atenção Integrada à Saúde
CFM	–	Conselho Federal de Medicina
CID	–	Código Internacional de Doenças
CSV	–	<i>Comma-Separated-Values</i>
DRS	–	Departamento Regional de Saúde
EVA	–	Entidade, Atributo, Valor
HAS	–	Hipertensão Arterial Sistêmica
HTML	–	<i>Hypertext Markup Language</i>
IS	–	Informática em Saúde
JSON	–	<i>Java Script Object Notation</i>
MR	–	Modelo de Referência
MOA	–	Modelo de Objetos de Arquétipos
MI	–	Modelo de Informação
MP	–	Mineração de Processos
OO	–	Orientação a Objetos
RAPS	–	Rede de Atenção Psicossocial
RES	–	Registro Eletrônico em Saúde
SBIS	–	Sociedade Brasileira de Informática em Saúde
SIS	–	Sistema de Informação em Saúde
SIM	–	Sistema de Informações sobre Mortalidade
SINAN	–	Sistema de Informação de Agravos de Notificação
SGBD	–	Sistema de Gerenciamento de Banco de Dados
SQL	–	<i>Structured Query Language</i>
UML	–	<i>Unified Modeling Language</i>

SUMÁRIO

1 INTRODUÇÃO.....	17
1.1 Contextualização.....	17
1.2 Definição do problema.....	18
1.3 Objetivo.....	19
1.4 Motivação.....	19
1.5 Hipótese.....	19
1.6 Organização da tese.....	20
2 FUNDAMENTAÇÃO TEÓRICA.....	21
2.1 Matriz de Saúde Mental.....	21
2.2 Mineração de Processos.....	23
2.3 Padronização de processos clínicos.....	28
2.4 <i>Big data</i> em saúde.....	34
2.5 MongoDB.....	37
2.6 ETL.....	40
2.7 Gestão da informação na atenção à saúde mental.....	42
3 MINERAÇÃO DE PROCESSOS NA ATENÇÃO À SAÚDE.....	45
4 MATERIAL E MÉTODOS.....	47
4.1 Descrição técnica (infraestrutura para obtenção de logs)	47
4.2 Aplicativo DISCO.....	50
4.3 Modelagem DUAL openEHR.....	50
4.4 API GDC-Arquetype e EDC-Templates para gerenciamento e execução de definições clínicas.....	53
4.5 EASY- EHR.....	55
5 RESULTADOS.....	59
5.1 Modelagem do SISAM-13 para geração dos logs (reengenharia de software)	59
5.2 Mineração de processos aplicada aos logs gerados pelo modelo.....	62
5.3 <i>Big data</i> para saúde mental.....	68
5.4 API para gestão da informação e criação da <i>timeline</i>	71
6 DISCUSSÃO.....	77
6.1 Trabalhos Relacionados.....	77
6.2 Contribuições.....	81

6.3 Limitações e Trabalhos Futuros.....	82
7 CONCLUSÃO.....	83
REFERÊNCIAS.....	85
APÊNDICE A – Script para obtenção de log de interações.....	95
APÊNDICE B – Script para checagem e tratamento dos dados.....	96
APÊNDICE C – Codificação referente ao arquétipo: openEHR-EHR-CLUSTER.mental _health_matrix_inputs.v1.adl.....	97
APÊNDICE D– Codificação referente ao arquétipo: openEHR-EHR-CLUSTER.mental _health_matrix_process.v1.adl.....	101
APÊNDICE E– Codificação referente ao arquétipo: openEHR-EHR-CLUSTER.mental _health_matrix_outputs.v1.adl.....	103
APÊNDICE F – Aplicativo DISCO.....	105
APÊNDICE G – GDC-ARQUETYPES.....	107
APÊNDICE H – EDC-TEMPLATES.....	113
APÊNDICE I – Modelo do formulário clínico openEHR para extração dos indicadores e estruturação da matriz de saúde.....	115
APÊNDICE J – Código fonte parcial de classe que realiza comunicação com <i>big data</i> ..	116
APÊNDICE L – Trecho parcial da codificação openEHR.....	117
ANEXO A – Parecer do comitê de ética em pesquisa.....	123

1 INTRODUÇÃO

Neste Capítulo, são descritas as principais motivações que apoiam o uso de modelos e conceitos das áreas computacionais de Mineração de Processos (MP) e *big bata* como ferramenta de repositório e tratamento de dados para o apoio à gestão na atenção à saúde mental. Os objetivos desta investigação são então apresentados, indicando a relevância do estudo realizado.

1.1 Contextualização

A saúde é um campo de investigação interdisciplinar e atuação multiprofissional, no qual se busca utilizar conhecimentos gerados a partir de dados reais para apoiar a tomada de decisão. A Informática em Saúde (IS) é uma área da ciência que lida, entre outros temas, com recursos, procedimentos e mecanismos para propiciar o armazenamento, a recuperação e análise de grandes Bancos de Dados (BD), visando apoiar a tomada de decisão por especialistas. De acordo com Gartner (2011), BD faz referência ao grande volume de dados que demandam formas inovadoras e rentáveis de processamento da informação, para melhor percepção e gestão do conhecimento.

Dentre diferentes abordagens possíveis, como *machine learning*, *text process* e *option mining*, a investigação de um BD na área da saúde pode ser realizada por meio de técnicas de mineração de processos ou por meio da identificação e classificação de eventos de processos, através da análise do registro do histórico de eventos que ocorrem em um SIS, capturado no formato de *logs*. (MEDEIROS, 2006). Dentre as principais vantagens em relação as abordagens anteriormente citadas estão:

- A extração e organização não supervisionada de conhecimento a partir de dados textuais;
- Possibilidade de auxiliar processos automáticos de recuperação de informação de procedimentos, a partir dos descritores associados nos *logs*;
- Descoberta de padrões, tendências e correlações em fragmentos de registros de *logs* de sistemas de informação;

Nesse contexto, a proposta da investigação aqui apresentada buscou avaliar a aplicabilidade e efetividade do uso da MP como ferramenta de extração e organização de dados (indicadores) para o apoio à gestão na atenção à saúde mental, tendo como referência o Sistema

de Informação em Saúde Mental (SISAM-13) (MIYOSHI et al., 2018; YOSHIURA et al., 2017), que integra dados das ações de atenção à saúde mental no Departamento Regional de Saúde de Ribeirão Preto - DRS XIII.

1.2 Definição do problema

Na região de Ribeirão Preto os serviços de saúde que oferecem apoio na área de saúde mental trabalharam de forma adequada até 2004. A partir desse período a demanda em atenção à saúde mental aumentou progressivamente e a coordenação das informações geradas nos diferentes serviços tornou-se problemática. Até então, todas as regulações na área de saúde mental eram feitas de modo descentralizado e sem o apoio de um sistema logístico informatizado, o que dificultava todo o processo de referência e contra referência, aumentando o tempo de espera para o tratamento (YOSHIURA et al., 2017).

Com o estabelecimento da portaria do Ministério da saúde Nº 3.088, de 23 de dezembro de 2011 (BRASIL, 2011) que instituiu a RAPS (Rede de Atenção Psicossocial), os serviços se organizaram e foi iniciado o desenvolvimento e a implantação do Sistema de Informação em Saúde Mental (SISAM-13) (YOSHIURA et al., 2017).

Embora o uso do SISAM-13 tenha propiciado uma melhor estruturação da coordenação dos serviços e dos processos na atenção à saúde mental e, por consequência, organizado a RAPS, os dados existentes nos diversos sistemas de informação que oferecem apoio nos diferentes níveis de atenção à saúde mental ainda não estavam integrados. Visando minimizar as dificuldades na continuidade do cuidado dos usuários da RAPS, foi projetada, desenvolvida e testada uma plataforma de interoperabilidade em saúde, denominada eHealth-Interop (MYIOSHI et al., 2018).

Para tanto foi adotada uma proposta de interoperabilidade por meio de serviços Web com um modelo de integração de dados baseado em um banco de dados centralizador. No total foram integrados dados de 31340 registros de pacientes existentes no SISAM-13, no sistema e-SUS AB de Santa Cruz da Esperança, no sistema de informação do Centro de Atenção Integrada à Saúde (CAIS) de Santa Rita do Passa Quatro e do sistema de informação do Hospital Santa Tereza e do Barramento do CNS do Ministério da Saúde.

Com a implementação da plataforma eHealth-Interop e sua integração com SISAM-13 o DRS XIII passou a contar com um conjunto integrado de dados sobre atenção à saúde mental, que pode ser considerado um BD. Considerando esse cenário, a pesquisa aqui apresentada teve como foco a investigação da aplicabilidade do conceito e dos métodos computacionais de

Mineração de Processos como ferramenta de apoio à gestão da atenção à saúde mental, por meio da extração de indicadores para construção de uma matriz de saúde (THORNICROFT; TANSELLA, 1999). Até onde conhecem os autores, não foi encontrado nenhum estudo anterior que realizasse esse tipo de estudo.

1.3 Objetivos

O objetivo desse estudo foi investigar a aplicabilidade do conceito e dos métodos computacionais de mineração de processos como ferramenta de apoio à gestão da atenção à saúde mental, por meio da extração de indicadores para construção de uma matriz de saúde.

1.4 Motivação

A motivação é usar a mineração de processos para extração de indicadores na construção da matriz de saúde (THORNICROFT; TANSELLA, 1999). O número crescente de aplicações de informática para a área da saúde tem aumentado a complexidade para modelagem de dados, fazendo com que os dados do paciente estejam armazenados em várias bases e em formatos distintos, geralmente utilizando um padrão proprietário. Além disso, o domínio de saúde é reconhecido pela quantidade e complexidade de conceitos existentes, assim como pela velocidade com que estes se alteram (GODOI, 1995). Nesse contexto, a MP pode ser uma ferramenta de apoio à gestão do processo de atenção à saúde.

1.5 Hipótese

O presente trabalho estuda a hipótese de que a MP pode ser uma ferramenta adequada para a extração de indicadores para a construção da matriz de saúde, a partir de *logs* de eventos existentes no sistema de informação. A hipótese se apoia no entendimento dos autores Orellana, Pérez e Alfonso (2018) de que os SIS possuem uma grande quantidade de dados, mas não dispõem de mecanismos para analisar a execução de seus processos. As técnicas de mineração de processo analisam os dados existentes nos sistemas e podem ser utilizadas para a detecção da variabilidade na execução de processos de atenção à saúde.

1.6 Organização da Tese

Esta tese apresenta 7 capítulos, 1 anexo, 11 apêndices e está estruturada da seguinte forma:

Capítulo 1 – Introdução;

Capítulo 2 – Fundamentação teórica: são apresentadas e discutidas as bases conceituais relacionadas com o estudo;

Capítulo 3 – Mineração de processos na atenção à saúde;

Capítulo 4 – Material e métodos da pesquisa;

Capítulo 5 – Resultados: são apresentadas as simulações dos processos e os resultados obtidos;

Capítulo 6 – Discussões bem como as limitações encontradas e expectativas de trabalhos futuros;

Capítulo 7 – Conclusão;

2 FUNDAMENTAÇÃO TEÓRICA

Este capítulo aborda os temas considerados relevantes para o entendimento do objeto de estudo, introduzindo os conceitos relacionados a matriz de saúde, mineração de processos, padronização de processos clínicos, BD em saúde. Por fim, são apresentados princípios relacionados com a gestão da informação na atenção à saúde mental e API Google Maps para rastreabilidade na rede de saúde.

2.1 Matriz de Saúde Mental

A matriz de saúde adaptada para a área de atenção à saúde mental é o elemento primordial empregado como modelo de referência para a Rede de Atenção Psicossocial (RAPS) na cidade de Ribeirão Preto/SP. Elaborada conceitualmente e publicada inicialmente pelos autores Thornicroft e Tansella no ano de 1999, a matriz de saúde mental tem o seu uso voltado para suprir a necessidade de se desenvolver modelos de sistemas de apoio logístico para integração de informações, bem como coordenação de atendimento nas redes de saúde.

A matriz estabelece um modelo que pode ser usado para aumentar a eficácia clínica através da prática baseada em evidências, estruturada em dois eixos: temporal e espacial (geográfico). A matriz também permite avaliar as áreas que necessitam de melhoria nos serviços e é útil para toda a linha de operação e para apoiar o processo de decisão. A Figura 1 apresenta o modelo de matriz de saúde mental adaptada à XIII Rede Regional de Atenção à Saúde (DRS XIII) (VINCI et al., 2016).

Figura 1 – Modelo de matriz de saúde mental adaptado à XIII Rede Regional de Atenção à Saúde (DRS XIII)

Fonte: Adaptado de Vinci et al. (2016).

O eixo temporal da matriz de saúde é composto por indicadores do processo de atenção à saúde dos tipos dados de entrada (*inputs*), processamento (*process*) e saída (*outputs*) acerca do acompanhamento dos pacientes. Os dados de *input* retratam a visão clínica do caso: o profissional que realizou a solicitação da internação, diagnóstico primário e diagnóstico secundário, a data que ocorreu a internação, o serviço que realizou a solicitação da internação, bem como informações demográficas do paciente, como o código de identificação, sexo, idade e estado civil. Além disso, são observados registros de comorbidades durante o processo de atenção à saúde.

Dados que compõem o processamento (*process*) dizem respeito ao tempo entre a liberação da vaga e a internação do paciente no hospital, o tempo de duração de internação, a unidade onde o paciente foi internado e, existindo transferência intra-hospitalar, o registro também traz esse dado.

A dimensão final do modelo temporal, os “*outputs*”, trabalha com dados psicossociais relacionados ao destino final do paciente ao retratar para onde o paciente foi após a sua alta, a data e a hora da alta, o diagnóstico primário e diagnóstico secundário, tipo da alta e o profissional que realizou a alta. Os indicadores *inputs*, *process* e *outputs* podem ser investigados na matriz de saúde tanto individualmente quanto em conjunto. A avaliação dos dados em conjunto é capaz de projetar ao gestor de saúde uma visão da linha do tempo do processo de cuidado clínico e psicossociais dos pacientes.

No âmbito da dimensão espacial (geográfico), a matriz de saúde produz indicadores relacionados com os níveis de movimentação, destino e serviços utilizados: usuário, profissionais de saúde, serviços de saúde, cidades, regiões de saúde e departamento regional de saúde. Assim a matriz disponibiliza a possibilidade de realizar mapeamento da atenção à saúde que pode ser definido inclusive com base em uma posição georreferenciada.

Os indicadores da matriz de saúde traduzem os indícios da ocorrência de um determinado fato, podendo responder às perguntas formuladas localmente, visto que são resultados da composição de um conjunto de dados coletados em algum momento, nos mais variados lugares e espaços.

A análise temporal da matriz de saúde na RAPS atualmente tem seu monitoramento feito por representantes de serviços de saúde mental e entidades governamentais por meio de comitês municipais, regionais e estaduais. No entanto, não existem arranjos ou indicadores padronizados para a avaliação e gestão do RAPS instituídos pela legislação governamental. Os únicos pactos existentes referem-se à implementação, financiamento e padrões operacionais para serviços específicos da RAPS (VINCI et al., 2016).

No Brasil, existem SIS consolidados, a exemplo do Sistema de Informações sobre Mortalidade (SIM) e do Sistema de Informação de Agravos de Notificação (SINAN), que disponibilizam dados para análises de situação de saúde. Contudo, o georreferenciamento desses dados em cidades brasileiras ainda é um desafio. Entre os problemas mais comuns destacam-se a qualidade dos dados, a falta de Sistemas de Informações Geográficas (SIG) municipais estruturados, a falta de bases cartográficas digitais e cadastros oficiais de endereços, além da complexidade da infraestrutura urbana, principalmente, em áreas carentes como ocupações irregulares e favelas, onde é comum a não padronização dos endereços (SILVEIRA; OLIVEIRA; JUNGUER, 2017).

Os SIG são usados para descrever, analisar e prever padrões espaciais, dispondo de diversas aplicações em epidemiologia, como o mapeamento de doenças, a investigação de surtos e análises espaciais. Oferece também oportunidade de acompanhamento da atenção recebida pelo sujeito de toda a rede básica de saúde. Seu uso depende, sobretudo, da disponibilidade de dados georreferenciados. O georreferenciamento, por sua vez, é o processo pelo qual informações textuais descritivas de uma localidade, como o endereço relacionado a um evento de saúde, são convertidas em representações geográficas válidas. Ele pode ser realizado por meio da associação a um par de coordenadas geográficas ou a unidades espaciais, como setores censitários, bairros, municípios, entre outras (GOLDBERG; WILSON; KNOBLOK, 2007).

Os dados necessários para a construção conceitual da matriz de saúde estão presentes no SISAM-13, representados através de *logs* dos registros dos processos de atenção à saúde e podem ser utilizados para a construção de indicadores, assim como, para o georreferenciamento. O tópico a seguir traz os conceitos de ciência de dados e mineração de processos que se apresentam como solução alternativa possível ao retrabalho manual para extração de indicadores para a construção da matriz e seu georreferenciamento das pessoas atendidas em toda rede de saúde bem como da gestão desses serviços.

2.2 Mineração de Processos

A MP é um campo de pesquisa que surgiu inicialmente em 1999, a partir de um projeto de pesquisa desenvolvido por Wil Van Der Aalst e Ton Weijters (VAN DER AALST, 2011). Os mesmos autores, posteriormente, em Van Der Aalst e Weijters (2004), definem MP como “uma metodologia para destilar uma descrição estruturada de processo a partir de um conjunto de execuções reais”. A MP assume que é possível coletar um *log* de processos contendo a ordem

em que os eventos de cada instância são executados. Pode ser usado de maneira mais ampla para se referir não apenas a técnicas para descobrir modelos de processo, mas também a técnicas de auditoria de conformidade de processos de negócios e análise de desempenho com base em *logs* de eventos. Assim, em linhas gerais, a MP permite estabelecer um *link* entre os modelos de processo e a “realidade”, representada por uma sequência de eventos em um sistema de informação, conforme esquematizado na Figura 2.

Figura 2 – Visão geral da MP

Fonte: adaptada de Van Der Aalst (2011).

A MP é constituída por uma família de técnicas no campo do gerenciamento de processos que suporta a análise de processos de negócios com base em logs de eventos. Durante a mineração de processo, os algoritmos de mineração de dados especializados são aplicados aos dados do registro de eventos para identificar tendências, padrões e detalhes contidos nos registros de eventos presentes em um sistema de informação. A mineração de processo visa melhorar a compreensão dos processos, possibilitando atuar para aumentar sua eficiência.

As instituições em geral têm se mostrado cada vez mais eficientes para capturar, organizar e armazenar grandes quantidades de dados, obtidos de suas operações diárias. Porém, a maioria ainda não utiliza adequadamente essa quantidade de dados para transformá-la em conhecimento, que possa ser utilizado para melhoria de suas próprias atividades. (SONG; GÜNTHER, AALST, 2008)

Técnicas tradicionais de descoberta de processos incluem *brainstorming*, entrevistas, análise de documentos, observação passiva ou ativa e amostragem do trabalho (SCHEDLBAUER, 2010). Estas técnicas são manuais, não envolvendo uma análise rigorosa de dados previamente existentes (VAN DER AALST, 2011). Além disso, elas demandam uma grande quantidade de recursos humanos e de tempo.

Ainda segundo Van Der Aalst (2011), a modelagem tradicional está também sujeita a diversos erros. Por exemplo, um analista que projeta um modelo geralmente se concentra no comportamento padrão do processo, deixando de fora, por exemplo, os 20% de casos menos frequentes, os quais podem ser aqueles que tendem a ser mais problemáticos durante a execução do processo, embora sejam mais raros.

Um *log* de eventos pode ser obtido a partir de qualquer sistema de informação que armazena estes dados. Para que seja possível realizar atividades de MP, é necessário que cada registro de evento contenha informações sobre a instância na qual este evento ocorreu, à qual atividade ele corresponde e em que momento ele foi registrado. Partindo desse conjunto de dados, diversas técnicas têm sido investigadas, desenvolvidas e publicadas na literatura especializada.

A Figura 3 mostra um esboço geral da aplicação da mineração de processo em saúde. Conforme Rojas et al. (2016), normalmente, qualquer atividade realizada em um hospital por um médico, enfermeiro, técnico ou com base em qualquer outro recurso hospitalar para cuidar de um paciente é armazenada em um Sistema de Informação Hospitalar (*Hospital Information System – HIS*), que é composto por bancos de dados, programas de computadores, protocolos, eventos, entre outros. No HIS, as atividades são registradas em *logs* de eventos para suporte, controle e análise adicional.

Os modelos de processo são criados para especificar a ordem em que diferentes profissionais de saúde devem realizar suas atividades dentro de um determinado fluxo de trabalho ou para analisar criticamente o projeto do processo. Além disso, os modelos de processo também são usados para apoiar o desenvolvimento do HIS, por exemplo, para entender como o sistema de informação deve suportar a execução de um determinado processo.

Figura 3 – MP em saúde

Fonte: adaptado de Rojas et al. (2016).

Para Bernard et al. (2016), os sistemas de informação capturam atividades que acontecem no “mundo real” e, para se executar a mineração de processo nesses dados, os

rastreamentos digitais capturados nos sistemas de informação devem ser extraídos e transformados em *logs* de eventos. Existem três requisitos mínimos para isso:

- Primeiro deve ser possível definir um nome de atividade para etapas bem definidas;
- Segundo deve ser possível identificar casos distintos. O caso fornece os meios para agrupar instâncias do mesmo processo, enquanto a atividade descreve etapas bem definidas;
- Terceiro as atividades devem estar disponíveis de maneira ordenada;

Além disso, a análise pode ser ampliada enriquecendo os eventos com informações adicionais, como uma indicação de qual recurso está executando determinada atividade ou qualquer outro dado relevante relacionado ao caso.

O modelo do ciclo de vida de um projeto de mineração de processo pode ser dividido em cinco etapas: planejar e justificar; extrair; criar o modelo de fluxo de controle e conectar o *log* de eventos; criar o modelo de processo integrado; e, fornecer suporte operacional. Ele é usado como a metodologia sugerida para orientar projetos de mineração de processo, especificamente no *Cluster* de Conformidade Médica Baseado em Evidências (CCMBE), conforme Figura 4 (BERNARD et al., 2016).

Figura 4 – Ciclo de vida de um projeto de MP

Fonte: adaptado de Bernard et al. (2016).

Ao realizar um estudo sobre os principais *softwares* para realizar MP que analisa e visualiza processos de negócios com base em dados de tecnologia da informação elencamos:

- Celonis: é uma tecnologia inteligente de *big data* que analisa e visualiza todos os processos em qualquer empresa. Empresas que utilizam o *software*: Siemens, Uber, L'Oréal, Cisco e Vodafone. Disponível em: <https://www.celonis.com>;
- ARIS *Process Mining*: é projetado para revelar pontos fracos e tornar os processos mais transparentes, mais rápidos e mais econômicos. Disponível em: <https://www.ariscommunity.com/>;
- Minit: fornece *insights* visuais sobre os processos de negócios e revela como aumentar as receitas, a economia e a eficiência. Disponível em: <https://www.minit.io>;
- *Apromore Advanced Process Analytics*: é uma plataforma de análise de código aberto que oferece recursos avançados de MP. Disponível em: <http://apromore.org/>;
- Disco: Transforma os dados da empresa em valor criando mapas de processos interativos de como pessoas, máquinas e *software* se inter-relacionam. Disponível em: <https://fluxicon.com/disco/>. O uso do aplicativo, dispensa experiência prévia ou conhecimento específico do negócio, a única informação necessária é o *log* de eventos. As etapas de uso do aplicativo estão representadas no Apêndice F;
- Kofax Insight: proporciona um monitoramento e análise de processos com visualizações, análises e integração de dados para oferecer visibilidade completa do processo. Disponível em: <https://www.kofax.com>;
- LANA Process Mining.: detecta possíveis riscos e pontos fracos em seus processos de negócios. Disponível em: <https://lana-labs.com/en/>;
- ProM: é uma estrutura extensível que suporta uma ampla variedade de técnicas de mineração de processo na forma de plug-ins. Disponível em: <http://www.promtools.org/>;

O objetivo da MP é fornecer informações baseadas em fatos sobre a execução de processos. Um aspecto importante disso é a descoberta de modelos de processo baseados no comportamento capturado em dados de eventos. Esses modelos geralmente mostram as atividades que podem ser executadas durante o processo e como elas são ordenadas.

Um dos aspectos mais importantes da descoberta de processos é deduzir os estados do processo operacional no *log*. Muitos algoritmos de mineração de dados só têm uma noção implícita de estado, ou seja, o foco está em aprender a ordenação de atividades. No entanto, as informações do estado do processo podem, na verdade, estar presentes explicitamente nos sistemas de informação. Exemplos de tal informação explícita do estado são o diagnóstico de um paciente em um processo de assistência médica.

A etapa posterior pós cadastro, é o uso do aplicativo. Esta etapa consiste de seis fases, a saber:

- Preparar o Log: o log de eventos é extraído do sistema de informação;
- Inspeccionar o Log: para ganhar entendimento inicial do processo;
- Analisar o fluxo de controle: para conseguir visualizá-lo de uma forma ilustrativa e amigável;
- Analisar desempenho: para que, por exemplo, os possíveis gargalos sejam identificados;
- Analisar papéis: para identificar os recursos que realizam as tarefas e se trabalham em conjunto;
- Transferir resultados: repassando-os apropriadamente ao cliente para que ele consiga obter um melhor entendimento do SI envolvido, e assim torná-lo mais eficiente ou ágil quando aplicável;

2.3 Padronização de processos clínicos

Embora ainda haja questionamentos dos pesquisadores da área da Saúde Mental quanto à confiabilidade acerca dos registros relacionados aos transtornos mentais nos SIS brasileiros, estudos vêm apontando o seu uso como um meio simplificado para a obtenção de dados sobre a situação dos transtornos mentais e comportamentais, permitindo aos profissionais formular e executar ações importantes na área (PEREIRA et al., 2012). Tais dados, quando padronizados podem se tornar um importante objeto de investigação de diversos estudos.

No campo da saúde, a evolução tecnológica, com a adoção da padronização de registros clínicos destaca-se principalmente quando provê uma melhoria na prestação de assistência, com a redução de erros e a duplicidade de dados, além de reduzir o custo com investigações diagnósticas desnecessárias (BRAILER, 2005; HEARD, 2010; PIRES, 2007).

Dentro do escopo de otimização de processos a arquitetura openEHR provê uma padronização semântica no contexto da saúde, representando a habilidade dos sistemas de reconhecerem o significado da informação que compartilham, permitindo uma melhor comunicação, podendo ser usada até por sistemas mais antigos, chamados de sistemas legados (AVILA, 2011). Tal padronização é alcançada por meio do uso de arquétipos e templates que representam conceitos e documentos clínicos, que podem ser partilhados por várias entidades.

O openEHR é baseado no paradigma de modelagem de informação em dois níveis, no qual o primeiro nível representa o Modelo de Referência (MR) e o segundo compreende as definições formais e o conteúdo clínico (componente semântico). Este segundo nível é baseado em arquétipos e modelos de conteúdo (OPENEHR, 2019). O ambiente utilizado para compilar a estrutura de um MR se faz com a adoção de uma linguagem de programação. Em OpenEHR (2019), as plataformas: Java, JWSDP, WebSphere, AXIS e “.NET” são recomendadas para esta finalidade.

Como consequência deste paradigma, os sistemas assim construídos, são codificados de forma mais simples, com menos linhas de programação e apresentam maior facilidade para manutenção do que os sistemas modelados apenas em um nível, nos quais modelo e definições estão integrados na mesma estrutura. Estes sistemas também são facilmente adaptáveis, possibilitando a adição de novos conteúdos na forma de arquétipos.

O Modelo de Referência (MR), utilizado pela Fundação openEHR, representa as características globais dos componentes do registro em saúde, como eles são agregados, assim como a informação contextual necessária para atender os requisitos éticos, legais e de proveniência. Este modelo define o conjunto de classes que formam os blocos genéricos para construir um Registro Eletrônico em Saúde (RES), permitindo a separação do conteúdo clínico e do conteúdo demográfico (OPENEHR, 2019).

A separação entre os dois níveis da arquitetura openEHR permite também o envolvimento direto de profissionais de saúde no desenvolvimento dos sistemas de informação, algo já identificado como necessário na realidade brasileira (BRASIL, 2011a; 2011b), uma vez que toda a implementação das regras de negócio e procedimentos para manipulação dos dados, que caracterizam a arquitetura, fica separada da definição do conteúdo clínico do sistema. Com o paradigma da modelagem multinível, a parte central do sistema é baseada em um modelo extremamente estável: o MR.

Existem implementações de software que possibilitam a construção de SIS apenas por meio do conteúdo do sistema, utilizando arquétipos e templates, ficando sob responsabilidade do software todo o desenvolvimento e implementação do sistema funcional (NATIONAL E-

HEALTH TRANSITION AUTHORITY - NEHTA, 2011; OPENEHR, 2019; ROTZSCH, 2010). Estes softwares permitem que os profissionais definam os conteúdos dos formulários dos sistemas, que serão modelados utilizando-se arquétipos para os conceitos em saúde necessários, e posteriormente agrupados em formulários, que formarão de fato as fichas eletrônicas de cadastro de informações.

A arquitetura openEHR atende aos requisitos da ISO 18308. Essa ISO define um conjunto de requisitos para a arquitetura de um RES. Do ponto de vista de documento, os arquétipos compõem as entradas, que são agrupadas em seções e estas, por sua vez, em composições. As pastas representam os agrupamentos de diversas composições segundo algum critério: especialidade, evento, etc. A Figura 5 apresenta a representação do mapa mental de um arquétipo clínico geral disponibilizado pela fundação openEHR.

Figura 5 – Mapa mental geral de arquétipo

Fonte: openEHR-CKM (2019).

Conforme OpenEHR-CKM (2019), o mapa mental geral de arquétipo é composto por:

- *Observation*: classe que engloba tudo que é dito pelo paciente (sintoma, evento, preocupação), achados de exames e resultados de testes ou medidas. Estes dados podem ser registrados repetidamente ao longo do tempo mesmo dentro de uma *composition*. O conteúdo de uma *observation* é composto por quatro partes: *data* (dados), *protocol* (protocolo), *state* (estado do paciente) e *Events* (eventos). As *observations* geram opiniões por parte dos profissionais de saúde;
- *Evaluation*: Esta classe de arquétipo registra as interpretações das observações. Estão incluídos os achados interpretados clinicamente, opiniões, resumos clínicos, bem como ideias, rotulações, ou visões a partir da mente do clínico.

São exemplos de *evaluations*: “diagnóstico”, “avaliação e indicadores de risco”. Seu conteúdo é composto por dados e protocolos;

- *Instruction*: Instruções sobre como dar seguimento aos cuidados de saúde são modeladas por esta classe de arquétipos. São exemplos de *instructions*: “prescrições”, “solicitação de exames” e “recomendações”;
- *Action*: Os conceitos das atividades clínicas e os seus resultados são realizados com o arquétipo da classe *action*. Os arquétipos desta classe são muitas vezes confundidos com *observation*, a diferença é que as *actions* são a uma intervenção e geralmente precedidos por uma *instruction*. São exemplos de *action*: um procedimento cirúrgico, a administração de medicamento intravenoso. O seu conteúdo inclui a descrição de uma atividade (*activity*), o caminho a ser percorrido para sua execução (*pathway*) e o protocolo;
- *Admin*: Os arquétipos desta classe não são relevantes clinicamente. Eles servem para guardar informações administrativas sobre o processo clínico, como os dados administrativos da admissão hospitalar (p ex. seguro de saúde);
- *Composition*: A maneira mais fácil de entender esta classe de arquétipos é considerá-la como documentos clínicos comumente utilizados (p. ex. relatório de alta);
- *Section*: Tem a finalidade de auxiliar e organizar o conteúdo do arquétipo em compartimentos específicos, separando-os em sessões e subseções. Em exemplo das classes *section*: “exame físico” e “dados vitais”;
- *Structure*: Classe de arquétipos responsável pela organização da estrutura de dados necessária para o registro de informação;
- *Cluster*: São arquétipos que podem ser utilizados dentro de outros arquétipos dos tipos descritos anteriormente. Um *cluster* é um fragmento de um arquétipo. Os conceitos clínicos que são classificados como um *cluster* são aqueles que podem ser reutilizados em diversos contextos;
- *Element*: Classe responsável pelos elementos para entrada de dados, como por exemplo pressão sistólica, pressão diastólica;
- *Generic Entry*: São classes de arquétipos que possibilitam realizar a entrada de dados de modo genérico como por exemplo: data e hora;

- Demgraphic Model archetypes: tem a finalidade de proporcionar uma modelagem de dados demográficos do paciente, como por exemplo, a idade, sexo, data de nascimento, estado civil e endereço completo do paciente;

A Fundação openEHR propõe uma modelagem de dois níveis ou multinível para os sistemas de informação. A principal característica oferecida pela modelagem de dois níveis é a separação entre informação e conhecimento. A informação é modelada pelo Modelo de Referência (MR), enquanto o conhecimento é modelado usando o Modelo de Objetos de Arquétipos (MOA). Na modelagem de dois níveis, o Modelo de Objetos do openEHR é dividido no MR, representando a visão da informação, e o Modelo de Objetos de Arquétipos (MOA) representando o modelo de domínio (conhecimento) por meio de arquétipos (BEALE, 2008).

Um modelo de informação ou também chamado de modelo de referência é requerido, e este é especificado pelo padrão openEHR. O termo modelo de informação (MI) foi utilizado nesta tese. Tomando como elementos constituintes os MI, que permitem representar conceitos clínicos gerais, chega-se ao modelo de conhecimento do padrão openEHR, que tem por objetivo representar conceitos clínicos particulares (ao contrário do MI tradicional, que representa conceitos clínicos gerais e reside dentro do software).

Os conceitos clínicos particulares são representados como um conjunto de restrições sobre o MI genérico. A abordagem do openEHR, através da modelagem de dois níveis, habilita especialistas clínico a determinar as características do registro de saúde mais adequadas para suas necessidades, ilustra a separação entre a atividade dos especialistas de domínio (neste caso, em específico, a área médica) na criação dos arquétipos através do modelo genérico de informação e que, por sua vez, irão compor as bases de conhecimento clínico conforme a Figura 6.

Figura 6 – Construção de Arquétipos por Especialistas Médicos.

Fonte: Gutiérrez e Carrasco (2013)

Na arquitetura proposta pelo padrão openEHR, o MR é responsável pela semântica da informação que será trabalhada via restrições na camada de conhecimento (BEALE, 2008). Assim, através da implementação dos modelos de informação definidos no modelo de referência, é possível expressar informações clínicas conforme a semântica openEHR nas diferentes plataformas de desenvolvimento de softwares. No caso específico deste trabalho de pesquisa, escolheu-se a plataforma Java.

A modelagem openEHR fornece especificações de SIS utilizando notações que representam conceitos da abordagem de Orientação a Objetos (OO). A OO é uma abordagem para desenvolvimento de software que trata os problemas do mundo real como um conjunto de objetos distintos. Leva-se em conta a estrutura e o comportamento dos dados para essa representação. A OO possui características específicas como identidade, classificação, encapsulamento, herança, polimorfismo e persistência (BEALE, 2008).

Descrito sob a ótica da restrição aos objetos do MR, os arquétipos podem ser vistos conforme Martínez-Costa et al. (2010): o modelo de referência define o conjunto de classes que formam os blocos genéricos de construção do registro eletrônico de saúde e contém suas características não voláteis. Arquétipos aplicam restrições aos objetos, que são descritores dos níveis ontológicos do domínio, definidos num modelo de referência. Os arquétipos fazem a ponte entre a generalidade dos conceitos do modelo de referência e a variabilidade da prática clínica, tornando-se assim, uma ferramenta para representar estes conceitos.

O Modelo de Objeto de Arquétipos (MOA) é um modelo genérico que pode ser usado para expressar arquétipos de qualquer modelo de referência. Este modelo pode ser usado como base para desenvolvimento de sistemas que se baseiam na modelagem multinível, citado anteriormente, e que processam arquétipos de forma independente. Como também, pode ser usado para desenvolver parsers, que interpretam arquétipos em um formato de linguagem de programação, como o *Archetype Definition Language* (ADL) (BEALE, 2008).

Arquétipos fornecem uma modelagem semântica e são expressos sob a forma de restrições (*constraints*) nos dados em que as instâncias estão em conformidade com o MR. O MOA define as relações que devem ser verdadeiras entre as partes de um arquétipo para que seja válido como um todo. Ou seja, todos os arquétipos devem estar em conformidade com MOA. Arquétipos podem ser compostos para formar estruturas maiores equivalentes semanticamente a um arquétipo grande. Isso permite aos arquétipos serem definidos de acordo com níveis naturais ou encapsulamentos de informação, e também permite a reutilização de arquétipos menores (BEALE, 2008).

Conforme Peçanha e Bax (2015), o modelo de arquétipos, portanto, pode ser visto como uma representação de metadados desenvolvida para organizar e padronizar dados de domínios de conhecimento. Através dos arquétipos os conceitos clínicos são capturados de modo estruturado fora do software. Ao se desenvolver o SIS como um sistema baseado em conhecimento, obtém-se um aplicativo robusto, de mais fácil manutenção e modificação, cujo acréscimo de conhecimentos é feito via arquétipos, que são a expressão computável de um conceito em nível de domínio na forma de declarações de restrições estruturadas, baseada em algum modelo de informação de referência.

2.4 *Big data* em saúde

O *big data* mudou a forma como gerenciamos, analisamos e aproveitamos dados em qualquer setor. Uma das áreas mais promissoras em que pode ser aplicado para fazer uma mudança é a saúde. A análise dos serviços de saúde tem o potencial de reduzir os custos do tratamento, prever surtos de epidemias, evitar doenças evitáveis e melhorar a qualidade de vida em geral (ZHANG et al., 2009).

Big data é um conceito que vem ganhando espaço e interesse na medicina, por conta do aumento das soluções em IS, com a implementação de modelos de e-Saúde (modelo de atenção à saúde integral e integrado com base nas Tecnologias da Informação e Comunicação – TIC) e a implantação de Sistemas de Informação em Saúde (SIS). Nas últimas décadas, o avanço tecnológico afetou consideravelmente a forma como as organizações e profissionais da área da saúde lidam com suas informações. Segundo Groves et al. (2014), o uso de *big data* na área da saúde pode ser categorizado em 5 tipos diferentes:

- Administração sobre Registros Eletrônicos de Saúde (RES), e informações baseadas em evidências observadas;
- Suporte à decisão clínica: os métodos de *big data* ajudam a lidar com o processo de tomada de decisão relacionado à medicina;
- Extração de informações clínicas sobre um conjunto de dados disponíveis a partir de vários tipos de sistemas de informação para análise de dados;
- Avaliação de comportamento: neste caso, técnicas de *big data* são usadas para análise sociodemográfica;

- Extração de informações gerais para suporte à ação: quando o uso de *big data* que não pode ser colocado em nenhuma das quatro categorias acima, é agrupado nesta categoria. A eficácia do tratamento, registros laboratoriais e de pesquisa, intervenções de saúde pública, medicamentos, detecção de erros, etc., estão localizados nesta categoria;

A investigação na área da saúde com adoção e utilização de *big data* se revela como uma área importante, compartilhada e trabalhada por diversos especialistas (como cientistas de dados e epidemiologistas), podendo atuar em diferentes tipos de pesquisas, por exemplo sobre causalidade, eventos adversos de medicamentos, avaliação de intervenções, impactos de políticas públicas, estudo de doenças raras e outras.

Soluções em *big data* são projetadas para trabalhar de forma distribuída, facilitando a adição de novos pontos de processamento e armazenamento de informações em tempo de execução, automatizando a distribuição e balanceamento de dados. Isto permite que a solução se adeque a crescente demanda de usuários. O *big data* suporta o armazenamento, acesso, disponibilidade, escalonamento e análise dos dados.

Um dos maiores obstáculos no caminho para usar um banco de dados na medicina é o fato dos dados médicos, em geral, estarem distribuídos por muitas fontes administradas por diferentes estados, hospitais e departamentos administrativos. A integração dessas fontes de dados exigiria o desenvolvimento de uma nova infraestrutura na qual todos os provedores de dados colaborassem entre si. Os recursos de *big data* são frequentemente caracterizados originalmente por volume, velocidade, variedade, variabilidade e veracidade (GARTNER, 2011).

Os desafios de "volume" geralmente podem ser abordados com a disponibilidade e a capacidade atual de armazenamento de informações. O desafio de "velocidade" ainda está por ser resolvido porque a maioria desses recursos de *big data* é estática e atualizada periodicamente, com poucos aplicativos "em tempo real" atualmente desenvolvidos; no entanto, essa situação mudará se os aplicativos de suporte à decisão forem implementados e / ou se o aprendizado de máquina e a inteligência artificial começarem a ser incorporados nos sistemas de registros. "Variedade" e "variabilidade", como velocidade, não são desafios atuais, mas provavelmente se tornarão cada vez mais importantes no futuro próximo. A "veracidade" continua a ser uma consideração fundamental e é improvável que se altere com os avanços tecnológicos, porque o uso de dados secundários continua a depender dos dados realmente registrados na prática clínica.

Big data está sendo usado para pesquisa em saúde mental em muitas partes do mundo e para muitos propósitos diferentes (STEWART; DAVIS, 2016). O uso secundário de dados administrativos, especialmente onde informações de diagnóstico de rotina são incluídas, provavelmente se tornará cada vez mais adotado para pesquisa, uma vez que esses recursos de informação são relativamente de baixo custo e escaláveis.

Na modelagem de dados tradicional, cada atributo do objeto que se deseja armazenar é representado por uma coluna no esquema do banco de dados. No entanto, essa modelagem é bastante inflexível e não acompanha o crescimento da aplicação, de modo que qualquer alteração dos atributos provoca uma modificação do esquema do banco de dados (PATRICK; TRUNAN, 2006). Segundo Beale (2008), essa abordagem não atende as necessidades do meio de saúde, e a abordagem EAV (*Entity, Attribute, Value*) foi utilizada na modelagem do esquema proposto na investigação. A utilização dessa abordagem é útil em um banco de dados relacional ou objeto-relacional, no entanto, estudos precisam ser realizados para uma melhor aplicação da referente modelagem.

Conceitualmente a modelagem EAV pode ser representada por três tabelas no banco de dados: tabela Entidade ligada à tabela Atributos que por sua vez é ligada a uma tabela de Valor. A relação entre as tabelas está representada na Figura 7. Na modelagem EAV, os atributos que eram armazenados em colunas na modelagem convencional são armazenados em linhas nas tabelas, permitindo um vasto armazenamento de dados, *big data*, em apenas algumas tabelas.

Figura 7 – Modelagem EAV.

Fonte: Patrick, Trunan (2006)

A importância de *big data* na saúde é imensa, mas ela não gira apenas em torno da quantidade de dados que uma instituição possui, mas sim o que cada uma extrai deles. O fundamental é compreender como isso impacta no dia a dia do gestor clínico: essa quantidade de informações pode ser analisada e processada para gerar insights e ajudar gestores em decisões estratégicas de que forma é possível utilizar melhor as informações que fluem para dentro das organizações para reduzir custos, otimizar tempo e melhorar o atendimento aos pacientes.

Segundo Patrick e Trunan (2006) o modelo EAV pode apresentar problemas de desempenho em um ambiente operacional. A principal causa desses problemas é a natureza genérica das estruturas de dados presentes no modelo de referência do openEHR. Inserção, recuperação e validação são as principais causas de problemas de desempenho nessa abordagem. A obtenção de informação da base de dados se torna dispendiosa devido à grande quantidade de pré e pós-processamento na decomposição e recomposição das estruturas de dados genéricas (PATRICK; TRUNAN, 2006).

A abordagem EAV possui a vantagem de representação das informações de forma genérica, além de ter um esquema bastante simplificado em relação a abordagens tradicionais, facilitando a manutenção de banco de dados. Devido à grande importância no desempenho dos sistemas, o armazenamento de dados e sua estrutura física representada no repositório de dados são classificados como bancos não relacionais, como é o exemplo do modelo de banco de dados orientado a objetos NoSQL (ANICETO; XAVIER, 2014). O paradigma de persistência de dados NoSQL é uma solução alternativa para os bancos de dados relacionais, são muito flexíveis, possuem alta escalabilidade e disponibilidade, além de sua alta performance.

2.5 MongoDB

Existe uma grande quantidade de bancos de dados *Not Only SQL*(NoSQL) no mercado, que trabalham de diferentes maneiras, como os bancos de dados orientados a documento, chave-valor, grafos e colunas. O MongoDB trabalha com documentos no formato *Java Script Object Notation* (JSON), para armazenar dados. Também fornece os recursos necessários para um ambiente de produção: balanceamento de carga, replicação, indexação, consulta e pode atuar como um sistema de arquivos.

Os sistemas de banco de dados NoSQL, diferentemente dos relacionais, são pensados e otimizados para serem executados em *clusters*, de modo que são mais apropriados para grandes quantidades de dados (SADALAGE; FOWLER, 2013). Ou seja, diferentemente dos bancos de dados SQL, os bancos de dados NoSQL foram pensados e projetados para utilizar computação distribuída. Esses sistemas de bancos de dados também permitem que sejam armazenadas estruturas mais complexas do que simples tabelas.

MongoDB é um banco de dados orientado a documentos. Foi escrito na linguagem C++ e possui algumas características: a chave que identifica um documento normalmente é a combinação do ID e um valor timestamp; implementa uma linguagem própria de consulta. Suas

características permitem que as aplicações modelem informações de modo muito mais natural, pois os dados podem ser aninhados em hierarquias complexas e continuar a ser indexáveis e fáceis de buscar.

Modelar uma estrutura para o MongoDB é mais simples do que uma modelagem para um banco de dados relacional, porém esse trabalho pode ficar complicado se pensarmos de forma relacional (MONGODB, 2019). A Figura 8 apresenta o modelo do banco de dados orientado a documentos MongoDB.

Figura 8 – Banco de dados orientado a documentos.

Fonte: MongoDB (2019)

O modelo disponibilizado pelo MongoDB oferece técnicas de distribuição de documentos em vários servidores através da replicação de dados, permitindo master-slave, que é mais utilizado em caso de falha de leitura ou escrita, e sharding, utilizado para a escalabilidade horizontal, o que garante a durabilidade dos dados. A Figura 9 apresenta a estrutura básica do banco de dados MongoDB.

Figura 9– Estrutura de armazenamento do MongoDB.

Fonte: MongoDB (2019)

Na abordagem relacional, alguns termos são utilizados para identificar os elementos que compõe a estrutura dos bancos de dados relacionais. É possível fazer uma analogia com o

MongoDB e mostrar a terminologia empregada nos elementos que compõe a estrutura do banco de dados. A Tabela 1 apresenta os termos que são empregados em elementos dos bancos e dados relacionais e seus correspondentes no banco de dados MongoDB. Alguns termos, tais como index, primary key e banco de dados tem o mesmo significado em ambos, enquanto as informações dos registros armazenados (tabela, coluna e linha) são diferentes (MONGODB, 2009).

O banco de dados MongoDB utiliza mecanismos diferentes de manipulação de dados em comparação aos bancos de dados relacionais. É possível realizar uma comparação sintática de como é feita a criação de uma relação, inserção e consulta dos dados em um banco de dados relacional e no banco de dados MongoDB, conforme Tabela 1.

Tabela 1 – Termos que são empregados em elementos dos bancos e dados relacionais e seus correspondentes no banco de dados MongoDB.

Termos SQL	Termos MongoDB
<i>Banco de dados</i>	Banco de dados
<i>Tabela</i>	Coleção
<i>Linha</i>	Documento ou JSON
<i>Coluna</i>	Campo
<i>Index</i>	Index
<i>Primary Key</i>	Primary Key

Fonte: O autor.

A categoria banco de dados orientado a colunas é o modelo que mais se assemelha com bancos de dados relacionais, que armazena as informações em relações, mas não permite relacionamentos entre elas. Os dados são armazenados separadamente em colunas e cada coluna de dados é um índice do banco de dados. As colunas podem ser agrupadas em famílias de colunas, que permitem sua distribuição em vários nós, o que torna importante para a organização e particionamento dos dados (HAAS, 2013).

Segundo Hass (2013), um BD é um conjunto de dados organizados, relacionados entre si e referentes a algum contexto organizacional. O projeto de um BD deve seguir algumas regras de construção, conhecidas como modelagem. A modelagem do BD pode ser comparada à documentação do projeto de construção de uma casa, que inclui sua planta, seu projeto elétrico e hidráulico, paisagismo, entre outros. O BD não é somente o conjunto de arquivos de computador armazenados em disco; ele consiste, também, de documentos que descrevem esses

arquivos e que são associados a eles, além de um sistema de gerenciamento de banco de dados, o SGBD.

Conforme Hass (2013), para que esses arquivos sejam criados é necessária a edição e execução de uma série de comandos, em linguagem de programação e consulta para banco de dados SQL, a partir da interface de um Sistema de Gerenciamento de Banco de Dados (SGBD), que interpreta tais comandos e os executa criando assim as estruturas lógicas e binárias necessárias para receber os dados e mantê-los armazenados, organizados e com todas as características pertinentes a um banco de dados consistente.

2.6 ETL

A sigla ETL significa Extração, Transformação e Carga (em inglês *Extract, Transform and Load*) e visa trabalhar com toda a parte de extração de dados de fontes externas, transformação para atender às necessidades de negócios e carga dos dados dentro do BD (KIMBALL; CASERTA, 2004). ETL é uma das etapas mais importantes do projeto, sendo necessária bastante atenção na integridade dos dados a serem carregados em um BD (RIBEIRO, 2011). A Figura 10 apresenta cada uma das etapas do ETL que serão descritas a seguir:

Figura 10– Etapas do ETL.

Fonte: O autor.

Para Elias e Ribeiro (2011), a coleta de dados dos sistemas de origem, que podem ser outras bases de dados relacionais ou arquivos, extraindo-os e transferindo-os para a “área de preparação” (área de transição ou área temporária), onde o processo de ETL pode operar de forma independente. É nesta etapa que são realizados os devidos ajustes dos dados extraídos, aplicando uma série de regras ou funções para assim melhorar a qualidade dos dados e

consolidar dados de duas ou mais fontes. Algumas fontes de dados necessitarão de muita pouca manipulação de dados. Em outros casos, diversos tipos de transformações são feitos para atender as regras do negócio. Dependendo das necessidades da organização, o tempo de execução deste processo pode variar bastante, podendo ter dados atualizados semanalmente ou a cada meia hora.

No processo tradicional de processamento, os dados são extraídos das suas fontes através de uma ferramenta de ETL, passando por uma fase de transformação e limpeza, para posteriormente serem inseridos num BD (LEBDAOUI; ORHANOU; ELHAJI, 2014). De acordo com o período da sua extração (diário, mensal, anual, entre outros), os dados são refrescados, permitindo aos utilizadores ter dados atuais e atualizados com informação útil para o processo de tomada de decisão, havendo sincronismo da informação atualizada com a que está no BD. A necessidade de uma análise de dados com mais brevidade motivou a redução da latência existente no processo de ETL, pois é esta que impossibilita a disponibilização de serviços em tempo real (FREUDENREICH et al., 2013; VAISMAN; ZIMÁNYI, 2012)

O processo ETL tem que ser repetido periodicamente, de forma a atualizar a informação disponível até ao momento. A periodicidade dependerá da taxa de refrescamento dos dados em questão, sendo que o sistema terá que ser suficientemente eficiente para que estes refrescamentos não prejudiquem o normal funcionamento do sistema de acesso a um DB. É desejável que a periodicidade do refrescamento seja cada vez menor para que a informação analisada reflita mais convenientemente a realidade (WAAS et al., 2013).

Após o processo de ETL as informações são transferidas para área de apresentação dos dados. Nessa área é realizada a organização, o armazenamento e a disponibilidade de consultas, relatórios e outras aplicações. Geralmente a área de apresentação dos dados é referida como *data mart*. Um *data mart* representa um subconjunto de dados de um projeto de BD classificados segundo os temas seleccionados de um determinado negócio. (CHAUDHURI; DAYSAL, 2000)

Conforme Santos et al. (2001) os *data marts* devem conter dados detalhados e atômicos, sumarizados e agregados. Com o conceito de *data mart*, o BD pode ser compreendido como sendo a base de dados que representa integridade dos mesmos. Diversos *data marts* são construídos de maneira isolada para depois serem integrados e formar um amplo BD.

2.7 Gestão da informação na atenção à saúde mental

Muitos centros de atenção à saúde mental estão presentemente confrontados com a necessidade de implementar um sistema de informação de gestão. Para Lowe e Sugarman (1978), se um centro escolhe adaptar um sistema existente ou desenvolver um sistema próprio, um planejamento cuidadoso antes da implementação do sistema pode ajudar a garantir que ele atenda às necessidades do centro e opere com sucesso. As diretrizes são organizadas nas categorias de considerações sobre dados, considerações sobre pessoas e considerações sobre o sistema. As duas primeiras categorias são de interesse geral, enquanto a última categoria é mais técnica por natureza.

Desde 1988, um complexo sistema nacional de saúde (altamente descentralizado) vem sendo implementando no Brasil com responsabilidades no financiamento e gestão divididos entre os três níveis federados de âmbito nacional (cidades, estados e país). O sistema nacional de saúde também é composto por uma combinação de serviços públicos e privados, com fins lucrativos e sem fins lucrativos. Como em outros países, no Brasil a prevalência de transtornos mentais é alta e a comorbidade entre transtornos mentais e problemas de saúde não psiquiátricos é comum. Além disso, uma vez que a integração regional entre os serviços ainda é limitada e extremamente desafiadora, há uma necessidade urgente de desenvolver estratégias/soluções de integração apropriadas e viáveis para os serviços de saúde mental. (MIYOSHI et al., 2018)

Segundo a World Health Organization - WHO (2005, p.13), um Sistema de Informação em Saúde Mental (SISM) é “um sistema para coletar, processar, analisar, disseminar e usar informações sobre um serviço de saúde mental e as necessidades de saúde mental da população que atende. Além disso, conforme Jamal, Mckenzie e Clark (2009), um SISM, pode ajudar a desenvolver sistemas de cuidados adequados e equitativos dentro de um contexto específico e pode ajudar a melhorar a comunicação do médico, a adesão às normas de melhores práticas e a redução de erros no atendimento clínico. Ainda segundo WHO (2005), o SISM afeta, portanto, o planejamento, o orçamento e a avaliação em todos os níveis de atenção à saúde. Também é útil para apresentar dados de saúde mental de forma compreensível e acessível às partes interessadas - incluindo provedores de serviços, usuários de serviços, formuladores de políticas e a população em geral.

Informações sobre transtornos específicos, sua prevalência e o resultado de seu tratamento em contextos específicos são fundamentais na formulação de políticas e planejamento, a fim de compreender a situação de saúde mental em qualquer país. Uma das formas mais bem-sucedidas de obter essas informações é através do uso de sistemas de informações de gerenciamento de saúde, nos quais as informações e os planos de cuidados do

paciente são inseridos uniformemente em um banco de dados para facilitar ainda mais o gerenciamento do paciente (KPOBI; SWARTZ; OFORI-ATTA, 2018).

Embora existam ferramentas como o Instrumento de Avaliação da OMS para Sistemas Mentais de Saúde (WHO-AIMS) e o Atlas da Saúde Mental da OMS, que poderiam ser utilizados periodicamente para avaliar o progresso em direção às metas de cuidado acordadas internacionalmente pelos estados membros da OMS (WHO, 2005), pesquisas ainda recentes indicam que a informação sobre saúde mental é frequentemente da mais baixa prioridade em sistemas de informação de gestão de saúde (SIGS) em países de baixa e média renda.

Os indicadores de saúde mental estão ausentes ou são minimamente incluídos no monitoramento rotineiro da saúde e geralmente se limitam ao atendimento ambulatorial e à taxa de ocupação por pacientes internados (JORDANS et al., 2016). Mesmo assim, iniciativas como a publicação do Atlas de Saúde Mental da OMS em 2001, tem possibilitado identificar que 60% dos países associados apontam para um conjunto de cinco indicadores principais para avaliação da garantia da qualidade na atenção à saúde mental, sendo eles: existência de políticas e leis específicas para a saúde mental, existência de programas de promoção e prevenção voltados para a saúde mental, disponibilidade de serviços e força de trabalho em saúde mental.

Para Kpobi, Swartz e Ofori-Atta (2018), um SISM funcional é um fator importante para os sistemas de saúde mental em qualquer país. Com um sistema operacional e implementado, é possível fornecer dados concretos aos formuladores de políticas e planejadores, e permitir que intervenções e avaliações direcionadas sejam desenvolvidas. Além disso, com o uso de um SISM, os profissionais de saúde mental são capazes de identificar padrões na apresentação de queixas e necessidades de medicação de sua população, o que, por sua vez, possibilita estruturar a prestação de serviços e a prática do cuidado

Um SISM é, portanto, uma ferramenta útil para o gerenciamento e monitoramento de pacientes. Ele permite que as informações sejam mantidas de maneira uniforme e centralizada, o que é valioso para os cuidados de longo prazo, particularmente quando os médicos ou as unidades de saúde mudam, uma vez que esses dados podem ser mais fáceis de acessar sempre que novos tratamentos forem necessários. O SISM também pode ser usado para gerar relatórios epidemiológicos, criando assim um melhor quadro da situação de saúde mental num hospital, distrito ou país (GLOVER, 2000).

Os SISM usados para medir e gerenciar o fornecimento de serviços de saúde mental, tornaram-se cada vez mais importantes para melhorar a eficácia dos cuidados de saúde mentais (JORDANS et al., 2016). OS SISM incluindo os de saúde mental, asseguram a coleta, o

processamento e a comunicação de dados e são projetados especificamente para auxiliar a política e o planejamento da gestão de saúde (WHO, 2004).

A Saúde Mental foi incluída como prioridade no Pacto pela Vida em 2007 quando foram elencados dois indicadores importantes, através da Portaria MS/GM nº 325, de 13 de março de 2006: a taxa de cobertura de CAPS/100.000 habitantes e a taxa de cobertura do Programa de Volta para Casa (MATTA; MORENO, 2014).

Contudo, além de possuir uma relevância técnica, o indicador, assim como a informação, pode ter seu uso político, estabelecendo sentidos aos sujeitos e aos cuidados em saúde. Muitas das vezes tomados como “verdades científicas” os indicadores e as informações podem ser instrumentos para o convencimento na construção de argumentos políticos, por vezes, atribuindo-lhes um grande poder de utilidade e resolutividade para diversas questões de saúde globalmente (MATTA; MORENO, 2014; MORENO, 2015).

Como exemplificação de plataforma de SISM, Miyoshi et al. (2018), modelaram e implementaram em ambiente web uma plataforma de integração (eHealth-Interop), seguindo uma arquitetura cliente-servidor, considerando que o atendimento regional é constituído por diferentes atores localizados em diferentes pontos de uma determinada região. No escopo da plataforma, a gestão da informação na atenção à saúde mental no âmbito nacional é apresentada por Miyoshi et al. (2018) através do desenvolvimento, implantação e gerenciamento de processos, e estruturação de repositório de dados através de técnicas computacionais de *big data*, com a plataforma de interoperabilidade desenvolvida para integração de cuidados mentais na região metropolitana da cidade de Ribeirão Preto/SP.

3 MINERAÇÃO DE PROCESSOS NA ATENÇÃO À SAÚDE

Neste capítulo são apresentadas as principais pesquisas que se relacionam com a proposta da tese.

Em 1995, Cook e Wolf publicaram na conferência internacional sobre tecnologia de banco de dados os primeiros algoritmos para determinar modelos de processos a partir de logs de eventos de software. O uso de MP para explorar modelos de processos de negócios, foi proposto por Agraval, Gunopulos e Leymann (1998). Segundo Neumuth et al. (2011), a comunidade de MP trabalhou ativamente nos últimos anos para formalizar a descoberta de modelos de processo com base em registros de eventos.

Conforme Huang, Lu e Duan (2012), a representação do conhecimento na área da saúde pode ser realizado através de um desenho dos processos clínicos, otimização de processos, decisão clínica, apoio e gestão de negócios, e assim por diante. A MP pode ser uma forma objetiva de analisar os caminhos clínicos, pois não é influenciada por percepções ou comportamentos normativos. Observa-se que os comportamentos médicos em jornadas de atendimento ao paciente podem ser registrados em registros de fluxo de trabalho clínico por meio de vários tipos de SIS.

Para Rebugue e Ferreira (2012), realizar mineração de processos em organizações de saúde é particularmente difícil devido à natureza altamente dinâmica, complexa e multidisciplinar em procedimentos da assistência médica. Mesmo assim, os autores concluem que a MP é uma abordagem promissora para obter um melhor entendimento sobre essas demandas, analisando dados de eventos registrados em SIS. Os sistemas estudados foram: Futura Reflect e ARIS PPM. Tratam-se de sistemas comerciais de *Business Process Management* (BPM) para automatizar a gestão de processos de negócio (modelagem, execução, controle e monitoração).

Van Genuchten et al. (2014), propõem o uso de MP para provar que a atualização de uma versão de software fornece benefícios quantificáveis para o usuário final. A MP aproveita o fato de que os logs de eventos capturam informações sobre processos. Esses eventos podem ser usados para tornar os processos visíveis e mostrar os benefícios de usar a próxima versão do produto de software. Segundo o autor, três grupos se beneficiam desse processo: usuários finais, fornecedores de software e pesquisadores. Alguns fornecedores de software, como a Mozilla, começaram a coletar métricas para entender melhor como seus produtos estão sendo

usados. Seu foco é tipicamente em aspectos como elementos de interface do usuário e uso de memória e não nos processos suportados pelos produtos e nos tempos de uso.

Conforme Garcia, Ramirez e Larrea (2015) o processo de mineração é responsável por gerar conhecimento e descobrir processos a partir de logs de eventos que são extraídos de sistemas de informação, para encontrar erros, inconsistências e vulnerabilidades. Para melhorar seu desempenho, as organizações buscam uma abordagem de gerenciamento de processos melhor, que, como primeiro passo, requer uma modelagem precisa destes. Para Garcia, Ramirez e Larrea (2015), Van Genuchten et al. (2014) e Zhou et al., (2014) a MP no setor da saúde é uma área ainda amplamente inexplorada.

A literatura especializada distingue dois tipos de dados de evento: dados históricos referem-se a logs de eventos completos do passado, enquanto dados atuais representam processos contínuos normalmente usados para executar suporte operacional. Conforme Bernard et al. (2016), um modelo de processo é usado principalmente para expressar a ordenação de atividades. Nesta mesma linha de investigação de Orellana, Estrada e Alfonso (2018) relatam em suas pesquisas como a mineração de processo foi aplicada em muitos domínios, incluindo manufatura, cadeia de suprimentos, governo, saúde e engenharia de software. Segundo os autores, particularmente na área de saúde, onde a maioria dos processos são complexos, variáveis, dinâmicos e multidisciplinares, a aplicação dessa técnica é cada vez mais desafiadora.

4 MATERIAL E MÉTODOS

Este capítulo apresenta cinco etapas desenvolvidas durante esta investigação para se alcançar o objetivo proposto tal como apresentadas na Figura 11. As mesmas retratam: obtenção dos *logs* dos eventos dos processos da plataforma SISAM-13. O período dos registros de episódios de internações dos processos presentes nos *logs*, compreende registros nos anos de 2013, 2014, 2015 e 2016. Após esta etapa, são descritos os métodos empregados para obtenção dos indicadores necessários para a construção da matriz de saúde a partir dos logs, utilizando o aplicativo *open source* DISCO. Posteriormente, é apresentado o método de estruturação dos dados (indicadores) para composição da matriz de saúde utilizando o modelo canônico openEHR. Em etapa seguinte, é descrita a utilização do modelo de geração de *big data* baseado nas evidências encontradas. Por fim, descreve-se a organização dos indicadores necessários para criação da linha do tempo do paciente utilizando os conceitos da matriz de saúde ao longo dos episódios de internação investigados.

Figura 11 – Etapas dos materiais e métodos utilizados

Fonte: O autor

A sessão a seguir descreve como os registros de episódios de internação (*logs*) foram obtidos a partir da sequência de eventos existentes no SISAM-13.

4.1 Descrição técnica (infraestrutura para obtenção de *logs*)

Um ambiente de repositório de registros obtidos por meio de um sistema de informação é representado por uma infraestrutura composta por computador (servidor) e SIS. Para obtenção dos *logs* são utilizadas linguagem de programação de *scripts* de banco de dados. Conforme

Miyoshi et al. (2018), a arquitetura do SISAM-13 segue um modelo com 3 camadas principais: Camada de Dados, Camada Semântica e Camada de Comunicação.

A camada de dados é onde os dados a serem integrados são representados e armazenados; Camada Semântica, na qual são descritos os esquemas mediadores, terminologias, ontologias, codificações locais e mapeamentos entre eles; finalmente, a Camada de Comunicação, onde os aspectos de segurança e comunicação com as fontes de dados são gerenciados (Figura 12).

Figura 12 – Arquitetura conceitual do SISAM-13.

Fonte: Adaptado de Miyoshi et al. (2018).

Para obtenção dos *logs* do SISAM-13 e posterior extração dos indicadores para a construção da matriz de saúde, foi elaborado um processo de preparação dos dados conhecido como ETL (*extraction-transformation-load*), ou seja, extração, transformação e carga contemplando por meio de um *script* utilizando a linguagem *Structured Query Language* (SQL). Trata-se de uma linguagem de pesquisa declarativa padrão para banco de dados relacional.

A linguagem SQL é o recurso mais conhecido por programadores de sistemas para a execução de comandos em bancos de dados (BD). É por meio dela que é possível criar tabelas, colunas, índices, atribuir permissões a usuários, bem como realizar consultas a dados. Enfim, é

utilizando a SQL que “conversamos” com o banco de dados. O *script* desenvolvido está disponível no Apêndice A. A programação para obtenção dos logs foi realizada levando-se em consideração os episódios de internação, identificando a passagem dos pacientes nos vários serviços das tabelas do BD disponibilizados pela plataforma SISAM-13 e exportados para um conjunto de logs.

O conjunto de comandos, *script*, apresentado no Apêndice B se executado em qualquer Sistema Gerenciador de Banco de Dados (SGBD) que tenha acesso a plataforma SISAM-13 gera *logs* compatíveis com os eventos ali registrados, referentes aos dados que a plataforma SISAM-13 armazenou originalmente conforme a arquitetura conceitual do SIASM. Ao inserir e executar comandos em linguagem SQL (utilizada pelo MySQL, que é o SGBD do SISAM-13) para serem executados no SGBD, estes são armazenados em um histórico configurado no próprio SGBD. Esse histórico é gravado em forma de arquivo texto com os comandos executados em sua sequência original e pode ser recuperado, gerando o conjunto de *logs*.

Os registros dos *logs* estão disponíveis em arquivos no formato de Valores Separados por Vírgula (VSV). Trata-se de um tipo de arquivo possível de ser interpretado por aplicativos que realizam MP. Em vez de armazenar dados em colunas, os arquivos CSV armazenam informações separadas por vírgulas. Quando o texto e os números são salvos em um arquivo CSV, é fácil utilizá-los em diferentes programas. O método ETL estudado e implementado para extração de indicadores a partir dos *logs* conforme Figura 13.

Figura 13 – ETL para SISAM-13.

Fonte: O autor.

A **extração** é o primeiro passo do conjunto de processos que consiste na leitura e no entendimento das fontes de dados para área de preparação de dados do SISAM-13. Uma vez

que os dados estão na área de preparação dos dados ocorrem as ordenações e processamentos sequencias preliminares a transformações. A **extração** implica nas correções de inconsistência de dados tais como limpeza de dados como a correção, p. ex.: de conflitos de nomes ou análise de padrões de formato, eliminação de redundância de dados.

O segundo passo, **carregamento**, é responsável por disponibilizar a extração dos dados referente aos episódios de internação dos pacientes para construção da matriz de saúde adaptada à DRS XIII através de um barramento *data mart* para ferramenta de acesso aos dados.

A última etapa, ocorrem o **acesso** do aplicativo DISCO; modelagem DUAL openEHR para os dados minerados, e conectividade do *software* EASY EHR. A sessão a seguir apresenta a utilização do aplicativo DISCO, incluindo sua linguagem de programação, infraestrutura de *hardware* e requisitos de *software* para o seu funcionamento.

4.2 Aplicativo DISCO

O aplicativo open source Disco, é uma ferramenta para MP, que pode oferecer informações valiosas sobre as funcionalidades que usuários acessam em um sistema de informação e como eles interagem com seus aplicativos. A utilização do software Disco pode auxiliar na criação de mapas de processo inteligentes diretamente de logs através de métricas de processo que são projetadas diretamente em um mapa de atividades (ou de eventos).

Para a utilização do aplicativo DISCO na pesquisa em questão, foi utilizado um computador do tipo Pentium (R) Dual-Core CPU T4500 2.30 GHz, 2GB (gigabytes) de memória RAM e sistema operacional Windows 7. No equipamento descrito foi instalada a Máquina Virtual Java (MVJ) versão 1.8.

4.3 Modelagem DUAL openEHR

Realizar a modelagem DUAL openEHR com o propósito final de estruturar a matriz de saúde proposta por Thornicroft e Tansella (1999) e adaptada por Vinci et al. (2016) exigiu, primeiro, que fosse especificado de modo abstrato do conhecimento a respeito do domínio (requisitos de *input*, *process* e *outputs*, contendo os atributos, tipos de dados, especificação de cada atributo), para que então, a codificação na linguagem de arquétipos, ADL fosse realizada. Essa etapa inicial foi realizada com a colaboração de dois profissionais, sendo um especialista

clínico (médico) e um informata biomédico. Após esta etapa, foi possível estruturar o conhecimento clínico em clusters. Um cluster é um conjunto de elementos relacionados em um mesmo arquétipo.

A escolha por elaborar arquétipos do tipo cluster decorre do fato que a matriz de saúde mental apresenta em sua composição dados de *input*, *output* e *process*, para que fosse possível realizar uma investigação em qualquer uma das suas camadas de modo independente, assim como de modo único em conformidade com a matriz. Para tanto foram criados três arquétipos novos específicos:

- *openEHR-EHR-CLUSTER.mental_health_matrix_inputs.v1.adl*: esse arquétipo contém dados referentes ao código do paciente, profissional que realizou a solicitação da internação, data da solicitação, sexo, idade, estado civil, serviço que solicitou a internação, município de onde veio o paciente, município onde mora o paciente, diagnóstico principal, e as comorbidades: problemas respiratórios, acidente cardiovascular, convulsão, HAS, gravidez, diabetes e doença infectocontagiosa. A Figura 14 ilustra o arquétipo para representar a entrada de dados, “INPUTS” que é do tipo cluster.

Figura 14 – Mapa mental do arquétipo de entrada de dados: *openEHR-EHR-CLUSTER.mental_health_matrix_inputs.v1.adl*.

Fonte: o autor.

O Apêndice C apresenta a codificação referente ao arquétipo de entrada de dados: *openEHR-EHR-CLUSTER.mental_health_matrix_inputs.v1.adl*.

- *openEHR-EHR-CLUSTER.mental_health_matrix_process.v1.adl*: este arquétipo contém dados referentes ao tempo de liberação da vaga, unidade de

internação, transferência intra hospitalar e duração da internação. A Figura 15 ilustra o arquétipo “PROCESS” que é do tipo *cluster*.

Figura 15– Mapa mental do arquétipo de processamento de dados: openEHR-EHR-CLUSTER.mental_health_matrix_process.v1.adl.

Fonte: o autor.

A codificação do arquétipo “PROCESS” está descrita no Apêndice D. O arquétipo para composição de dados de processamento da matriz de saúde mental após ser elaborado, fica armazenado fisicamente em um repositório de arquétipos. Para facilitar a localização e identificação dos arquétipos, ao realizar a criação de um novo a nomenclatura “openEHR-EHR” é indicada pra identificação dos arquivos em seu formato físico:

- *openEHR-EHR-CLUSTER.mental_health_matrix_outputs.v1.adl*: este arquétipo contém dados referentes ao código do profissional que fez a solicitação, destino do paciente, diagnóstico primário, diagnóstico secundário e profissional que fez a alta. A Figura 16 ilustra o arquétipo “OUTPUTS” que é do tipo *cluster*;

Figura 16 – Mapa mental arquétipo de saída de dados: openEHR-EHR-CLUSTER.mental_health_matrix_process.v1.adl.

Fonte: o autor.

A codificação do arquétipo final, *outputs*, está descrita no Apêndice E. A integração dos arquétipos de entrada, processamento e saída da matriz foi feita através da criação de um

template. O *template* é um formulário especializado no modelo canônico openEHR que possibilita ao profissional clínico após a elaboração do(s) arquétipo(s) realizar a gravação, alteração e consulta (também chamada de persistência) de dados mediante a sua necessidade de investigação. A Figura 17 ilustra o *template* criado, que foi batizado de *mental_health_matrix.oet*.

Figura 17 – Formulário clínico especializado para saúde mental.

Fonte: o autor.

4.4 API GDC-Arquetype e EDC-Templates para gerenciamento e execução de definições clínicas

A etapa posterior à definição e codificação dos arquétipos e *templates* em conformidade com a modelagem DUAL openEHR foi gerar a estrutura da matriz de saúde. Para esta etapa, foram utilizadas as APIs GDC- Arquetype, Apêndice G, e EDC-Templates, Apêndice H. Codificadas utilizando as linguagens de programação Java e Grails, as APIs foram implementadas no ambiente de desenvolvimento Eclipse. As APIs são os programas computacionais responsáveis por fazer com que os arquétipos e *template* anteriormente modelados funcionem na prática, executando seus códigos. Ambas as APIs utilizam os seguintes plugins opensource:

- Tomcat:7.0.47: é o servidor web Java utilizado para o funcionamento das APIs. Após inicializar este *plugin*, os *softwares* entram em funcionamento;
- Scaffolding:2.0.1: *Plugin* que permite gerar automaticamente interfaces Web para uma determinada classe de domínio, incluindo com operações de criação/leitura/atualização/exclusão de páginas HTML;

- JQuery:1.10.2.2: é uma biblioteca de funções JavaScript que interage com o HTML, desenvolvida para simplificar os *scripts* interpretados no navegador de internet;

Embora os *plugins* mencionados anteriormente sejam utilizados por ambas as APIs GDC- Arquetype e EDC-Templates, existem algumas particularidades que são específicas de cada API, a saber:

- GDC- Arquetype: é a API que foi desenvolvida com o propósito de ser um gerenciador de definições clínicas. Apresenta interface visual com usuário através de uma única execução do sistema. A partir daí, com o uso da internet a API estabelece a forma de interação da entidade com o mundo exterior através de um *web service*, que é uma solução utilizada na integração de sistemas e na comunicação entre aplicações diferentes. Com esta tecnologia é possível que novas aplicações possam interagir com aquelas que já existem e que sistemas desenvolvidos em plataformas diferentes sejam compatíveis. A API contém os seguintes módulos implementados:
 - br.com: módulo responsável por armazenar classes que estabelecem comunicação com repositório openEHR através de *web service* de wsdl;
 - archetypeservlet: módulo que gerencia as classes provenientes de *web service* de wsdl. Um modelo de análise, projeto e programação de sistemas de software baseado na composição e interação entre diversas unidades de *software* chamadas de objetos;
 - archetypeservlet.xsd: módulo responsável por realizar a criação de instâncias de objetos com o uso de padrão de projetos de *software*;
- EDC-Template: API que faz uso do gerenciador de definições clínicas openEHR. O seu funcionamento é suportado pelos seguintes componentes de *software*:
 - adl-parser-1.0.5.jar: componente de *software* utilizado para interpretar os dados lidos da MP utilizando-se os arquétipos openEHR-EHR-CLUSTER.mental_health_matrix_inputs.v1.adl, openEHR-EHR-CLUSTER.mental_health_matrix_outputs.v1.adl e openEHR-EHR-CLUSTER.mental_health_matrix_process.v1.adl. Após realizar a leitura dos arquétipos é gerado um objeto computacional na forma de modelo de objeto de arquétipo para composição da estrutura da matriz de saúde. Durante o uso do componente, todos os elementos presentes

no *<input>*, *<process>* e *<outputs>* são validados de acordo com o seu tipo;

- oet-parser-1.0.5.jar: é o componente utilizado para realizar a conversão do arquétipo escrito em linguagem ADL, para a linguagem de programação orientada a objetos Java. O componente, apresenta dependência de outros componentes para o seu funcionamento, sendo estes: openehr-rm-core.jar, openehr-rm-domain.jar, openehr-aom.jar, adl-parser.jar, xmlbeans.jar, jsr173.jar, measure-serv.jar, mini-termserv.jar, adl-serializer.jar, log4j.jar, commons-io.jar,
- openehr-aom-1.0.5.jar: contém a instrução definitiva da semântica do arquétipo, na forma de um modelo de objeto para arquétipos. O modelo pode ser usado como base para a construção de *softwares* que processam arquétipos, independentemente de sua representação persistente;
- openehr-rm-core-1.0.5.jar: biblioteca utilizada para geração de leitura de *templates* openEHR de extensão oet. A estrutura da matriz de saúde, batizada de *mental_health_matrix.oet* é gerada com o uso deste componente;

A codificação utilizada na API GDC-Arquetype encontra-se no Apêndice F e para a API EDC-Template no Apêndice G. O Apêndice H apresenta um exemplo de sequência de telas com as etapas do funcionamento das APIs e a exibição do modelo do formulário clínico para extração dos indicadores e estruturação da matriz de saúde mental.

4.5 EASY- EHR

Trata-se de um sistema desenvolvido em conformidade com a modelagem openEHR que realiza integração com banco de dados noSQL, implementado utilizando a linguagem Java e Grails. O projeto arquitetural do EASY-EHR leva em consideração o modelo de *software Model View Controller* (MVC). Diferentemente das APIs apresentadas na sessão anterior o EASY-EHR foi desenvolvido para manipular formulários clínicos baseados em arquétipos. O sistema foi implementado utilizando o ambiente de desenvolvimento *eclipse* e utiliza os seguintes *plugins open source*:

- Tomcat:7.0.47: é o servidor web Java utilizado para o funcionamento das APIs. Após inicializar este plugin, os softwares entram em funcionamento;

- Scaffolding:2.0.1: Plugin que permite gerar automaticamente interfaces Web para uma determinada classe de domínio, incluindo com operações de criação/leitura/atualização/exclusão de páginas HTML;
- Hibernate: 4.1.11.6: é um *framework* utilizado para o mapeamento objeto-relacional escrito na linguagem Java;
- Database-migration:1.3.8: *plugin* utilizado para realizar a migração de esquema bancos de dados noSQL;
- JQuery:1.10.2.2: é uma biblioteca de funções JavaScript que interage com o HTML, desenvolvida para simplificar os scripts interpretados no navegador internet;
- MongoDB:2.0.0: SGBD utilizada para manipulação *big data* NoSQL;

O sistema funciona tanto em ambiente Windows quanto em distribuição Linux. Para realizar a sua instalação basta copiar o projeto para a pasta “*bin*” do *plugin* Toncat e executar o ícone start-toncat. Contudo é importante que as seguintes instruções sejam seguidas para o seu uso:

- Realizar a instalação da MVJ no computador;
- Instalar o banco de dados MongoDB;
- Ao realizar o primeiro acesso com o EASY-EHR o sistema verifica automaticamente a existência do repositório de dados para o seu uso. Caso não exista, o sistema cria uma comunicação com o MongoDB e cria o repositório MENTAL_HEALTH_OPENEHR. As configurações para funcionamento do *big data* são: host = "localhost", port = 27017, username = "MENTAL_HEALTH_OPENEHR" e password = "SISAM-13";

A Figura 18 ilustra a *integração das ferramentas* por meio do projeto arquitetural do EASY-EHR com a edição do *template* de matriz de saúde mental.

Figura 18 – *Template* para matriz de saúde mental.

Hilton Vicente César ▾

Meu Template

EHR Archetypes

- Cluster
- Composition
- Element
- Entry
 - Action
 - Evaluation
 - Observation
 - Instruction
 - Admin
- Section
- Structure

Demographic Model

PAINEL > MEU TEMPLATE

Matriz de Saúde: Template openEHR

ENCOUNTER

T 433 ▾

INPUTS

ID: ID

profissional_solicitacao: T

data_solicitacao: UI

cod_paciente: Q

idade: T

estado_civil: T

sexo: T

servico_solicitacao_internacao: T

municipio_origem_paciente: T

municipio_paciente: T

diagnostico_principal: T

traumatismo: T

prob_respiratorio: T

avc: T

convulsao: T

has: T

gravidez: T

diabetes: T

doenca_infecto: T

PROCESS

ID: ID

tempo_liberacao_vaga: UI

unidade_internacao: T

transferencia_intrahospitalar: T

duracao_internacao: UI

OUTPUTS

ID: ID

destino_paciente: T

diagnostico_primario: T

diagnostico_secundario: T

profissional_alta: T

Fonte: O autor.

5 RESULTADOS

Neste capítulo são detalhados os resultados obtidos. O resultado principal da tese é a modelagem do processo para a extração de indicadores para a matriz de saúde a partir dos *logs* do sistema de informação, através do uso de mineração de processo. Como resultados secundários pode-se citar estruturação dos dados segundo o modelo de arquétipos openEHR e a criação de um grande banco de dados (*big data*) que abrange informações demográficas, de episódios de internação/reinternação e comorbidades da população, cujos procedimentos de atenção à saúde foram realizados e registrados no SISAM-13 no período no qual os *logs* foram estudados.

5.1 Modelagem do SISAM-13 para geração dos *logs* (reengenharia de *software*)

A investigação do modelo conceitual SISAM-13 na camada de DATALAYER possibilita um entendimento de como os dados são armazenados e como a geração de logs pode ser realizada. Para o entendimento desta camada, a modelagem de dados *Unified Modeling Language* (UML) se apresenta como a linguagem-padrão para este propósito. A UML é empregada para a visualização, a especificação, a construção e a documentação de artefatos que façam uso de sistemas complexos de *software*.

A Figura 19 apresenta um recorte da camada DATALAYER responsável pela aquisição de *logs* do SISAM-13. Esse modelo retrata as relações das tabelas que estão presentes no repositório de dados relacionadas aos episódios de internação do paciente. As tabelas investigadas neste modelo são: internações, regulações, pacientes, usuários, unidades, profissionais e pessoas. Cada tabela representada através de um diagrama UML contempla um número específico de atributos para manipulação dos dados. As setas que fazem ligações entre as tabelas indicam suas relações. A organização das tabelas está distribuída da seguinte forma:

- Internações apresenta um valor único, chave primária, representado através do atributo que armazena o código de internação. Além disso a tabela de internações armazena registros importantes como o código da regulação, código do profissional, data da internação, código de transferência para outro hospital, status da internação, tipo da internação, tipo da alta, código de alta e profissional que realizou a alta;

- A tabela de regulações contempla através do código de regulação a identificação de seus registros. É possível identificar nesta tabela registros de códigos que retratam a relação entre a tabela de regulação com outras tabelas como p. ex.: código da pessoa, código do paciente, código do solicitador, código da unidade de origem, código da unidade de destino, tipo de regulação realizada e comorbidades;
- A tabela de usuários apresenta os códigos de solicitador, código do profissional e código de usuário. Dos três atributos apenas o último é único nesta tabela;
- A tabela de unidades disponibiliza um código unitário, além do nome das unidades e o município;
- Na tabela de profissionais apenas o código do profissional está disponibilizado;
- A tabela de pacientes contempla o código da pessoa, o nome, unidade de realização de cadastro, o estado civil, etnia e arranjo domiciliar;
- Na tabela de pessoas, a identificação única é através do código da pessoa. Além do código único, a tabela disponibiliza informações demográficas como o sexo e a data de nascimento;

Figura 19– Modelagem DATALAYER SISAM-13.

Fonte: O autor.

A Tabela 2 apresenta os atributos por registro relacionados a dados demográficos, presentes nos *logs* do sistema.

Tabela 2 – Atributos demográficos do paciente.

Atributo	Tipo	Descrição
<i>Cod_pessoa</i>	Numérico	Código da paciente
<i>Sexo</i>	Literal	Sexo
<i>Data_nascimento</i>	Data	Data de nascimento
<i>Município</i>	Literal	Município de origem
<i>Cidade_residencia</i>	Literal	Cidade de residência
<i>Estado_residencia</i>	Literal	Estado de residência
<i>Logradouro</i>	Literal	Endereço completo
<i>Profissão</i>	Literal	Profissão
<i>Idade</i>	Numérico	Idade
<i>Estad_civil</i>	Literal	Estado civil
<i>Etnia</i>	Literal	Etnia
<i>Arranjo_domiciliar</i>	Literal	Arranjo domiciliar
<i>Estad_civil</i>	Literal	Estado civil
<i>Estad_civil</i>	Literal	Estado civil

Fonte: O autor.

A Tabela 3 apresenta os atributos complementares ao registro do Log relacionado a episódios de internação do paciente.

Tabela 3 – Atributos de episódios de internação.

Atributo	Tipo	Descrição
<i>Unidade_cadastro</i>	Literal	Unidade onde o paciente foi cadastrado
<i>Cod_regulacao</i>	Numérico	Código da regulação
<i>Cod_internacao</i>	Numérico	Código de internação
<i>Cod_transferencia</i>	Numérico	Código de transferência
<i>Cod_alta</i>	Numérico	Código de alta
<i>Cod_unidade_origem</i>	Numérico	Código da unidade de origem
<i>Sts_internacao</i>	Literal	Status da internação
<i>Sts_regulacao</i>	Literal	Status da regulação
<i>Via_direta</i>	Literal	Via direta
<i>Tipo_transferencia</i>	Literal	Tipo de transferência
<i>CID_diagnostico_principal</i>	Literal	CID diagnóstico principal
<i>CID_diagnostico_secundario</i>	Literal	CID diagnóstico secundário
<i>CID_diagnostico_principal_alta</i>	Literal	CID diagnóstico principal da alta
<i>CID_diagnostico_secundario_alta</i>	Literal	CID diagnostico secundário da alta
<i>Data_hora_vaga_internacao</i>	Data	Data e hora em que foi disponibilizado a vaga de internação
<i>Data_hora_admissao</i>	Data	Data e hora em que foi realizado a admissão do paciente
<i>Data_hora_internacao</i>	Data	Data e hora em que foi realizado a internação

Fonte: O autor.

Atributo	Tipo	Descrição
<i>Data_hora_cancelamento</i>	Data	Data e hora em que foi realizado o cancelamento da solicitação de internação
<i>Data_hora_transferencia</i>	Data	Data e hora em que foi realizado a transferência do paciente para outra unidade
<i>Tipo_alta</i>	Literal	Tipo da alta
<i>Tipo_internacao</i>	Literal	Tipo de internação
<i>Nome_unidade_solicitacao</i>	Literal	Nome da unidade de solicitação
<i>Nome_unidade_destino</i>	Literal	Nome da unidade de destino
<i>Municipio_solicitacao</i>	Literal	Município da unidade de solicitação
<i>Municipio_destino</i>	Literal	Município da unidade de destino
<i>Cod_profissional_solicitacao</i>	Numérico	Código do profissional que realizou a solicitação
<i>Cod_profissional_internacao</i>	Numérico	Código do profissional que realizou a internação
<i>Cod_profissional_alta</i>	Númerico	Código do profissional que realizou a alta
<i>Via_internacao</i>	Literal	Via de internação
<i>Tempo_espera_internacao</i>	Numérico	Tempo de espera de internação
<i>Tempo_internacao</i>	Numérico	Tempo de internação
<i>Grupo_diagnostico_alta</i>	Literal	Grupo diagnostico da alta

Fonte: O autor.

Os atributos relacionados as comorbidades do paciente estão presentes na lista da Tabela 4.

Tabela 4 – Atributos de comorbidades do paciente

Atributo	Tipo	Descrição
<i>Traumatismo</i>	Literal	Traumatismo
<i>Prob_respiratorio</i>	Literal	Problemas respiratórios
<i>AVC</i>	Literal	Acidente vascular cerebral
<i>Convulsão</i>	Literal	Convulsão
<i>HAS</i>	Literal	Hipertensão arterial sistêmica
<i>Gravidez</i>	Literal	Gravidez
<i>Diabetes</i>	Literal	Diabetes
<i>Doença infeto contagiosa</i>	Literal	Doença infecto contagiosa

Fonte: O autor.

5.2 Mineração de processos aplicada aos *logs* gerados pelo modelo

Conforme apresentado na parte introdutória da tese, a matriz de saúde em seu formato genérico estabelece um modelo que pode ser usado para aumentar a eficácia clínica através da prática baseada em evidências, sendo estruturada em dois eixos: temporal e espacial (geográfico). O modelo de matriz de saúde mental, adaptado à DRS XIII por Vinci et al. (2016) (Figura 1) é composto no eixo temporal por indicadores do processo de atenção à saúde referentes aos dados de entrada (*inputs*), como por exemplo referentes ao profissional que

realizou a solicitação da internação, diagnóstico primário e diagnóstico secundário, entre outros; processamento (*process*), como por exemplo, o tempo entre a liberação da vaga e a internação do paciente no hospital, o tempo de duração de internação, entre outros; e saída (*outputs*), como por exemplo, o destino final do paciente, a data e a hora da alta, o profissional que realizou a alta, entre outros. No âmbito da dimensão espacial (geográfico) a matriz de saúde adaptada utiliza indicadores relacionados com os níveis de usuário, profissionais de saúde, serviços de saúde, cidades, regiões de saúde e departamento regional de saúde. Ainda segundo o esquema da Figura 1, todos os indicadores necessários para a estruturação da matriz nos diferentes níveis temporais e espaciais estão presentes no SISAM-13. A seguir são apresentados exemplos variados de como esses indicadores podem ser gerados a partir dos *logs* de eventos existentes no SISAM-13 a partir do uso de mineração de processos. Para fins da pesquisa os dados referentes aos pacientes e aos profissionais foram anonimizados.

A Figura 20 mostra um exemplo da geração de indicadores para a matriz de saúde considerando como nível geográfico o paciente, tendo como *input* a unidade de saúde HC-EPIB e um único episódio de solicitação de internação, como *process* no início da internação e como *output* o tempo de duração da internação.

Figura 20 – Exemplo de geração de indicadores no nível do paciente.

Fonte: O autor

Conforme apresentado na Figura 17, foi possível identificar trezentos e noventa e cinco pacientes internados uma vez nessa unidade, durante o período analisado. O aplicativo DISCO, possibilita a navegação através de um menu interativo, conforme representado, com a listagem

de todos os pacientes, sendo possível ter acesso às informações do paciente por meio do seu código de identificação.

A opção do menu, “*Complete log*”, possibilita ao gestor clínico uma visão completa das internações e a opção variant 1 o acesso aos dados referentes aos pacientes que foram internados uma única vez.

O item do menu, representado por *cases*, apresenta paciente por paciente em uma linha do tempo. Os números 3400, 3365, 3409, 3305 e 2802 são códigos para identificação dos pacientes. Por fim é apresentada, como exemplo, a linha do tempo do paciente representado pelo código 3400, com informações relacionadas a data e hora em que ocorreu o episódio de internação desse paciente.

O paciente 3400 foi internado no dia dois de janeiro de 2014, as 17h14min. Além disso, é possível verificar o tempo que o paciente ficou internado na unidade HC EPIB, nesse caso por 23 dias e 1 hora.

A Figura 21 apresenta um exemplo da geração de indicadores para a matriz de saúde considerando como nível geográfico o paciente, tendo como *input* a unidade de saúde HC-EPIB, porém com até dois episódios de solicitação de internação, como *process* o início da internação e como *output* o número de internações, a duração de cada internação e o intervalo entre elas.

Figura 21 – Exemplo de geração de indicadores no nível do paciente.

Fonte: o autor

Conforme apresentado na Figura 19, foi possível identificar 447 pacientes internados nesta unidade, sendo 395 com uma única internação e 52 com duas internações cada um, totalizando 499 internações no período estudado.

A opção do menu Complete log possibilita ao gestor clínico a visão completa dos pacientes e a opção variant 1 e variant 2, os indicadores do fluxo de pacientes. Nesta visão de resultados é possível observar dois fluxos em que ocorreram internações. No fluxo variant 1, são apresentados 395 casos de pacientes internados uma única vez. Na variant 2, são apresentados os 52 casos de pacientes internados duas vezes cada um. O item do menu representado por *case* apresenta paciente por paciente em uma linha do tempo.

Os números 3393, 3410, 3443, 3122, 3522 e 3678 são códigos para identificação dos pacientes. A linha do tempo contemplando o código 3678, apresenta informações relacionadas a data e hora em que ocorreu o episódio de internação deste paciente. No dia 26 de janeiro de 2014 às 09h53min o paciente 3678 foi internado, permanecendo internado por 20 dias e 23 horas. O paciente realizou uma segunda internação 13 dias e 17 min depois da primeira internação na unidade HC EPIB, no dia 4 de abril de 2014, às 9h20min19s permanecendo por 13 dias e 7 horas internado.

A Figura 22 apresenta um exemplo da geração de indicadores para a matriz de saúde considerando como nível geográfico o paciente, tendo como *input* a unidade de saúde HC-EPIB, porém com até três episódios de solicitação de internação, como *process* o início da internação e como *output* o número de internações, a duração de cada internação e o intervalo entre elas.

Figura 22 – Exemplo de geração de indicadores no nível do paciente.

Fonte: o autor.

Conforme apresentado na Figura 20, foi possível identificar 458 pacientes internados nesta unidade, sendo 395 com uma internação, 52 com duas internações cada um e 11 com três internações cada um, totalizando 532 internações no período estudado.

A opção do menu “*Complete log*” possibilita ao gestor clínico a visão completa dos pacientes e a opção variant 1, variant 2 e variant 3, os indicadores dos fluxos de pacientes. Nesta visão de resultados é possível observar três fluxos em que ocorreram internações. No fluxo variant 1, são apresentados 395 casos de pacientes que foram internados apenas uma vez. Na variant 2, são apresentados 52 casos de pacientes com duas internações cada um. Na variant 3, são apresentados 11 casos de pacientes com três internações cada um.

O item do menu representado por case apresenta paciente por paciente em uma linha do tempo. Os números 3588, 3619, 2817, 1672, 5240, 5331 e 6414 são códigos para identificação dos pacientes. Por fim a linha do tempo contemplando o código 3678, apresenta informações relacionadas a data e hora em que ocorreu o episódio de internação. No dia 16 de fevereiro de 2014 às 12h42min7s o paciente 3619 foi internado, permanecendo internado por 18 dias e 21 horas. O paciente realizou uma segunda internação 8 dias e 2h depois na unidade HC EPIB, no dia 15 de março de 2014 às 12h35min46s permanecendo por 3 dias e 20 horas internado. Após 35 dias e 54 minutos da sua segunda internação, o paciente 3619 foi submetido a uma terceira internação no HC EPIB, no dia 23 de abril de 2014 às 09h57min, permanecendo internado por 20 dias e 6h na unidade

A Figura 23 apresenta exemplo da geração de indicadores para a matriz de saúde considerando como nível geográfico o profissional, tendo como *input* a unidade de saúde HC-EPIB e até três solicitações de internação por paciente, como *process* o número de solicitações de internações por paciente e como *output* o número de pacientes.

Figura 23 – Exemplo de geração de indicadores no nível do profissional.

Fonte: o autor.

A Figura 23 apresenta um profissional clínico (4375) que fez duas solicitações de internação para seus pacientes no HC EPIB, sendo um dos pacientes o de código 374. Somando-se o total de “cases” apresentados em cada “variant” totaliza-se 40 solicitações de internações, com pacientes internados uma, duas ou três vezes no HC EPIB. Os números 4201, 4199, 4184, 4401 e 4375 são códigos para identificação dos profissionais de saúde.

Note-se que a granularidade mais fina da informação representa a linha de tempo de eventos do paciente (timeline), o que na matriz de saúde se refere ao nível geográfico do paciente, tendo como *input* a solicitação de internação e o profissional solicitante, como *process* a internação e seu período de duração e como *output* o momento da alta. No caso de pacientes com duas ou mais internações, o *output* traz também o intervalo de tempo entre as internações.

A Figura 24 apresenta exemplo da geração de indicadores para a matriz de saúde considerando como nível geográfico o departamento regional de saúde, tendo como *input* profissionais que realizaram até nove internações de um mesmo paciente, como *process* os profissionais que realizaram internações e número de internações e como *output* os pacientes internados por cada profissional.

Figura 24 – Exemplo de geração de indicadores no nível do departamento regional de saúde.

Fonte: o autor.

A Figura 24 apresenta a lista de profissionais clínicos do departamento regional de saúde que realizaram até nove internações de um mesmo paciente. Foi possível identificar 1480 pacientes internados. Considerando-se as ocorrências de reinternação (593 no total) para um conjunto de 9 variant, verifica-se um total de 2.073 internações realizadas nos serviços do

departamento regional de saúde. A opção do menu ‘*Complete log*’ possibilita ao gestor clínico a visão completa do número de pacientes internados e as opções variant 1, variant 2, variant 3, variant 4, variant 5, variant 6, variant 7, variant 8 e variant 9, detalham as reinternações.

Os números 4401, 4201, 4199, 4032, 4184, 4199, 4032, 4379, 4787 e 4395 são códigos para identificação dos profissionais de saúde. Note-se que a granularidade mais fina aqui também representa a linha de tempo de eventos do paciente (*timeline*), o que na matriz de saúde se refere ao nível geográfico do paciente, tendo como *input* a internação e o profissional que fez a internação, como *process* o número de internações, momento da internação e seu período de duração e como *output* o momento da alta. No caso de pacientes com duas ou mais internações, o *output* traz também o intervalo de tempo entre as internações.

5.3 Big data para saúde mental

A criação do *big data* para saúde mental foi realizado levando-se em consideração os indicadores extraídos com o uso da mineração de processos aplicada aos *logs* do SISAM-13. Tais indicadores abrangem informações demográficas, de episódios de internação/reinternação e comorbidades da população e se apresentaram aderentes a modelagem openEHR realizada com a adoção de arquétipos clínicos de *input*, *process* e *outputs* da matriz de saúde que foram extraídos das APIs GDC-Arquetypes e EDC-Templates para essa finalidade.

A solução para possibilitar o acesso/manipulação aos dados foi elaborada sobre o sistema de banco de dados MongoDB. A infraestrutura utilizada para realização dos experimentos foram um computador com processador Intel Core i6- 3517U CPU de 2.5 GHz, 4 GB de RAM, com o sistema operacional Windows 10 Pro de 64 bits. A versão do MongoDB foi a 3.2.3.

Para implementar os testes, a linguagem Java, versão 1.8, foi utilizada juntamente com os respectivos *drivers* necessários para a conexão com o banco de dados. Para o MongoDB foi utilizado o *driver* Java Driver MongoDB. Assim o primeiro passo foi criar a estrutura para armazenamento dos dados. O comando utilizado com essa finalidade foi:

- `use MENTAL_HEALTH_OPENEHR;`

Após a criação do banco de dados, foram criadas as coleções: Document e SISAM_INTERNACOES. A coleção system.indices é gerada automaticamente pelo MongoDB para realizar o gerenciamento dos índices para formulários armazenados. Os comandos utilizados foram:

- `db.createCollection("Document");`

- `db.createCollection("SISAM_INTERNACOES");`

A Figura 25 mostra o esquema de dados orientado a documentos, utilizando o banco de dados MongoDB, criado a partir dos quatro arquétipos usados nessa pesquisa.

Figura 25 – Esquema de dados orientado a documentos utilizando MongoDB.

Name	Count	Size	Storage	Index	AvgObj
Document	48	6 KB	8 KB	7 KB	147,33
SISAM_INTERNACOES	4,814	16,2 MB	32,2 MB	0,2 MB	3336,71
system.indexes	3	0,3 KB	8 KB	0 bytes	90,67

Fonte: o autor.

Um exemplo de comando para realizar pesquisas aos episódios de internação, no repositório *big data* para saúde mental:

- `db.SISAM_INTERNACOES.find()`. Ao executar esse comando, são disponibilizados o acesso de todos os formulários armazenados. Foram armazenados no mongoDB um total de 4.814 registros na base SISAM_INTERNACOES.

A base contempla os indicadores dos episódios de internações extraídos da mineração de processos para composição da matriz de saúde, conforme proposto nesta investigação. Porém, se a consulta aos formulários receber como parâmetro um valor solicitado, será disponibilizado um documento de retorno.

A Figura 26 apresenta o resultado obtido ao executar o comando de exemplo:

- `db.SISAM_INTERNACOES.find(1);`

Figura 26: Documento openEHR, exemplo de episódio de internação, no repositório *big data* para saúde mental.

The screenshot shows a web application interface for a database explorer. At the top, there are two tabs: 'MENTAL_HEALTH_OPENEHR' (active) and 'SISAM_INTERNAÇÕES'. Below the tabs are three view options: 'Tree View' (selected), 'Table View', and 'Text View'. A status bar indicates '100 Documents (0 to 99)'. The main area displays a hierarchical tree structure of a document. The root is '{ } {..}'. It has several children, including '_id' (5ce982a57faa9827db742fa8), 'limit' (1), 'category_defining_code_0' (433), and a series of 'ehr/items[at...]/value[...]' entries. The last entry is 'ehr/items[at0044]/value[21]' with the value '02/01/2014 12:38:37'.

Key	Value
{ } {..}	
... _id	5ce982a57faa9827db742fa8
... limit	1
... category_defining_code_0	433
... ehr/items[at0032]/value[1]	02/01/2014 17:48:00
... ehr/items[at0033]/value[2]	3400
... ehr/items[at0034]/value[3]	PA
... ehr/items[at0035]/value[4]	30/12/2013 00:00:00
... ehr/items[at0036]/value[5]	M
... ehr/items[at0037]/value[6]	12/01/1968 00:00:00
... ehr/items[at0038]/value[7]	SP
... ehr/items[at0039]/value[8]	RIBEIRÃO PRETO
... ehr/items[at0040]/value[9]	RUA MANTIQUEIRA 934
... ehr/items[at0041]/value[10]	INATIVO
... ehr/items[at0043]/value[11]	3400
... ehr/items[at0044]/value[12]	38
... ehr/items[at0044]/value[13]	0
... ehr/items[at0044]/value[14]	SOLTEIRO
... ehr/items[at0044]/value[15]	BRANCO
... ehr/items[at0044]/value[16]	
... ehr/items[at0044]/value[17]	4305
... ehr/items[at0044]/value[18]	38
... ehr/items[at0044]/value[19]	AGITAÇÃO PSICOMOTORA
... ehr/items[at0044]/value[20]	
... ehr/items[at0044]/value[21]	02/01/2014 12:38:37

Fonte: o autor.

O esquema de dados apresenta um *database explorer* com a possibilidade de estabelecer conexão com o *big data*, batizado MENTAL_HEALTH_OPENEHR. É possível navegar no ambiente *big data* através do menu de opções *collections*, *Stored JavaScript*, *GridFS* e *User*. A visão TREE_VIEW representada pelo conceito da modelagem openEHR para armazenamento de formulários “chave” *key*, “valor” *value* é um exemplo do primeiro episódio registrado na base.

Cada registro apresenta uma identificação única *primary key* e o atributo *_id* é responsável por armazenar tais registros. A identificação única em *big data* é representada por um tipo de dado alfanumérico e diferentemente das bases de dados relacionais são gerados dinamicamente, minimizando possíveis erros quando comparados com modelos legados de repositório de dados.

Além da *primary key* o formulário apresenta como descritivos dos atributos que armazenam os dados as iniciais *ehr_itens_[at_x]_value[x]*. Onde o valor de *x* refere-se a um dado valor atribuído durante a elaboração do arquétipo. Por exemplo, o atributo do formulário *ehr_itens_[at_0036]_value[5]* refere-se ao sexo do paciente e o *ehr_itens_[at_0044]_value[14]* ao estado civil do paciente.

O código fonte parcial de classe que realiza comunicação com *big data* com o *mongoDB* está disponível no Apêndice I. A implementação openEHR para integração com *big data* está disponível no Apêndice J.

5.4 API para gestão da informação e criação da *timeline*

A API desenvolvida, batizada de Easy-EHR, disponibiliza um conjunto de interfaces para auxiliar a atividade do gestor em saúde. A Figura 25 apresenta a *timeline* do paciente em conformidade com a matriz de saúde proposta por Thornicroft e Tansella (1999) e adaptada por Vinci et al. (2016) para a DRS XIII, com base nos *logs* registrados no SISAM-13. O conjunto de informações retrata os dados demográficos, entradas <inputs>, processamento <process> e destino <outputs> dos episódios de internação que foram identificados nos *logs* dos registros na ordem cronológica. Dentre as funcionalidades oferecidas pela API existe um menu autoexplicativo que possibilita:

- Visualizar os arquétipos elaborados pelo profissional clínico;
- Realizar pesquisa de informações provenientes da MP;
- Acessar o *big data* dos dados utilizados para composição dos indicadores de saúde em conformidade com o modelo canônico openEHR;

Com a informatização da matriz de saúde, o profissional da saúde consegue realizar uma pesquisa de rastreabilidade do paciente durante a realização de cada episódio de internação. Para cada evento registrado no *big data* utilizando o banco de dados NoSQL MongoDB é possível caracterizar os indicadores:

- *Input*: profissional que fez a solicitação, unidade de solicitação, município de solicitação, diagnóstico principal, diagnóstico secundário e comorbidades;
- *Process*: tempo entre a liberação da vaga e a internação do paciente no hospital, tempo de internação e transferência intra-hospitalar;
- *Output*: para onde o paciente foi após a alta, diagnóstico principal da alta, diagnóstico secundário da alta, tipo da alta e profissional que fez a alta;

Para exemplificar a *timeline* com o uso do Easy-EHR utilizamos o paciente código 5240 (Figura 27), sexo feminino, moradora do município de Serra Azul/SP. O paciente apresenta 6 episódios de internações. No dia 18 de novembro de 2014 às 19:46:07 foi realizada a primeira internação na unidade HC EPIB. O profissional que fez a solicitação foi: 4424, unidade de solicitação: Unidade Mista de Saúde Leôncio do Nascimento, município de solicitação: Serra Azul/SP, diagnóstico principal: outros transtornos psicóticos agudos essencialmente delirantes e sem comorbidades. O tempo entre a liberação da vaga e a internação do paciente no hospital: 11h18min, tempo de internação: 12 dias e 11h e transferência intra-hospitalar realizada em: 29

de novembro de 2014 às 11h01min09s; Pós alta o paciente foi para Serra Azul/SP, com diagnóstico principal da alta de esquizofrenia paranoide. O tipo de alta: Clínica e o profissional que realizou a alta foi: 4199.

No dia 05 de maio de 2015 às 18:05:00 foi realizada a segunda internação na unidade CAIS NAIC. O profissional que fez a solicitação foi: 4424, unidade de solicitação: Unidade Mista de Saúde Leôncio do Nascimento, município de solicitação: Serra Azul/SP, diagnóstico principal: esquizofrenia - transtornos esquizofrênicos e transtornos delirantes. O tempo entre a liberação da vaga e a internação do paciente no hospital: 10h04min, tempo de internação: 6 dias e 4h e não houve transferência intra-hospitalar realizada; Após a alta, o paciente foi para Serra Azul, com diagnóstico principal da alta de Psicose não orgânica - não especificada e diagnóstico secundário: transtorno efetivo bipolar, episódio atual maníaco com sintomas psicóticos. O tipo de alta: Clínica e o profissional que realizou a alta foi: 4401.

Figura 27 – Exemplo de *Timeline* do paciente.

Fonte: o autor.

No dia 11 de outubro de 2015 às 17:41:14 foi realizada a primeira internação na unidade CAIS NAIC. O profissional que fez a solicitação foi: 4424, unidade de solicitação: Unidade Mista de Saúde Leôncio do Nascimento, município de solicitação: Serra Azul/SP, diagnóstico

principal: outros transtornos psicóticos agudos essencialmente delirantes e sem comorbidades. O tempo entre a liberação da vaga e a internação do paciente no hospital: 9h06min, tempo de internação: 9 dias e 7h e não houve transferência intra-hospitalar. Após a alta o paciente foi para Serra Azul/SP, com diagnóstico principal da alta de esquizofrenia - transtornos esquizofrênicos e transtornos delirantes. O tipo de alta: Clínica e o profissional que realizou a alta foi: 4401.

No dia 29 de novembro de 2015 às 16:43:50 foi realizada a primeira internação na unidade HC EPIB. O profissional que fez a solicitação foi: 4423, unidade de solicitação: Unidade Mista de Saúde Leôncio do Nascimento, município de solicitação: Serra Azul, diagnóstico principal: esquizofrenia não especificada. O tempo entre a liberação da vaga e a internação do paciente no hospital: 14h20min, tempo de internação: 1 dias e 12h e transferência intra-hospitalar realizada em: 19 de dezembro de 2015 _às 12h04min56s. Após a alta o paciente foi para Serra Azul/SP, com diagnóstico principal da alta de esquizofrenia hebefrênica. O tipo de alta: transferência e o profissional que realizou a alta foi: 4786.

No dia 28 de dezembro de 2015 às 16:43:50 foi realizada a primeira internação na unidade HST Agudos Geral Feminino. O profissional que fez a solicitação foi: 4199, unidade de solicitação: HC EPIB, município de solicitação: Ribeirão Preto/SP, diagnóstico principal: esquizofrenia paranoide. O tempo entre a liberação da vaga e a internação do paciente no hospital: 13h, tempo de internação: 2 dias e 1h e não houve transferência intra-hospitalar. Após a alta o paciente foi para Serra Azul/SP, com diagnóstico principal da alta de esquizofrenia paranoide, diagnóstico secundário: transtorno afetivo bipolar atualmente em remissão. O tipo de alta: clínica e o profissional que realizou a alta foi: 4184.

No dia 26 de julho de 2016 às 20:50:58 foi realizada a primeira internação na unidade HST Primeiras Internações Feminino. O profissional que fez a solicitação foi: 4379, unidade de solicitação: Unidade Mista de Saúde Leôncio do Nascimento, município de solicitação: Serra Azul/SP, diagnóstico principal: esquizofrenia paranoide. O tempo entre a liberação da vaga e a internação do paciente no hospital: 13h, tempo de internação: 2 dias e 1h e não houve transferência intra-hospitalar. Após a alta o paciente foi para Serra Azul/SP, com diagnóstico principal da alta de esquizofrenia paranoide, diagnóstico secundário: transtorno afetivo bipolar atualmente em remissão. O tipo de alta: clínica e o profissional que realizou a alta foi: 4184. A Figura 28 apresenta o complemento da *timeline* do paciente código: 5240.

Figura 28 – Exemplo de *Timeline* do paciente.

11/08/2018 17:41:14	<p>Processo Tempo entre a liberação da vaga e a internação do paciente no hospital: 10hora(s),04minuto(s) Tempo de duração da internação: 06dia(s),04hora(s) Transferência intrahospitalar:</p> <p>Output Para onde foi o paciente após a alta? SERRA AZUL. Diagnóstico principal: Psicose não-orgânica não especificada / Diagnóstico secundário: Transtorno afetivo bipolar, episódio atual maníaco com sintomas psicóticos Tipo Alta: Clínica Profissional que fez a alta: 4461</p> <p>Unidade onde o paciente foi internado: CAUS NAIC</p>
29/11/2018 16:43:50	<p>Processo Tempo entre a liberação da vaga e a internação do paciente no hospital: 09hora(s),06minuto(s) Tempo de duração da internação: 09dia(s),07hora(s) Transferência intrahospitalar:</p> <p>Output Para onde foi o paciente após a alta? SERRA AZUL. Diagnóstico principal: Esquizofrenia - transtornos esquizotípicos e transtornos delirantes / Diagnóstico secundário: Tipo Alta: Clínica Profissional que fez a alta: 4401</p> <p>Unidade onde o paciente foi internado: HC EPIB</p> <p>Output Profissional solicitação: 4423 Unidade Solicitação: Unidade Mista de Saúde Leônido do Nascimento Município solicitação: Serra Azul Diagnóstico principal: Esquizofrenia não especificada / Diagnóstico secundário: Comorbidades: Traumatismo: NÃO / Problemas Respiratórios: NÃO / AVC: NÃO Consultas: NÃO / HAS: NÃO / Gravidez: Diabetes: NÃO / Doença Infecto: NÃO</p> <p>Processo Tempo entre a liberação da vaga e a internação do paciente no hospital: 14hora(s),20minuto(s) Tempo de duração da internação: 01dia(s),12hora(s) Transferência intrahospitalar: 19/12/2018 12:04:38</p> <p>Output Para onde foi o paciente após a alta? SERRA AZUL. Diagnóstico principal: Esquizofrenia hebefrênica / Diagnóstico secundário: Tipo Alta: Transferência Profissional que fez a alta: 4786</p> <p>Unidade onde o paciente foi internado: HST Apoio Geral Feminino</p>
28/12/2018 13:05:21	<p>Processo Tempo entre a liberação da vaga e a internação do paciente no hospital: 13hora(s),00minuto(s) Tempo de duração da internação: 02dia(s),01hora(s) Transferência intrahospitalar:</p> <p>Output Para onde foi o paciente após a alta? SERRA AZUL. Diagnóstico principal: Esquizofrenia paranóide / Diagnóstico secundário: Tipo Alta: Clínica Profissional que fez a alta: 4184</p> <p>Unidade onde o paciente foi internado: HST Primeiras Interações Feminino</p> <p>Output Profissional solicitação: 4199 Unidade Solicitação: HC EPIB Município solicitação: Ribeirão Preto Diagnóstico principal: Esquizofrenia paranóide / Diagnóstico secundário: Comorbidades: Traumatismo: NÃO / Problemas Respiratórios: NÃO / AVC: NÃO Consultas: NÃO / HAS: NÃO / Gravidez: Diabetes: NÃO / Doença Infecto: NÃO</p> <p>Processo Tempo entre a liberação da vaga e a internação do paciente no hospital: 13hora(s),00minuto(s) Tempo de duração da internação: 02dia(s),01hora(s) Transferência intrahospitalar:</p> <p>Output Para onde foi o paciente após a alta? SERRA AZUL. Diagnóstico principal: Esquizofrenia paranóide / Diagnóstico secundário: Transtorno afetivo bipolar atualmente em remissão Tipo Alta: Clínica Profissional que fez a alta: 4184</p> <p>Unidade onde o paciente foi internado: HST Primeiras Interações Feminino</p>
26/07/2016 20:50:58	<p>Processo Tempo entre a liberação da vaga e a internação do paciente no hospital: 13hora(s),00minuto(s) Tempo de duração da internação: 02dia(s),01hora(s) Transferência intrahospitalar:</p> <p>Output Para onde foi o paciente após a alta? SERRA AZUL. Diagnóstico principal: Esquizofrenia paranóide / Diagnóstico secundário: Tipo Alta: Clínica Profissional que fez a alta: 4184</p> <p>Unidade onde o paciente foi internado: HST Primeiras Interações Feminino</p> <p>Output Profissional solicitação: 4379 Unidade Solicitação: Unidade Mista de Saúde Leônido do Nascimento Município solicitação: Serra Azul Diagnóstico principal: Esquizofrenia paranóide / Diagnóstico secundário: Comorbidades: Traumatismo: NÃO / Problemas Respiratórios: NÃO / AVC: NÃO Consultas: NÃO / HAS: NÃO / Gravidez: Diabetes: NÃO / Doença Infecto: NÃO</p>

Fonte: o autor.

No dia 26 de setembro de 2016 às 19:09:19 foi realizada a primeira internação na unidade HC EPIB. O profissional que fez a solicitação foi: 4383, unidade de solicitação: HC Unidade de Emergência, município de solicitação: Ribeirão Preto/SP, diagnostico principal: Transtorno afetivo bipolar, episódio atual maníaco com sintomas psicóticos. O tempo entre a liberação da vaga e a internação do paciente no hospital: 11h16min, tempo de internação: 11

dias e 1h e houve transferência intra-hospitalar no dia 04/11/2016 às 11h15min. Após a alta o paciente foi para Serra Azul, com diagnóstico principal da alta de Transtorno afetivo bipolar, episódio atual maníaco com sintomas psicóticos. O tipo de alta: clínica e o profissional que realizou a alta foi: 4199.

No dia 28 de setembro de 2016 às 10:15:29 foi realizada a primeira internação na unidade HST Primeiras Internações Feminino. O profissional que fez a solicitação foi: 4383, unidade de solicitação: HC Unidade de Emergência, município de solicitação: Ribeirão Preto/SP, diagnóstico principal: Transtorno afetivo bipolar, episódio atual maníaco com sintomas psicóticos. O tempo entre a liberação da vaga e a internação do paciente no hospital: 18h04min, tempo de internação: 11 dias e 6h e não houve transferência intra-hospitalar. Após a alta o paciente foi para Serra Azul, com diagnóstico principal da alta de transtorno de personalidade com instabilidade emocional. O tipo de alta: clínica e o profissional que realizou a alta foi: 4184. A Figura 29 apresenta o complemento da *timeline* do paciente código: 5240.

Figura 29 – Exemplo de *Timeline* do paciente

Fonte: o autor.

Embora não seja essa a finalidade do estudo aqui apresentado, a mineração de processos possibilita também realizar estáticas descritiva. Por exemplo, o serviço da UPA Treze de Maio apresentou 645 encaminhamentos, unidade de solicitação, que representa 13.14% da internação

total. O município de origem de pacientes com maior número de internações apresentou 1561 registros (32.43%), o segundo 773 (16.06%) e o terceiro 233 (4.84%) no período analisado.

A análise com a mineração de processos permitiu também quantificar um total de 60% de pacientes do sexo masculino, 39.87 % feminino e 12% transexual. Desse acumulado, 63.87% são de etnia branca, 20.9% negra, 14.71% parda e 0.52 amarela, cuja internação voluntária ao longo do período analisado corresponde a 67,48%, involuntária 21.75% e compulsória 10.76%. Dentre as principais vias de internação destacam-se a “por prestador” que corresponde 69.98% dos casos.

A unidade de destino CAIS NAD Masculino, HC EPIB, HST Primeiras Internações Masculinas e CAIS NAT apontam um acumulado de 35.3% das internações. HST Atenção psicossocial feminino e CAIS-NAT misto aparecem em último, com acumulado de 0.21%. O tipo de alta com maior frequência observado foi a alta clínica (81.36%), seguido de a pedido” (7.04%), transferência da UCP (4.16%), transferência (3.93%), evasão (3.1%) e administrativa (0.42%).

6 DISCUSSÃO

Nesta sessão são inicialmente apresentados os trabalhos relacionados. A busca por trabalhos relacionados, publicados na literatura científica, envolveu a arquitetura de mineração de processos, aplicações seguindo a modelagem proposta pela fundação openEHR para elaboração de arquétipos clínicos, e avaliações de tecnologias de informação para o contexto da saúde mental. Em seguida as contribuições, achados e experiências adquiridas são discutidos assim como as dificuldades, limitações e possíveis trabalhos futuros decorrentes da tese apresentada.

6.1 Trabalhos Relacionados

A busca na literatura pelos trabalhos relacionados foi baseada nos seguintes descritores: mapa de saúde mental, mineração de processos, openEHR e *big data*. Em geral, as investigações realizadas na linha da ciência de dados trazem uma aproximação de um entendimento teórico de como é possível realizar o armazenamento de dados para operacionalizar indicadores em saúde mental. Trata-se de um grande desafio, pois a tradição de indicadores nesta área é mais restrita quando comparada a outras áreas do campo da saúde, como a atenção básica e a atenção hospitalar.

Conforme Huang, Lu e Duan (2012), a representação do conhecimento na área da saúde pode ser realizada através de um desenho dos processos clínicos, otimização de processos, decisão clínica, apoio e gestão de negócios, e assim por diante. Os resultados apresentados pelos autores retratam um esboço da linha do tempo de processos clínicos, obtido a partir de entrevistas e relatos da equipe médica do Hospital Central de Zhejiang Huzhou, na China. Os autores concluem que a abordagem de mineração realizada: (1) permite que as atividades clínicas sejam claramente distribuídas ao longo da linha do tempo do cuidado do paciente; (2) permite que certas relações temporais existentes entre as atividades sejam explicitadas; e (3) permite que um padrão de processo clínico enumere comportamentos médicos regulares que se espera que ocorram em jornadas de cuidado ao paciente, que servem como pontos de verificação para avaliação da qualidade do processo realizado. A Figura 30 apresenta o desenho final dos processos clínicos obtidos pela mineração.

Figura 30: Desenho dos processos clínicos

Fonte: Huang, Lu e Duan (2012)

Com a API desenvolvida, batizada de Easy-EHR, a pesquisa aqui relatada apresenta contribuições inovadoras em relação aos resultados apresentados por Huang, Lu e Duan (2012), disponibilizando um conjunto de interfaces para auxiliar a extração automática, a partir dos *logs* registrados no SISAM-13, de indicadores para apoiar a estruturação da matriz de saúde proposta por Thornicroft e Tansella (1999) e adaptada por Vinci et al. (2016) para a DRS XIII.

Rebuge e Ferreira (2012) aplicaram técnicas de *script* para extração dos *logs* de um banco de dados chamado Medtrix. Segundo os autores, a base do Medtrix contempla mais de 400 tabelas que não estão documentadas. A solução utilizada pelos autores foi selecionar manualmente e mover a parte relevante desse banco de dados para uma nova base, documentada, contendo os eventos dos cuidados de emergência de janeiro a julho de 2009, conforme ilustrado na Figura 31

Figura 31: Eventos dos cuidados de emergência

Fonte: Rebuge e Ferreira (2012)

A modelagem realizada do SISAM-13 para geração dos *logs* (reengenharia de software), traz contribuições relevantes em comparação com o trabalho de Rebuge e Ferreira (2012), pois a abordagem utilizada com o SISAM-13 possibilita a visão de toda a base investigada na camada de armazenamento de dados do próprio Sistema (camada de DATALAYER). Isso possibilitou um entendimento de como os dados são armazenados e como a geração de *logs* poderia ser realizada de maneira automatizada. Esse modelo gerado retrata as relações das tabelas que estão presentes no repositório de dados relacionadas aos episódios de solicitação de internação, internação e alta dos pacientes.

Mans et al. (2013) utilizaram uma metodologia orientada a processos para avaliar o impacto do uso do *software* CPN Tools para realizar simulação de processos na área da saúde, para investigar os impactos da odontologia digital. Para ilustrar as ferramentas CPN Tools, na Figura 32, é apresentado um *screenshot* em que um pequeno processo é simulado para possibilitar o entendimento das atividades realizadas dentro do processo de negócio da "prótese" e por quem as atividades são realizadas, os dados de execução foram obtidos de dois serviços holandeses. Os autores concluem que a ferramenta CPN foi adequada para possibilitar uma boa compreensão das atividades que são realizadas em todo o processo de tratamento odontológico

Figura 32: Eventos dos cuidados de emergência.

Fonte: Mans et al. (2013)

Os resultados obtidos na tese aqui apresentada corroboram os achados de Mans et al. (2013), porém com a utilização de uma outra ferramenta de mineração de processos que é o DISCO. Com funcionalidade similar a ferramenta CPN, foi possível extrair indicadores no nível de paciente, profissional e do departamento regional de saúde para representação da matriz de saúde.

Garcia, Ramirez e Larrea (2015), utilizaram mineração de processos para a extração de indicadores de permanência hospitalar: tempo máximo de permanência, tempo médio de permanência e tempo mínimo de permanência do paciente na unidade de saúde. O SIS utilizado foi o XAVIA HIS, desenvolvido pelo Centro de Informática Médica da UCI, o qual permite a coleta, o armazenamento, o processamento e interpretação de dados clínicos e administrativos, para todas as atividades relacionadas à instituição de saúde. O XAVIA HIS como sistema de informação possui um número considerável de dados armazenados que são utilizados para a tomada de decisões. A linguagem de programação utilizada foi Java, o Eclipse 3.4.2 Integrated Development Environment, o framework JBoss Seam 2.1.1 e para o mapeamento dos dados, o Hibernate 3.3.

Os resultados apresentados por Garcia, Ramirez e Larrea (2015) são compatíveis com os obtidos na pesquisa aqui apresentada. Porém, os métodos de desenvolvimento das soluções são diferentes. O EASY-EHR, implementado nessa pesquisa, adota o modelo canônico dual openEHR, possibilitando a separação entre a construção do modelo clínico dos indicadores, que deve ser feito com apoio do profissional de saúde, do modelo de gerenciamento de dados para composição da matriz, apresentando maior versatilidade e facilidade de ampliação e manutenção.

Orellana, Estrada e Alfonso (2018) desenvolveram um modelo de componentes de *software* para geração de *log* de eventos e uso de técnicas de mineração de processo para a detecção de variabilidade em processos hospitalares usando mineração de processos a partir de registros de eventos de SIS (Figura 33).

Figura 33: Modelo de componentes de software para melhoria de processos.

Fonte: Orellana, Estrada e Alfonso (2018)

Os autores avaliaram diferentes abordagens de mineração de processos e concluíram pela sua utilidade para apoiar a detecção de variabilidade nos processos hospitalares. Os resultados obtidos na pesquisa aqui relatada corroboram o entendimento dos autores em relação ao potencial do uso da mineração de processos para a detecção de variabilidade nos processos de atenção à saúde, embora esse não tenha sido o objetivo principal da investigação.

Embora modelada e implementada para o SISAM-13, a solução aqui apresentada para extração de indicadores por meio da mineração de processos é generalizável para qualquer sistema de informação do qual seja possível fazer a extração de *logs* através de *scripts* específicos, ou que já tenha os *logs* disponíveis através de uma trilha de auditoria. As trilhas de auditoria, em geral, devem conter informações relacionadas minimamente aos eventos do tipo criação, consulta, acréscimo ou substituição de registros do sistema de informação, importação e exportação de dados. Em geral, no que se refere aos sistemas de informação em saúde, exige-se a existência da trilha de auditoria para conformidade com modelos de certificação, como o desenvolvido pela Sociedade Brasileira de Informática em Saúde (SBIS) em parceria com o Conselho Federal de Medicina (CFM) para Sistemas de Registro Eletrônico em Saúde (SBIS, 2019). A certificação SBIS-CFM exige em seu requisito obrigatório de Nível Geral de Segurança NGS1.0.8 a presença da trilha de auditoria de *logs* do sistema. Dessa forma, a solução de mineração de processos aqui apresentada poderia ser imediatamente aplicada a qualquer sistema de informação em saúde que esteja em conformidade com a certificação SBIS-CFM.

6.2 Contribuições

A principal contribuição da pesquisa foi evidenciar a aplicabilidade da mineração de processos como método de apoio para a extração automática de indicadores para a construção da matriz de saúde, a partir de eventos armazenados em *logs* de sistemas de informação em saúde. Porém, para se chegar a essa contribuição principal, várias outras contribuições parciais foram também realizadas, a saber:

- Implementação do programa GDC-ARQUETYPES: API para realizar a sincronização de arquétipos clínicos do repositório CKM openEHR e gerenciar os arquétipos de *input*, *output* e *process*;
- Implementação do programa EDC-TEMPLATES: API editor de formulário clínico especializado que utiliza arquétipos do GDC-ARQUETYPES para construção da matriz de saúde mental;

- Implementação do programa EASY-EHR: *Software* que realiza persistência de formulários clínicos disponibilizados da API EDC-TEMPLATES e armazena em *big data*;
- Modelagem e criação do banco de dados MENTAL_HEALTH_OPENEHR: *big data* para armazenamento de indicadores para matriz de saúde mental utilizando o banco MondoDB;

6.3 Limitações e Trabalhos Futuros

Embora a modelagem realizada para a extração dos *logs* quando utilizada em o conjunto com os programas desenvolvidos (API) possibilite a construção da *timeline* do paciente e a partir desta a organização dos indicadores da matriz de saúde em diferentes níveis de granularidade, a saída dos dados apresentada pela ferramenta DISCO não possui uma representação visual amigável, necessitando de algum conhecimento específico para sua interpretação.

Essa limitação pode ser solucionada através da implementação de um novo conjunto de arquétipos que representem os diversos níveis geográficos da matriz de saúde. Porém, a modelagem dessa implementação é bastante complexa e provavelmente exigirá uma modelagem adicional dos *logs* em relação à que foi implementada para a geração da *timeline* do paciente. Sugere-se como trabalhos futuros a investigação dessa possibilidade.

7 CONCLUSÃO

O aumento de informações armazenadas em grandes repositórios de dados (*big data*), gera a oportunidade no setor da saúde de obter, a partir destas, conhecimento sobre o desempenho e real funcionamento de seus processos em rede de saúde. Nesse contexto, pode-se considerar a principal contribuição desta investigação o uso de mineração de processos para extração de indicadores de um sistema de informação para a estruturação de um modelo de matriz de saúde mental. A contribuição secundária, se apresenta por meio da utilização da representação openEHR como modelo canônico de dados, através da criação de novos arquétipos para concepção do conhecimento clínicos.

Esta pesquisa confirma também que a mineração de processos pode extrair automaticamente indicadores para a construção da matriz de saúde. A construção da *timeline* do paciente é a unidade de informação mais fina, que possibilita a construção dos diferentes níveis geográficos da matriz e extração de seus respectivos indicadores para caracterização da entrada (input), processo (process) e saída (output) nos seus diferentes níveis de granularidade.

REFERÊNCIAS

AGRAWAL, R.; GUNOPULOS, D.; LEYMAN, F. Mining process models from workflow logs. In: INTERNATIONAL CONFERENCE ON EXTENDING DATABASE TECHNOLOGY, 6., 1998, Valencia. **Proceedings...** Berlin: Springer, 1998. p. 469-483. Disponível em: <<https://link.springer.com/chapter/10.1007/BFb0101003>>. Acesso em: 18 abr. 2017.

ANICETO, R.C.; XAVIER, R.F. **Um Estudo sobre a utilização do banco de dados NoSQL cassandra em dados biológicos**. 2014. 61p. Monografia (Graduação) - Universidade de Brasília, Brasília, 2014.

AVILA, C. Exportação de sumários hospitalares do sistema SONHO utilizando a norma openEHR. In: IBERIAN CONFERENCE ON INFORMATION SYSTEMS AND TECHNOLOGIES, 6., 2011, Chaves. **Proceedings...** Piscataway: IEEE, 2011. 1 CD-ROM.

BEALE, T. **Archetype object model**. London: OpenEHR, 2008. Disponível em: <<http://www.openehr.org/svn/specification/TRUNK/publishing/architecture/am/aom1.4.pdf>>. Acesso em: 20 jun. 2017.

BERLER, A.; PAVLOPOULOS, S.; KOUTSOURIS, S. Design of an interoperability framework in a regional healthcare system. In: ANNUAL INTERNATIONAL CONFERENCE OF THE IEEE EMBS, 26., 2007, San Francisco. **Proceedings...** Piscataway: IEEE, 2007. p.3093–3096. Disponível em: <<http://ieeexplore.ieee.org/xpls/abs/all.jsp?arnumber=1403874>>. Acesso em: 15 jan. 2017.

BERNARD, G. et al. When sales meet process mining: a scientific approach to sales process and performance management research-in-progress. In: INTERNATIONAL CONFERENCE ON INFORMATION SYSTEMS, 37., 2016, Dublin. **Proceedings...** New York: AIS, 2016. Disponível em: <https://www.researchgate.net/publication/308940401_When_Sales_Meet_Process_Mining_A_Scientific_Approach_to_Sales_Process_and_Performance_Management_Research-in-Progress>. Acesso em: 26 abr. 2017.

BRAILER, D. Interoperability: the key to the future health care system. **Health Affairs**: web exclusive, v.24, Suppl.1, p.w5.19-w5.21, 2005. DOI: 10.1377/hlthaff.w5.19.

BRASIL. Ministério da Saúde. **Gestão de sistemas e tecnologias da informação em hospitais**: panorama, tendências e perspectivas em saúde. 2011a. Disponível em: <<http://bvsms.saude.gov.br/bvs/PoliticaInformacaoSaude29032004.pdf>>. Acesso em: 18 set. 2017.

BRASIL. Ministério da Saúde. **Portaria nº2.073, de 31 de agosto de 2011**. Regulamenta o uso de padrões de interoperabilidade e informação em saúde para sistemas de informação em saúde no âmbito do sistema único de saúde, nos níveis municipal, distrital, estadual e federal, e para os sistemas privados e do setor de saúde suplementar. 2011b. Disponível

em:<http://bvsmms.saude.gov.br/bvs/saudelegis/gm/2011/prt2073_31_08_2011.html>. Acesso em: 13 out. 2017.

BRASIL. Ministério da Saúde. **Plano de ações estratégicas para o enfrentamento das doenças crônicas não transmissíveis (dcnt) no Brasil**. 2011c. Disponível em:<http://bvsmms.saude.gov.br/bvs/publicacoes/plano_acoes_enfrent_dcnt_2011.pdf>. Acesso em: 18 jan. 2018.

CARON, F. et al. A Process mining-based investigation of adverse events in care processes health. **Health Information Management**, v.43, n.1, p.16-25, 2014. Disponível em:<<https://www.ncbi.nlm.nih.gov/pubmed/27010685>>. Acesso em: 18 jan. 2018.

CÉSAR, H.V.; AZEVEDO-MARQUES, P.M. **Projeto GDC – arquetypes**. São Paulo: AUSPIN - Agência USP de Inovação, 2017. Disponível em:<http://www.patentes.usp.br/tech/PROJETO_GDC_-_ARQUETYPES>. Acesso em: 18 fev. 2018.

CÉSAR, H.V.; AZEVEDO-MARQUES, P.M. **EDC _ templates**. São Paulo: AUSPIN - Agência USP de Inovação, 2017. Disponível em:<http://www.patentes.usp.br/tech/EDC_-_TEMPLATES>. Acesso em: 18 fev. 2018.

CHARLSON, F.J. et al. The Burden of mental, neurological, and substance usedisorders in China and India: a systematic analysis of communityrepresentative epidemiological studies. **Lancet**, v.388, n.10042, p.376–389, July 2016.

CHAUDHURI, S.; DAYAL, U. **An Overview of data warehousing and OLAP technology**. 2000. Disponível em:<<https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/sigrecord.pdf>>. Acesso em: 28 mai. 2018.

COOK, J.E.; WOLF, A.L. Automating process discovery through event-data analysis. In: INTERNATIONAL CONFERENCE ON SOFTWARE ENGINEERIN, 17., 1995, Seattle. **Proceedings...** New York:ACM, 1995. p.73-82. Disponível em:<<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5071093&isnumber=5071086>>. Acesso em: 18 abr. 2018.

DISCO ver your processes. Disponível em :<<https://fluxicon.com/disco/>>. Acesso em: 11 jun. 2019.

EATON, J. et al. Scale up of services formental health in low-income and middle-income countries. **Lancet**, v.378, n.9802, p.1592–1603, Oct. 2011.

ELIAS, D. **Entendendo o processo de ETL**. 2014. Disponível em:<<http://corporate.canaltech.com.br/noticia/business-intelligence/entendendo-o-processo-de-etl-22850/>>. Acesso em: 18 set. 2018.

ERDOGAN, T. G.; TARHAN, A. Systematic mapping of process mining studies in healthcare. **IEEE Access**, v.6, p.24543-24567, 2018. Disponível em:<<https://ieeexplore.ieee.org/document/8352636>>. Acesso em: 08 abr. 2018.

FREDRICKSON, B.L. et al. Open hearts build lives: positive emotions, induced through loving-kindness meditation, build consequential personal resources. **Journal of Personality and Social Psychology**, v.95,n.5, p.1045-1062, Nov. 2008. Disponível em:<<http://psycnet.apa.org/doiLanding?doi=10.1037%2Fa0013262>>. Acesso em: 18 jan. 2019.

FREUDENREICH, T. et al. An On-Demand ELT architecture for real-time BI. In: CASTELLANOS, M.; DAYAL, U.; RUNDENSTEINER, E.A. (Ed.). **Enabling real-time business intelligence**. Berlin: Springer Berlin Heidelberg, 2012. p.50–59. DOI: https://doi.org/10.1007/978-3-642-39872-8_4.

GARCIA, A.O.; RAMIREZ, Y.E.P.; LARREA, O.U.A. Process mining in healthcare: analysis and modeling of processes in the emergency area. **IEEE Latin America Transactions**, v.13, n.5, p.1612-1618, May 2015. Disponível em:<<https://ieeexplore.ieee.org/document/7112022>>. Acesso em: 23 abr. 2019.

GARCIA, A.O.; SENTI, V.E.; ALFONSO, D.P. MVD: a model for the detection of variability in hospital processes using process mining. **IEEE Latin America Transactions**, v.16, n.3, p.966-974, Mar. 2018. Disponível em:<<https://ieeexplore.ieee.org/document/8358680>>. Acesso em: 25 abr. 2019.

GARTNER. **Gartner says solving ‘big data’ challenge involves more than just managing volumes of data**. Stamford, 2011. Disponível em:<<http://www.gartner.com/newsroom/id/1731916>>. Acesso em: 29 abr. 2019.

GODOY, A. S. Pesquisa qualitativa tipos fundamentais. 1995. Disponível em: <https://virtual.ufmg.br/20191/pluginfile.php/206135/mod_resource/content/1/Godoy_1995_Pesquisa-qualitativa--tipos-fu_12736.pdf>. Acesso em: 18 mai. 2019.

GLOVER, G. A Comprehensive clinical database for mental health care in England. **Social Psychiatric & Psychiatric Epidemiology**, v.35, n.11,p.523-529, Dec. 2000. DOI: 10.1007/s001270050275.

GOLDBERG, D.W. **A Geocoding best practices guide** [Internet]. Springfield: North American Association of Central Cancer Registries, 2008. Disponível em:<https://20tqtx36s1la18rvn82wcmnpn-wpengine.netdna-ssl.com/wp-content/uploads/2016/11/Geocoding_Best_Practices.pdf>. Acesso em:28 ago. 2017.

GOLDBERG, D.W.; WILSON, J.P.; KNOBLOK, C.A. From text to geographic coordinates: the current state of geocoding. **URISA Journal**, v.19, n.1, p.33-46, Jan. 2007.

GROVES, P. et al. **The ‘big data’ revolution in healthcare: accelerating value and innovation**. New Jersey: McKinsey, 2013. APUD HERMON; WILLIAMS (2014)

GUTIÉRREZ, P.P.; CARRASCO, L. Open EHR-gen framework generador de sistemas normalizados de história clínica electrónica basados en el estándar openEHR. In: JORNADAS DE SISTEMAS DE INFORMACIÓN EN SALUD DEL HOSPITAL ITALIANO DE BUENOS AIRES, 2013, Buenos Aires. **Oficina do Programa de Internacionalização do OpenEHR**. Buenos Aires, 2013. Disponível em: <<http://informatica-medica.blogspot.com.br/2013/12/talleres-de-openehr-enhiba-2013.html>>. Acesso em: 18 abr. 2017.

HAAS, L. (2013). **IBM AL maden Research Centre. Paper presented at the HISA Big Data 2013 Melbourne Presentation**, Melbourne;

HAN, J. et al. **Survey on nosql database**. In: IEEE INTERNATIONAL CONFERENCE ON PERVASIVE COMPUTING AND APPLICATIONS, 6., 2014, Budapest. **Proceedings...** Piscataway: IEEE, 2014. p.363–366.

HEARD, S. The Use of standard content specifications in a national health interoperability framework. **Electronic Journal of Health Informatics**, v.5, n.1, e.3, 2010. Special Issue Systemic Interoperability. Disponível em: <<http://www.ejhi.net>>. Acesso em: 18 mar. 2019.

HECHT, R.; JABLONSKI, S. Nosql evaluation - a use case oriented survey. In: INTERNATIONAL CONFERENCE ON CLOUD AND SERVICE COMPUTING, 2011, Hong Kong. **Proceedings...** Piscataway: IEEE, 2011. p.336–341. DOI: 10.1109/CSC.2011.6138544.

HERMON, R.; WILLIOAMS, P.H.H. **Big data in healthcare: what is it used for?** Perth: Cowan University, 2014.

HIMSS. **21st Annual 2010 HIMSS Leadership Survey**, 2010. Disponível em: <https://www.himss.org/21st-annual-2010-himss-leadership-survey-results>. Acesso em: 18 abr. 2018.

HINCAPIE, A.; WARHOLAK, T. The Impact of health information exchange on health outcomes. **Applied Clinical Informatics**, v.2, n.4, p.499-507, 2011. DOI: 10.4338/ACI-2011-05-R-0027.

HUANG, Z.; LU, X.; DUAN, H. On Mining clinical pathway patterns from medical behaviors. **Artificial Intelligence in Medicine**, v.56, n.1, p.35-50, Sept. 2012. DOI: <https://doi.org/10.1016/j.artmed.2012.06.002>.

HUANG, Z. et al. Summarizing clinical pathways from event logs. **Journal of Biomedical Informatics**, v.46, n.1, p.111-127, Feb. 2013. DOI: <https://doi.org/10.1016/j.jbi.2012.10.001>.

JAMAL, A.; MCKENZIE, K.; CLARK, M. The Impact of health information technology on the quality of medical and healthcare: a systematic review. **Health Information Management**, v.38, n.3, p.26–38, 2009.

JORDANS, M.J. et al. Indicators for routine monitoring of effective mental healthcare coverage in low-and middle-income settings: a delphi study. **Health Policy and Planning**, v.31, n.8, p.1100–1006, Oct. 2016.

KIMBALL, R.; CASERTA, J. **The Data warehouse ETL toolkit**. New York: Wiley, 2004.

KONDO, K. et al. Mineração de processos para a identificação de padrões de falhas para a racionalização de alarmes. In: CONGRESSO BRASILEIRO DE Automática, 19., 2012, Campina Grande. **Anais...** Campinas: SBA, 2012. P.2958-2965. Disponível em:<https://www.researchgate.net/publication/281295061_Mineracao_de_Processos_para_a_Identificacao_de_Padrees_de_Falhas_para_a_Racionalizacao_de_Alarmes>. Acesso em: 18 abr. 2018.

KPOBI, L.; SWARTZ, L.; OFORI-ATTA, A.L. Challenges in the use of the mental health information system in a resource-limited setting: lessons from Ghana. **BMC Health Services Research**, v.18, n.1, p.98, Feb. 2018. DOI: 10.1186/s12913-018-2887-2.

LEBDAOUI, I.; ORHANOU, G.; ELHAJJI, S. An Integration adaptation for real-time data ware housing. **International Journal of Software Engineering and Its Applications**, v.8, n.11, p.115–128, 2014. DOI: <https://doi.org/10.14257/ijseia.2014.8.11.10>.

LOWE, B.H.; SUGARMAN, B. Design considerations for community mental health management information systems. **Community Mental Health Journal**, v.14, n.3, p.216-23, Fall, 1978.

MANS, R. et al. A Process-oriented methodology for evaluating the impact of IT: a proposal and an application in healthcare. **Information Systems**, v.38, n.8, p.1097-1115, Nov. 2013. DOI: 10.1016/j.is.2013.06.005.

MARTÍNEZ-COSTA, C. et al. A Model-driven approach for representing clinical archetypes for semantic web environments. **Journal of Biomedical Informatics**, v.42, n.1, p.150-164, 2008.

MARTÍNEZ-COSTA, C. et al. Semantic Web Technologies for managing EHR - related clinical knowledge. In: WU, G. (Ed.). **Semantic Web**. Rijeka: InTech, 2010. P.201-218. Disponível em:<http://cdn.intechopen.com/pdfs/9396/InTechSemantic_web_technologies_for_managing_eh_related_clinical_knowledge.pdf>. Acesso em: 10 jun. 2019.

MATTA, G.C.; MORENO, A.B. Saúde global: uma análise sobre as relações entre os processos de globalização e o uso dos indicadores de saúde. **Interface: comunicação, saúde, educação**, v.18, n.48, p.9-22, 2014. DOI: <http://dx.doi.org/10.1590/1807-57622014.0230>.

MEDEIROS, A.K.A. **Genetic process mining**. 2006. Thesis (PhD) - Eindhoven University of Technology, Eindhoven, 2006.

MENDES, E.V. **As Redes de atenção à saúde**. Brasília: Organização Pan-Americana da Saúde, 2011.

MIYOSHI, N.S. et al. **eHealth-interop**: an interoperability platform developed for mental care integration. 2018. DOI: 10.2196/10129.

MONGODB. MongoDB: The most popular database for modern apps. 2019. Disponível em: <<https://www.mongodb.com/>>. Acesso em: 24 jun. 2019.

MORENO, A.B. Olá, base de dados, poderia me dar uma indicação?. In: MATTOS, R.A.; BAPTISTA, T.V.F. (Org.). **Caminhos para análise das políticas de saúde**. Porto Alegre: Rede Unida, 2015 .

NATIONAL E-HEALTH TRANSITION AUTHORITY. **Review of shared electronic health record standards** - version 1.0. Sydney: NEHTA, 2006. Disponível em:<http://www.nehta.gov.au/component?option=com_docman/task,catview/gid,130/Itemid,139/>. Acesso em: 15 nov. 2018.

NEUMUTH, T. et al. Analysis of surgical intervention populations using generic surgical process models. **International Journal of Computer Assisted Radiology and Surgery**, v.6, n.1, p.59-71, Jan. 2011. DOI: <https://doi.org/10.1007/s11548-010-0475-y>.

OPENEHR. **Terminology and openEHR**. 2019. Disponível em:<<http://www.openehr.org/wiki/display/term/Terminology+and+openEH>>. Acesso em: 18 abr. 2018.

OPENEHR-CKM. **Clinical knowledge manager**. 2019. Disponível em:<<http://www.openehr.org/ckm/>>. Acesso em: 10 fev. 2019.

ORELLANA GARCÍA, A.; PÉREZ ALFONSO, D.; LARREA ARMENTEROS, O.U. Analysis of hospital processes with process mining techniques. In: SARKAR, I.N. et al. **MEDINFO 2015: e Health-enabled health**. Landale: IOS Press, 2015. P.310-314. (Studies in Health Technology, 216). Disponível em:<http://www.ewh.ieee.org/reg/9/etrans/ieee/issues/vol16/vol16issue03March2018/16TLA3_38OrellanaGarcia.pdf>. Acesso em: 22 dez. 2018.

PATRICK, J.; LY, R.; TRURAN, D. Evaluation of a persistent store for openEHR. In: WESTBROOK, J.; CALLEN, J. (Ed.). **HIC 2006 bridging the digital divide: clinician, consumer and computer**. Sydney: Health Informatics Society of Australian, 2006.

PEÇANHA, C.P.; BAX, M.P. Implementando o prontuário eletrônico openehr em sistemas gestores de conteúdo: uma aproximação. **Tendências da Pesquisa Brasileira em Ciência da Informação**, v.8, n.2, p.321-339, jul./dez. 2015. Encontro Nacional de Pesquisa em Ciência da Informação, 16., 2015, João Pessoa. Disponível em:<www.brapci.inf.br/index.php/article/download/43726>. Acesso em: 18 jun. 2018.

PEREIRA, P.K. et al. Transtornos mentais e comportamentais no sistema de informações hospitalares do SUS (SIH-SUS) no estado do Rio de Janeiro no período de 1999 a 2010. 2016. **Cadernos Saúde Coletiva**, v.20, n.4, p.482-491, 2012. DOI: <http://dx.doi.org/10.1590/51414-462x2012000400012>.

PIRES, D. F. **Uma solução interoperável, baseada na umls, para apoiar a decisão diagnóstica colaborativa na web**. 2007. 166p. Tese (Doutorado) – Faculdade de Filosofia Ciências e Letras de Ribeirão Preto, Ribeirão Preto, 2007. Disponível em: <<http://www.projedoradix.com.br/arq/artigo/teses/PiresTesesUsp.pdf>>. Acesso em: 10 mar. 2017.

REBUGE, A.; FERREIRA, D.R. Business process analysis in healthcare environments: a methodology based on process mining. **Information Systems**, v.37,n.2, p.99-116, Apr. 2012.

Disponível em:<<https://www.sciencedirect.com/science/article/abs/pii/S0306437911000044>>. Acesso em: 23 ago. 2019.

RIBEIRO, V. **O que é ETL?** 2011. Disponível em: <<https://vivianeribeiro1.wordpress.com/2011/06/28/o-que-e-etl-2/>>. Acesso em: 23 abr. 2018.
dia mês ano.

RIGOLETA, D.M.D.; TIMOTHY, W.C.; SERGIO, M.F. Modeling healthcare authorization and claim submissions using the openEHR dual-model approach. **BCM Medical Informatics & Decision Making**, v.11, p.60-71, 2011. Disponível em:<<http://www.biomedcentral.com/content/pdf/1472-6947-11-60.pdf>>. Acesso em: 11 jan. 2019.

ROBERTS, K. et al. **Overview of the TREC 2016 clinical decision support track**. Disponível em:<<https://trec.nist.gov/pubs/trec25/papers/Overview-CL.pdf>>. Acesso em: 25 ago. 2018.

ROJAS, E. et al. Process mining in healthcare: a literature review. **Journal of Biomedical Informatics**, v.61, p.224-236, June 2016. DOI: <https://doi.org/10.1016/j.jbi.2016.04.007>.

ROTZSCH, J.M.P. **Agência Nacional de Saúde Suplementar**. 2010. Disponível em:<http://www.ans.gov.br/portal/upload/biblioteca/memoriaeventos/20090630jussara08062009_TISS.pdf>. Acesso em: 28 jun. 2018.

SADALAGE, P.J.; FOWLER, M. **NoSQL essencial**. São Paulo: Novatec, 2013.

SANTOS, M.Y. Data models in NoSQL databases for big data contexts. In: TAN, Y.; SHI, Y. (Ed.). **Data mining and big data**. Cham: Springer International, 2001. p.475–485. DOI: https://doi.org/10.1007/978-3-319-40973-3_48.

SCHEDLBAUER, M. **The Art of business process modeling: the business analyst's guide to process modeling with UML & BPMN**. Scotts Valley: CreateSpace, 2010.

SILVEIRA, I.H.; OLIVEIRA, B.F.A.; JUNGUER, W.L. Uso de google maps para georreferenciación de datos del sistema de informaciones sobre mortalidad en el municipio de Río de Janeiro, Brasil, 2010-2012. **Epidemiologia e Serviços da Saúde**, v.26, n.4, p. 881-886, out./dez. 2017. DOI: 10.5123/51679-49742017000400018.

SOCIEDADE BRASILEIRA DE INFORMÁTICA EM SAÚDE. **Manual de certificação para sistemas de registro eletrônico em saúde: versão 4.3**, instituído e regido pela resolução CFM nº 1821/2007. São Paulo: SBIS, 2019. Disponível em:<http://www.sbis.org.br/certificacao/Manual_Certificacao_SBIS-CFM_2019_v4-3.pdf>. Acesso em: 15 jul. 2019.

SONG, M.; GÜNTHER, C.; AALST, W.M.P. Trace clustering in process mining. In: WORKSHOP ON BUSINESS PROCESS INTELLIGENCE, 4., 2008, Milan. **Proceedings...** Milan: Politecnico di Milano, 2008. p.51-62. Disponível em:<<https://research.tue.nl/en/publications/trace-clustering-in-process-mining-2>>. Acesso em: 18 abr. 2018.

STEWART, R.; DAVIS, K. **'Big data' in mental health research**: current status and emerging possibilities. 2016. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4977335/>>. Acesso em: 18 fev. 2018.
dia mês ano.

THORNICROFT, G.; TANSELLA, M. **The Mental health matrix**. Cambridge: Cambridge University Press, 1999.

VAISMAN, A.; ZIMÁNYI, E. Data warehouses: next challenges. In: AUFAURE, M.-A.; ZIMÁNYI, E. (Ed.). **Business intelligence**. New York: Springer, 2011. Cap.1, p.1-26. (Lecture Notes in Business Information Processing, 96). DOI: https://doi.org/10.1007/978-3-642-27358-2_1.

VAN DER AALST, W.M.P. **Process mining**: discovery, conformance and enhancement of business processes. New York: Springer, 2011. Disponível em: <<https://www.springer.com/us/book/9783642193453>>. Acesso em: 29 fev. 2017.

VAN DER AALST, W.M.P.; WEIJTERS, A.J.M.M. Process mining: a research agenda. **Computers in Industry**, v.53, n.3, p.231-244, Apr. 2004.

VAN GENUCHTEN, M. et al. Is Your upgrade worth it? Process mining can tell. **IEEE Software**, v.31, n.5, p.94-100, Sept./Oct. 2014. DOI: 10.1109/ms.2014.20.

VEST, J.R.; JASPERSON, J. What should we measure? Conceptualizing usage in health information exchange. **Journal of the American Medical Informatics Association**, v.17, n.3, p.302-307, May 2010. DOI: <https://doi.org/10.1136/jamia.2009.000471>.

VINCI, A.L.T. et al. Proposal of an evaluation model for mental health care networks using information technologies for its management. *Procedia Computer Science*, v.100, p.826-831, 2016. DOI: <https://doi.org/10.1016/j.procs.2016.09.231>.

WAAS, F. et al. On-Demand ELT. architecture for right-time bi: extending the vision. **International Journal of Data Warehousing and Mining**, v.9, n.2, p.21-38, 2013. DOI: <https://doi.org/10.4018/jdwm.2013040102>.

WORLD HEALTH ORGANIZATION. **Developing health management information systems**: a practical guide for developing countries. Geneva: WHO, 2004. p.3. Disponível em: <http://www.wpro.who.int/health_services/documents/developing_health_management_information_systems.pdf>. Acesso em: 18 mar. 2018.

WORLD HEALTH ORGANIZATION. **World Health Organization Assessment Instrument for Mental Health Systems (WHO-AIMS)**. Geneva: WHO, 2005. Disponível em: <https://apps.who.int/iris/bitstream/handle/10665/70771/WHO_MSD_MER_05.2_eng.pdf?sequence=1>. 21 mai. 2018.

WORLD HEALTH ORGANIZATION. **Mental Health Atlas 2011**. Geneva: WHO, 2011. Disponível em: <https://apps.who.int/iris/bitstream/handle/10665/44697/9799241564359_eng.pdf?sequence=1>. Acesso em: 20 dez. 2018.

WORLD HEALTH ORGANIZATION. **Mental health action plan 2013–2020**. Geneva: WHO, 2013. Disponível em: <http://www.who.int/mental_health/publications/action_plan/en/>. Acesso em: 08 abr. 201. 9

YOSHIURA V.T. et al. A Web-based information system for a regional public mental healthcare service network in Brazil. **International Journal of Mental Health Systems**, v.11, 2017. DOI: 10.1186/s13033-016-0117-z.

ZHANG, Y. et al. Fault detection based on data mining theory. In: INTERNATIONAL WORKSHOP ON INTELLIGENT SYSTEMS AND APPLICATIONS, 2009, Wuhan. **Proceedings...** Piscataway: IEEE, 2009. DOI: 10.1109/IWISA.2009.5072864.

ZHOU, Z.; WANG, Y.; LI, L. Process mining based modeling and analysis of workflows in clinical care - a case study in a Chicago Outpatient Clinic. In: INTERNATIONAL CONFERENCE ON NETWORKING, SENSING AND CONTROL, 11., 2014, Miami. **Proceedings...** Piscataway: IEEE, 2014. DOI: 10.1109/icnsc.2014.6819692.

APÊNDICE A – Script para obtenção de log de internações

```

SELECT i.data_internacao as data_internacao_real, pe.cod_pessoa, pe.tipo_pessoa, pe.data_cadastro,
pe.sexo, pe.data_nascimento, pe.uf, pe.municipio, pe.logradouro, pe.profissao_ocupacao,
pa.cod_paciente, pa.unidade_cadastro, pa.idade_aprox, pa.estado_civil, pa.etnia,
pa.arranjo_domiciliar, r.cod_regulacao, r.cod_unidade_origem, r.motivo_solicitacao,
r.cod_unidade_regulacao, r.data_solicitacao, r.cod_unidade_destino, r.tipo_regulacao, r.status_reg,
r.status_reg_compl, r.via_direta, r.tipo_transferencia, r.cid10_diagnostico, r.cid10_seg_diagnostico,
r.data_lista_reg, r.data_vaga_reg, r.data_admissao_reg, r.data_lista_hospital, r.data_vaga_hospital,
r.data_internacao, r.data_cancel_solicitacao, r.data_transferencia, r.traumatismo,
r.prob_respiratorios, r.avc, r.convulsao, r.has, r.gravidez, r.diabetes, r.doenca_infecto,
r.tipo_internacao_de, r.data_expiracao, r.cod_cidade_residencia, r.cod_estado_residencia,
i.cod_internacao, i.cod_transf_1, i.cod_transf_2, i.cod_transf_alta, i.status_internacao, i.data_alta,
i.cid10_diagnostico_alta, i.cid10_seg_diagnostico_alta, i.tipo_alta, i.tipo_internacao,
unidade_solicitacao.nome as nome_unidade_solicitacao, unidade_destino.nome as
nome_unidade_destino, unidade_solicitacao.municipio as municipio_unidade_solicitacao,
pro_internacao.cod_profissional AS cod_profissional_internacao, pro_alta.cod_profissional AS
cod_profissional_alta, pro_solicitacao.cod_profissional AS cod_profissional_solicitacao
FROM internacoes AS i LEFT JOIN regulacoes AS r ON (r.cod_regulacao = i.cod_regulacao) LEFT JOIN
pessoas AS pe ON (r.cod_paciente = pe.cod_pessoa) LEFT JOIN pacientes AS pa ON (r.cod_paciente =
pa.cod_paciente) LEFT JOIN usuarios AS user_solicitacao ON (r.cod_solicitador =
user_solicitacao.cod_usuario) LEFT JOIN profissionais AS pro_solicitacao ON
(pro_solicitacao.cod_profissional = user_solicitacao.cod_pessoa) LEFT JOIN usuarios AS user_alta ON
(i.cod_profissional_alta = user_alta.cod_usuario) LEFT JOIN profissionais AS pro_alta ON
(pro_alta.cod_profissional = user_alta.cod_pessoa) LEFT JOIN pessoas AS pro_alta_pessoa ON
(pro_alta.cod_profissional = pro_alta_pessoa.cod_pessoa) LEFT JOIN usuarios AS user_internacao ON
(i.cod_profissional_alta = user_internacao.cod_usuario) LEFT JOIN profissionais AS pro_internacao ON
(user_internacao.cod_pessoa = pro_internacao.cod_profissional) LEFT JOIN pessoas AS
pro_internacao_pessoa ON (pro_internacao.cod_profissional = pro_internacao_pessoa.cod_pessoa)
LEFT JOIN unidades AS unidade_solicitacao ON (r.cod_unidade_origem =
unidade_solicitacao.cod_unidade) LEFT JOIN unidades AS unidade_destino ON
(r.cod_unidade_destino = unidade_destino.cod_unidade) WHERE i.data_internacao > '2012-01-01'
AND i.data_internacao < '2016-12-31'

```

APÊNDICE B – Script para checagem e tratamento dos dados

```
SELECT pessoas.*, profissionais.* FROM profissionais, pessoas WHERE profissionais.  
cod_profissional = pessoas.cod_pessoa
```

```
SELECT * FROM unidades;
```

APÊNDICE C – Codificação referente ao arquétipo: openEHR-EHR- CLUSTER.mental_health_matrix_inputs.v1.adl

```

archetype (adl_version=1.4)
  openEHR-EHR-CLUSTER.mental_health_matrix_inputs.v1

concept
  [at0000]      -- INPUTS
language
  original_language = <[ISO_639-1::pt]>
description
  original_author = <
    ["name"] = <" Hilton Vicente César;
                  João Mazzoncini de Azevedo Marques;
                  Newton Shydeo Brandao Miyoshi;
                  Paulo Mazzoncini de Azevedo Marques ">
    >
  details = <
    ["pt"] = <
      language = <[ISO_639-1::pt]>
      purpose = <"Contatos">
    >
  >
  lifecycle_state = <"Initial">
  other_contributors = <>
  other_details = <
    ["MD5-CAM-1.0.1"] = <"3BA9597ED8B89C50BEDC57385FFE211F">
  >

definition
  CLUSTER[at0000] matches {  -- INPUTS
    items cardinality matches {1..*; unordered} matches {
      ELEMENT[at0006] occurrences matches {0..1} matches {  -- ID
        value matches {
          DV_IDENTIFIER matches {*}
        }
      }
      ELEMENT[at0029] occurrences matches {0..1} matches {  --
        professional_solicitacao
          value matches {
            DV_TEXT matches {*}
          }
        }
      ELEMENT[at0024] occurrences matches {0..1} matches {  --
        data_solicitacao
          value matches {
            DV_DATE_TIME matches {
              value matches {yyyy-mm-ddTHH:MM:SS}
            }
          }
        }
      ELEMENT[at0025] occurrences matches {0..1} matches {  --
        cod_paciente
          value matches {
            C_DV_QUANTITY <

```

```

    >
    }
  }
  ELEMENT[at0027] occurrences matches {0..1} matches {      -- idade
    value matches {
      DV_TEXT matches {*}
    }
  }
  ELEMENT[at0028] occurrences matches {0..1} matches {      -- estado_civil
    value matches {
      DV_TEXT matches {*}
    }
  }
  ELEMENT[at0043] occurrences matches {0..1} matches {      -- sexo
    value matches {
      DV_TEXT matches {*}
    }
  }
  ELEMENT[at0030] occurrences matches {0..1} matches {      --
servico_solicitacao_internacao
    value matches {
      DV_TEXT matches {*}
    }
  }
  ELEMENT[at0031] occurrences matches {0..1} matches {      --
municipio_origem_paciente
    value matches {
      DV_TEXT matches {*}
    }
  }
  ELEMENT[at0032] occurrences matches {0..1} matches {      --
municipio_paciente
    value matches {
      DV_TEXT matches {*}
    }
  }
  ELEMENT[at0033] occurrences matches {0..1} matches {      --
diagnostico_principal
    value matches {
      DV_TEXT matches {*}
    }
  }
  ELEMENT[at0034] occurrences matches {0..1} matches {      --
traumatismo
    value matches {
      DV_TEXT matches {*}
    }
  }
  ELEMENT[at0036] occurrences matches {0..1} matches {      --
prob_respiratorio
    value matches {
      DV_TEXT matches {*}
    }
  }
  ELEMENT[at0037] occurrences matches {0..1} matches {      -- avc
    value matches {
      DV_TEXT matches {*}
    }
  }

```

```

    }
    ELEMENT[at0038] occurrences matches {0..1} matches {      -- convulsao
      value matches {
        DV_TEXT matches {*}
      }
    }
  }
  ELEMENT[at0039] occurrences matches {0..1} matches {      -- has
    value matches {
      DV_TEXT matches {*}
    }
  }
  ELEMENT[at0040] occurrences matches {0..1} matches {      -- gravidez
    value matches {
      DV_TEXT matches {*}
    }
  }
  ELEMENT[at0041] occurrences matches {0..1} matches {      -- diabetes
    value matches {
      DV_TEXT matches {*}
    }
  }
  ELEMENT[at0042] occurrences matches {0..1} matches {      --
doenca_infecto
    value matches {
      DV_TEXT matches {*}
    }
  }
}

ontology
  term_definitions = <
    ["pt"] = <
      items = <
        ["at0000"] = <
          text = <"INPUTS">
          description = <". ">
        >
        ["at0006"] = <
          text = <"ID">
          description = <"* Armazenar o código do profissional que fez
a solicitação">
        >
        ["at0024"] = <
          text = <"data_solicitacao">
          description = <"*Data da solicitação">
        >
        ["at0025"] = <
          text = <"cod_paciente">
          description = <"*Código do paciente">
        >
        ["at0027"] = <
          text = <"idade">
          description = <"*Idade">
        >
        ["at0028"] = <
          text = <"estado_civil">
          description = <"*Estado civil">

```

```

>
["at0029"] = <
    text = <"profissional_solicitacao">
    description = <"*Profissional que realizou a solicitação da
internação">
>
["at0030"] = <
    text = <"servico_solicitacao_internacao">
    description = <"*Serviço que solicitou a internação">
>
["at0031"] = <
    text = <"municipio_origem_paciente">
    description = <"*Município de onde veio o paciente">
>
["at0032"] = <
    text = <"municipio_paciente">
    description = <"*Município onde mora o paciente.">
>
["at0033"] = <
    text = <"diagnostico_principal">
    description = <"*Diagnóstico principal">
>
["at0034"] = <
    text = <"traumatismo">
    description = <"*Diagnóstico secundário">
>
["at0036"] = <
    text = <"prob_respiratorio">
    description = <"*">
>
["at0037"] = <
    text = <"avc">
    description = <"*">
>
["at0038"] = <
    text = <"convulsao">
    description = <"*">
>
["at0039"] = <
    text = <"has">
    description = <"*">
>
["at0040"] = <
    text = <"gravidez">
    description = <"*">
>
["at0041"] = <
    text = <"diabetes">
    description = <"*">
>
["at0042"] = <
    text = <"doenca_infecto">
    description = <"*">
>
["at0043"] = <
    text = <"sexo">
    description = <"*">
>>>>

```


APÊNDICE D – Codificação referente ao arquétipo: openEHR-EHR- CLUSTER.mental_health_matrix_process.v1.adl

```

archetype (adl_version=1.4)
  openEHR-EHR-CLUSTER.mental_health_matrix_process.v1

concept
  [at0000]      -- PROCESS
language
  original_language = <[ISO_639-1::pt]>
description
  original_author = <
    ["name"] = <"">
  >
  details = <
    ["pt"] = <
      language = <[ISO_639-1::pt]>
      purpose = <"Contatos">
      use = <" Hilton Vicente César;
        João Mazzoncini de Azevedo Marques;
        Newton Shydeo Brandao Miyoshi;
        Paulo Mazzoncini de Azevedo Marques ">
    >
  >
  lifecycle_state = <"Initial">
  other_contributors = <>
  other_details = <
    ["MD5-CAM-1.0.1"] = <"320A79DB5C8F4EB59AB1150C9EAF52B5">
  >

definition
  CLUSTER[at0000] matches {    -- PROCESS
    items cardinality matches {1..*; unordered} matches {
      ELEMENT[at0006] occurrences matches {0..1} matches {    -- ID
        value matches {
          DV_IDENTIFIER matches {*}
        }
      }
      ELEMENT[at0029] occurrences matches {0..1} matches {    --
tempo_liberacao_vaga
        value matches {
          DV_TIME matches {*}
        }
      }
      ELEMENT[at0030] occurrences matches {0..1} matches {    --
unidade_internacao
        value matches {
          DV_TEXT matches {*}
        }
      }
      ELEMENT[at0031] occurrences matches {0..1} matches {    --
transferencia_intrahospitalar
        value matches {
          DV_TEXT matches {*}
        }
      }
    }
  }

```

```

    }
    ELEMENT[at0033] occurrences matches {0..1} matches {
duracao_internacao
        value matches {
            DV_TIME matches {*}
        }
    }
}
}
}

ontology
term_definitions = <
    ["pt"] = <
        items = <
            ["at0000"] = <
                text = <"PROCESS">
                description = <"..">
            >
            ["at0006"] = <
                text = <"ID">
                description = <"* Armazenar o código do profissional que fez
a solicitação">
            >
            ["at0029"] = <
                text = <"tempo_liberacao_vaga">
                description = <"*">
            >
            ["at0030"] = <
                text = <"unidade_internacao">
                description = <"*">
            >
            ["at0031"] = <
                text = <"transferencia_intrahospitalar">
                description = <"*">
            >
            ["at0033"] = <
                text = <"duracao_internacao">
                description = <"*">
            >
        >
    >
>

```

APÊNDICE E – Codificação referente ao arquétipo: openEHR-EHR- CLUSTER.mental_health_matrix_outputs.v1.adl

```

archetype (adl_version=1.4)
  openEHR-EHR-CLUSTER.mental_health_matrix_outputs.v1

concept
  [at0000]      -- OUTPUTS
language
  original_language = <[ISO_639-1::pt]>
description
  original_author = <
    ["name"] = <" Hilton Vicente César;
                  João Mazzoncini de Azevedo Marques;
                  Newton Shydeo Brandao Miyoshi;
                  Paulo Mazzoncini de Azevedo Marques ">
    >
  details = <
    ["pt"] = <
      language = <[ISO_639-1::pt]>
      purpose = <"Contatos">
      use = <"">
      misuse = <"">
      copyright = <"">
    >
  >
  lifecycle_state = <"Initial">
  other_contributors = <>
  other_details = <
    ["MD5-CAM-1.0.1"] = <"BC4DC7D896EFA50DFF737E6B536AF649">
  >

definition
  CLUSTER[at0000] matches {  -- OUTPUTS
    items cardinality matches {1..*; unordered} matches {
      ELEMENT[at0006] occurrences matches {0..1} matches {  -- ID
        value matches {
          DV_IDENTIFIER matches {*}
        }
      }
      ELEMENT[at0027] occurrences matches {0..1} matches {  --
destino_paciente
        value matches {
          DV_TEXT matches {*}
        }
      }
      ELEMENT[at0028] occurrences matches {0..1} matches {  --
diagnostico_primario
        value matches {
          DV_TEXT matches {*}
        }
      }
      ELEMENT[at0030] occurrences matches {0..1} matches {  --
diagnostico_secundario
        value matches {
          DV_TEXT matches {*}
        }
      }
    }
  }

```

```

    }
  }
  ELEMENT[at0031] occurrences matches {0..1} matches {
    --
    value matches {
      DV_TEXT matches {*}
    }
  }
}
}
}

ontology
  term_definitions = <
    ["pt"] = <
      items = <
        ["at0000"] = <
          text = <"OUTPUTS">
          description = <"..">
        >
        ["at0006"] = <
          text = <"ID">
          description = <"* Armazenar o código do profissional que fez
a solicitação">
        >
        ["at0027"] = <
          text = <"destino_paciente">
          description = <"*Para onde foi o paciente após a alta?">
        >
        ["at0028"] = <
          text = <"diagnostico_primario">
          description = <"*Diagnóstico primário">
        >
        ["at0030"] = <
          text = <"diagnostico_secundario">
          description = <"*Diagnóstico secundário">
        >
        ["at0031"] = <
          text = <"profissional_alta">
          description = <"*Profissional que fez a alta">
        >
      >
    >
  >
>

```

APÊNDICE F –Aplicativo DISCO

A utilização do software Disco pode auxiliar na criação de mapas de processo inteligentes diretamente de logs através de métricas de processo que são projetadas diretamente em um mapa de atividades. A Figura 1 do Apêndice apresenta a projeção dos processos analisados através do aplicativo DISCO.

Figura 1 do Apêndice - ilustra a projeção de processos de log.

Fonte: DISCO (2019).

A Figura 2 do Apêndice é um exemplo de animação de mapa de processo. Com o Disco, você pode criar animações, visualizando seu processo como aconteceu, diretamente no seu mapa de processo. A animação pode ajudá-lo a identificar instantaneamente os afunilamentos onde o trabalho está se acumulando.

Figura 2 do Apêndice – ilustra a animação de mapa de processos de log.

Fonte: DISCO (2019).

A Figura 3 do Apêndice é um exemplo de uma estatística detalhada com informações sobre o log de processos. É possível observar a quantidade de vezes que uma atividade foi realizada

Figura 3 do Apêndice – ilustra uma estatística detalhada com informações sobre o log.

Fonte: DISCO (2019).

Através Figura 4 do Apêndice é possível acessar o ambiente de linha do tempo do caminho do processo.

Figura 4 do Apêndice – ilustra o ambiente de linha do tempo com informações sobre o log.

Fonte: DISCO (2019).

APÊNDICE G – GDC-ARQUETYPES

```

import java.io.*;
import javax.xml.ws.*;

public class ADLParser implements ADLParserConstants {
    private static final String CHARSET = "UTF-8";
    private static final String ATTRIBUTE_UNKNOWN = "unknown";
    private boolean missingLanguageCompatible;
    private boolean emptyPurposeCompatible;
    private MeasurementService measureServ;
    public ADLParserTokenManager token_source;
    SimpleCharStream jj_input_stream;
    public Token token;
    public Token jj_nt;
    private int jj_ntk;
    private Token jj_scanpos;
    private Token jj_lastpos;
    private int jj_las;
    private int jj_gen;
    private final int[] jj_lal;
    private static int[] jj_lal_0;
    private static int[] jj_lal_1;
    private static int[] jj_lal_2;
    private static int[] jj_lal_3;
    private static int[] jj_lal_4;
    private static int[] jj_lal_5;
    private static int[] jj_lal_6;
    private static int[] jj_lal_7;
    private final ADLParser.JJCalls[] jj_2_rtns;

    public ADLParser(File file) throws IOException {
        this(new FileInputStream(file), "UTF-8");
    }

    public ADLParser(String value) { this((Reader)(new BufferedReader(new StringReader(value)))); }

    public ADLParser(File file, boolean missingLanguageCompatible, boolean emptyPurposeCompatible) throws IOException {
        this(new FileInputStream(file), "UTF-8");
        this.missingLanguageCompatible = missingLanguageCompatible;
        this.emptyPurposeCompatible = emptyPurposeCompatible;
    }

    public ADLParser(String value, boolean missingLanguageCompatible, boolean emptyPurposeCompatible) {
        this((Reader)(new BufferedReader(new StringReader(value))));
        this.missingLanguageCompatible = missingLanguageCompatible;
        this.emptyPurposeCompatible = emptyPurposeCompatible;
    }

    public ADLParser(Reader reader, boolean missingLanguageCompatible, boolean emptyPurposeCompatible) {
        this((Reader)reader);
        this.missingLanguageCompatible = missingLanguageCompatible;
        this.emptyPurposeCompatible = emptyPurposeCompatible;
    }

    public final ConstraintRef constraint_ref_obj(String path, CAttribute parent) throws ParseException {
        String typeName = "CODE PHRASE";
        Interval occurrences = new Interval(Integer.valueOf(1), Integer.valueOf(1));
        Object nodeId = null;
        String reference = this.constraint_ref();
        return new ConstraintRef(path, typeName, occurrences, (String)nodeId, parent, reference);
    }

    public final ArchetypeInternalRef archetype_internal_ref(String path, CAttribute parent) throws ParseException {
        Interval occurrences = new Interval(Integer.valueOf(1), Integer.valueOf(1));
        String nodeId = null;
        this.jj_consume_token(80);
        String type = this.type_identifier();
        switch(this.jj_ntk == -1 ? this.jj_ntk() : this.jj_ntk) {
            case 158:
                nodeId = this.constraint_ref();
                break;
            default:
                this.jj_lal[101] = this.jj_gen;
        }
    }
}

```

@ServiceMode(value = javax.xml.ws.Service.Mode.PAYLOAD)

@WebServiceProvider(serviceName = "ArchetypeFinderBean", portName =
 "ArchetypeFinderBeanHttpSoap11Endpoint", targetNamespace = "http://archetypeServlet",
 wsdlLocation = "WEB-INF/wsdl/EHRWebServiceFromWSDL/openehr.org/ckm/services/ArchetypeFinderBean.wsdl")

```

public class EHRWebServiceFromWSDL implements
javax.xml.ws.Provider<javax.xml.transform.Source> {

    public javax.xml.transform.Source invoke(javax.xml.transform.Source source) {
        //TODO implement this method
        throw new UnsupportedOperationException("Not implemented yet.");
    }

    public static void main(String[] args) throws IOException
    {
        ArchetypeFinderBean afb = new ArchetypeFinderBean();
        List<String> allArchetypesIds = null;

        allArchetypesIds=afb.getArchetypeFinderBeanHttpSoap11Endpoint().getAllArchetypeIds();
        for(int i=0; i< allArchetypesIds.size(); i++){
            try{
                File file = new File("C:\\repositorio", allArchetypesIds.get(i)+".adl");
                if( file.createNewFile() ){
                    FileWriter fw = new
FileWriter("C:\\repositorio"+"\\"+allArchetypesIds.get(i)+".adl");

                    fw.write(afb.getArchetypeFinderBeanHttpSoap11Endpoint().getArchetypeInADL(allArchetype
esIds.get(i)));
                    fw.close();
                }else{
                    System.out.println("Arquivo não criado");
                }
            }catch(IOException ex){
                ex.printStackTrace();
            }
        }

    }
}

```



```

41 public void getAllArchetypeIds() throws FileNotFoundException, IOException {
42     File raiz = new File("C:\\repositorio");
43     File[] files = raiz.listFiles();
44
45     for(int i=0; i< files.length; i++){
46         File arquivoOrigem = new File("C:\\repositorio"+"\\"+files[i].getName());
47         File arquivoDestino = new File("C:\\repositorio2"+"\\"+files[i].getName());
48         FileChannel sourceChannel = null;
49         FileChannel destinationChannel = null;
50         try {
51             sourceChannel = new FileInputStream(arquivoOrigem).getChannel();
52             destinationChannel = new FileOutputStream(arquivoDestino).getChannel();
53             sourceChannel.transferTo(0, sourceChannel.size(),
54                                     destinationChannel);
55         } finally {
56             if (sourceChannel != null && sourceChannel.isOpen())
57                 sourceChannel.close();
58             if (destinationChannel != null && destinationChannel.isOpen())
59                 destinationChannel.close();
60         }
61     }
62 }
63 }

```

```

4
5
6
7 import archetypeserviet.ArchetypeFinderBean;
8 import java.io.*;
9 import java.util.List;
10 import javax.ejb.Stateless;
11 import javax.xml.ws.ServiceMode;
12 import javax.xml.ws.WebServiceProvider;
13
14 /**
15  *
16  * @author hilton
17  */
18
19 @ServiceMode(value = javax.xml.ws.Service.Mode.PAYLOAD)
20 @WebServiceProvider(serviceName = "ArchetypeFinderBean", portName = "ArchetypeFinderBeanHttpSoap11Endpoint", targetNamespace
21 public class EHDWebServiceFromWSDL implements javax.xml.ws.Provider<javax.xml.transform.Source> {
22
23     public javax.xml.transform.Source invoke(javax.xml.transform.Source source) {
24         //TODO implement this method
25         throw new UnsupportedOperationException("Not implemented yet.");
26     }
27
28     public static void main(String[] args) throws IOException
29     {
30         ArchetypeFinderBean afb = new ArchetypeFinderBean();

```

```

30 ArchetypeFinderBean afb = new ArchetypeFinderBean();
31 List<String> allArchetypesIds = null;
32 allArchetypesIds=afb.getArchetypeFinderBeanHttpSoap11Endpoint().getAllArchetypeIds();
33 for(int i=0; i< allArchetypesIds.size(); i++){
34     try{
35         File file = new File("C:\\repositorio", allArchetypesIds.get(i)+".adl");
36         if( file.createNewFile() ){
37             FileWriter fw = new FileWriter("C:\\repositorio"+"\\"+allArchetypesIds.get(i)+".adl");
38             fw.write(afb.getArchetypeFinderBeanHttpSoap11Endpoint().getArchetypeInADL(allArchetypesIds.get(i)));
39             fw.close();
40         }else{
41             System.out.println("Arquivo não criado");
42         }
43     }catch(IOException ex){
44         ex.printStackTrace();
45     }
46 }
47
48
49 }

```

```
File raiz = new File("C:\\repositorio");

File[] files = raiz.listFiles();

for(int i=0; i< files.length; i++){

    File arquivoOrigem = new File("C:\\repositorio"+"/"+files[i].getName());

    File arquivoDestino = new File("C:\\repositorio2"+"/"+files[i].getName());

    FileChannel sourceChannel = null;

    FileChannel destinationChannel = null;

    try {

        sourceChannel = new FileInputStream(arquivoOrigem).getChannel();

        destinationChannel = new FileOutputStream(arquivoDestino).getChannel();

        sourceChannel.transferTo(0, sourceChannel.size(),

            destinationChannel);

    } finally {

        if (sourceChannel != null && sourceChannel.isOpen())

            sourceChannel.close();

        if (destinationChannel != null && destinationChannel.isOpen())

            destinationChannel.close();

    }

}
```

}

The screenshot shows an IDE interface. At the top, a project tree for 'GDC-ARQUETYPES' is visible, with a bracket spanning from the tree to the bottom section. The tree includes folders for 'Web Pages', 'Source Packages', 'br.com', 'Generated Sources (jax-ws)', 'Libraries', 'Enterprise Beans', and 'Web Services'. Under 'Web Services', the 'ArchetypeFinderBean' is selected.

Below the tree, the details for the 'ArchetypeFinderBean' web service are shown, specifically for the 'HttpSoap11Endpoint'. Two operations are listed:

getArchetypeIds

Parameters		Description
Parameter Name	Parameter Type	
searchParams	java.util.List<java.lang.String>	

getDescriptionForArchetypes

Parameters		Description
Parameter Name	Parameter Type	
archetypeIds	java.util.List<java.lang.String>	
language	java.lang.String	
descriptionTerms	java.util.List<java.lang.String>

INPIINSTITUTO
NACIONAL
DA PROPRIEDADE
INDUSTRIAL**INPI**
Sistema
SISTEMA

REPÚBLICA FEDERATIVA DO BRASIL
Ministério Da Indústria, Comércio Exterior e Serviços
Instituto Nacional da Propriedade Industrial
Diretoria de Patentes, Programas de Computador e Topografias de Circuitos Integrados

Certificado de Registro de Programas de Computador

Processo nº: BR 51 2017 000882-5

O Instituto Nacional da Propriedade Industrial expede o presente certificado de Registro de Programas de Computador, válido por 50 anos a partir de 1º de janeiro subsequente à data de Criação: 09 de março de 2015, em conformidade com o parágrafo 2º, artigo 2º da Lei Nº 9.609, de 19 de Fevereiro de 1998.

Título: **GDC-ARQUETYPES**
Data de Criação: 09 de março de 2015
Titular(es): UNIVERSIDADE DE SÃO PAULO
Autor(es): HILTON VICENTE CESAR
/ PAULO MAZZONCINI DE AZEVEDO MARQUES
Linguagem: JAVA
Campo de Aplicação: IF-07, SD-03, SD-08
Tipo Programa: DS-01, GI-01, GI-05, SO-02
Expedido em: 01 de agosto de 2017
Aprovado por Liane Elizabeth Caldeira Lage

Liane Elizabeth Caldeira Lage
Diretor(a) Substituto(a) de Patentes, Programas de Computador e Topografia de Circuitos Integrados

APÊNDICE H – EDC-TEMPLATES

```

TemplateManager.getInstance().setArcheTypesDirectory("${grails
Application.config.ckm.directory.archetypes}")
TemplateManager.getInstance().setTemplatesDirectory("${grailsA
pplication.config.ckm.directory.templates}")

def templatesDirectory =
grailsApplication.config.ckm.directory.templates
System.out.println("templatesDirectory"+templatesDirectory)

def dir = new File(templatesDirectory);
def templates = [];
def name = "template"
def template = [:]

TEMPLATE ehrTemplate =
TemplateManager.getInstance().readTemplate(name)
template.name = name
template.type =
TemplateManager.getInstance().getType(ehrTemplate)

templates << template
return templates

package br.org.usp.openehr.models;

import org.openehr.am.archetype.constraintmodel.CObject;

public abstract class ArchetypeObjectCallback {

    private CObject object;
    private Object param;

    public ArchetypeObjectCallback() {

    }

    public void registerObjectAndParam(CObject object, Object
param) {
        this.object = object;
        this.param = param;
    }

    public CObject getObject() {
        return this.object;
    }

    public Object getParam() {

```

```

        return this.param;
    }

    public abstract void processArchetypeObject();
}

```

REPÚBLICA FEDERATIVA DO BRASIL
Ministério Da Indústria, Comércio Exterior e Serviços
Instituto Nacional da Propriedade Industrial

Diretoria de Patentes, Programas de Computador e Topografias de Circuitos Integrados

Certificado de Registro de Programas de Computador

Processo nº: BR 51 2017 000883-3

O Instituto Nacional da Propriedade Industrial expede o presente certificado de Registro de Programas de Computador, válido por 50 anos a partir de 1º de janeiro subsequente à data de Criação: 11 de outubro de 2014, em conformidade com o parágrafo 2º, artigo 2º da Lei Nº 9.609, de 19 de Fevereiro de 1998.

Título: **EDC-TEMPLATES**

Data de Criação: 11 de outubro de 2014

Titular(es): UNIVERSIDADE DE SÃO PAULO

Autor(es): HILTON VICENTE CESAR
/ PAULO MAZZONCINI DE AZEVEDO MARQUES

Linguagem: GRAILS, JAVA

Campo de Aplicação: IF-07, SD-03, SD-08

Tipo Programa: DS-01, GI-01, GI-05, SO-02

Expedido em: 01 de agosto de 2017

Aprovado por Liane Elizabeth Caldeira Lage

Liane Elizabeth Caldeira Lage

Diretor(a) Substituto(a) de Patentes, Programas de Computador e Topografia de Circuitos Integrados

APÊNDICE I – Modelo do formulário clínico openEHR para extração dos indicadores e estruturação da matriz de saúde.

Hilton Vicente César ▾

Meu Template

EHR Archetypes

- Cluster
- Composition
- Element
- Entry
- Section
- Structure

Demographic Model

PAINEL > MEU TEMPLATE

⚠ Matriz de Saúde: Template openEHR

ENCOUNTER

T 433 ▾

INPUTS

ID: ID

profissional_solicitacao: T

data_solicitacao: UI

cod_paciente: Q

idade: T

estado_civil: T

sexo: T

servico_solicitacao_internacao: T

municipio_origem_paciente: T

municipio_paciente: T

diagnostico_principal: T

traumatismo: T

prob_respiratorio: T

avc: T

convulsao: T

has: T

gravidez: T

diabetes: T

doenca_infecto: T

PROCESS

ID: ID

tempo_liberacao_vaga:

unidade_internacao: T

transferencia_intrahospitalar: T

duracao_internacao:

OUTPUTS

ID: ID

destino_paciente: T

diagnostico_primario: T

diagnostico_secundario: T

profissional_alta: T

APÊNDICE J – Código fonte parcial de classe que realiza comunicação com big data

```
def timeLine() {

    MongoClient mongoClient = new MongoClient('127.0.0.1',27017);
    DB db = mongoClient.getDB("MENTAL_HEALTH_OPENEHR"); // com.mongodb.DB;
    DBCollection coll = db.getCollection("SISAM_INTERNAcoes");
    mongoClient.setWriteConcern(WriteConcern.JOURNALED);

    DBObject dbo = new BasicDBObject();
    DBCursor cursor = coll.find()
    List<Document> document =cursor.toList()

}

class EhrBeliefDecisionNetworksController {

    def listRecords() {

        MongoClient mongoClient = new MongoClient('127.0.0.1',27017);
        DB db = mongoClient.getDB("template_openEHR"); // com.mongodb.DB;
        DBCollection coll = db.getCollection("EhrPaciente");
        mongoClient.setWriteConcern(WriteConcern.JOURNALED);
        BasicDBObject doc = new BasicDBObject();
        DBCursor cursor = coll.find();

        List<EhrBeliefDecisionNetwork> ehrBeliefDecisionNetwork =cursor.toList()
        [records:ehrBeliefDecisionNetwork.sort { it.id} ]

    }

}

Document documentInstance = Document.get(id)
if ( documentInstance == null) {
    flash.message = "Document not found."
    redirect (action:'list')
} else {
    response.setContentType("APPLICATION/OCTET-STREAM")
    response.setHeader("Content-Disposition", "Attachment;Filename=\"${documentInstance.filename}\"")

    def file = new File(documentInstance.fullPath)
    def fileInputStream = new FileInputStream(file)
    def outputStream = response.getOutputStream()

    byte[] buffer = new byte[4096];
    int len;
    while ((len = fileInputStream.read(buffer)) > 0) {
        outputStream.write(buffer, 0, len);
    }

    outputStream.flush()
    outputStream.close()
    fileInputStream.close()
}
```


APÊNDICE L–Trecho parcial da codificação openEHR

```

TemplateManager.getInstance().setArcheTypesDirectory("${grails
Application.config.easyehr.directory.archetypes}")
TemplateManager.getInstance().setTemplatesDirectory("${grailsA
pplication.config.easyehr.directory.templates}")

def templatesDirectory =
grailsApplication.config.easyehr.directory.templates
def dir = new File(templatesDirectory);
def templates = [];
def name = "Paciente"
def template = [:]

TEMPLATE ehrTemplate =
TemplateManager.getInstance().readTemplate(name)
template.name = name
template.type =
TemplateManager.getInstance().getType(ehrTemplate)

templates << template
return templates

public TEMPLATE readTemplate(String templateName) {

    try {

        TemplateReader reader = new TemplateReader();

        reader.setArchetypesDirectory(archetypesDirectory);
        reader.setTemplatesDirectory(templatesDirectory);
        reader.readTemplate(templateName);

        return reader.getTemplate();

    } catch (Exception e) {
        e.printStackTrace();
    }

    return null;
}

public ArchetypeHolder readADL(TEMPLATE template, String
archetypeId) {

    ArchetypeHolder archetypeHolder = new
ArchetypeHolder(template, archetypesDirectory + "\\\" +
archetypeId + ".adl");

```

```

        return archetypeHolder;
    }

    public String getType(TEMPLATE template) {
        Archetyped archetyped = template.getDefinition();

        if (archetyped != null) {
            Node node = archetyped.getDomNode();

            return
node.getAttributes().getNamedItem("xsi:type").getNodeValue();
        }

        return null;
    }

    public String getConceptName(TEMPLATE template) {
        Archetyped archetyped = template.getDefinition();

        if (archetyped != null) {

            Node conceptNameNode = null;
            Node node = archetyped.getDomNode();

            if ((conceptNameNode =
node.getAttributes().getNamedItem("concept_name")) != null) {
                return conceptNameNode.getNodeValue();
            }

        }

        return null;
    }

    public static TemplateManager getInstance() {

        if (instance == null) {
            instance = new TemplateManager();
        }

        return instance;
    }

```

```

<body>
<g:form controller="ehrPacientes" action="saveRecord">
<g:if test="${flash.message}">
<div class="message" role="status">${flash.message}</div>
</g:if>
<link rel="stylesheet" href="${resource(dir: 'css/ehr', file:
'style.css')}" type="text/css">
<g:if test="${ehrTemplateName}">

```



```

archetype (adl_version=1.4)
  openEHR-EHR-SECTION.mental_health_matrix.v1

concept
  [at0000]      --
language
  original_language = <[ISO_639-1::en]>
description
  original_author = <
    ["name"] = <" Hilton Vicente César;
                João Mazzoncini de Azevedo Marques;
                Newton Shydeo Brandao Miyoshi;
                Paulo Mazzoncini de Azevedo Marques ">
    >
  details = <
    ["en"] = <
      language = <[ISO_639-1::en]>
    >
  >
  lifecycle_state = <"0">
  other_contributors = <>
  other_details = <
    ["MD5-CAM-1.0.1"] = <"E23E327BD1C0397F8B30B1C45B2B2D54">
  >

definition
  SECTION[at0000] matches {
    --
    items cardinality matches {1..*; unordered} matches {
      SECTION[at0005] occurrences matches {0..1} matches {
        -- Section
        Temporal Dimension
          items cardinality matches {1..*; unordered} matches {
            allow_archetype SECTION[at0006] occurrences matches
            {0..*} matches {
              -- MENTAL HEALTH
              include
                archetype_id/value matches {././}
              }
            allow_archetype SECTION[at0007] occurrences matches
            {0..*} matches {
              -- MENTAL HEALTH
              include
                archetype_id/value matches {././}
              }
            allow_archetype SECTION[at0010] occurrences matches
            {0..*} matches {
              -- MENTAL HEALTH
              include
                archetype_id/value matches {././}
              }
            }
          }
        SECTION[at0008] occurrences matches {0..1} matches {*}
      }
    }
  }

ontology
  term_definitions = <
    ["en"] = <
      items = <
        ["at0005"] = <

```

```

        text = <"Section Temporal Dimension">
        description = <"*">
    >
    ["at0006"] = <
        text = <"MENTAL HEALTH">
        description = <"*">
    >
    ["at0007"] = <
        text = <"MENTAL HEALTH">
        description = <"*">
    >
    ["at0008"] = <
        text = <"Spatial (Geographical Dimension)">
        description = <"*">
    >
    ["at0010"] = <
        text = <"MENTAL HEALTH">
        description = <"*">
    >
>
>
>
>

```

Anexo A – Parecer do comitê de ética em pesquisa

**CENTRO DE SAÚDE ESCOLA
FACULDADE DE MEDICINA DE RIBEIRÃO PRETO
UNIVERSIDADE DE SÃO PAULO**

OF. 034/2016-CEP/CSE-FMRP-USP

Ribeirão Preto, 04 de Outubro de 2016.

Prezado Senhor,

Comunicamos que o projeto de pesquisa abaixo especificado foi analisado e **APROVADO** pelo Comitê de Ética em pesquisa do Centro de Saúde Escola da Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo, em reunião ordinária na data de 04 de Outubro de 2016.

Pesquisador: Newton Shydeo Brandão Miyoshi

CAAE: 59603416.7.0000.5414

Projeto de pesquisa: “Estudo e Análise de Métodos de Integração de Dados e Interoperabilidade em Saúde Mental no XIII Departamento Regional de Saúde do Estado de São Paulo ”

Em atendimento à Resolução 466/12, deverá ser encaminhado a este CEP o relatório final da pesquisa e a publicação de seus resultados, para acompanhamento, bem como comunicada qualquer intercorrência ou a sua interrupção.

Atenciosamente,

Prof. Dr. Laércio Joel Franco
Coordenador do CEP/CSE-FMRP-USP

Ilmo. Prof. Dr. João Mazzoncini de Azevedo Marques
Docente do Departamento de Medicina Social da Faculdade de Medicina de Ribeirão Preto/USP

RUA TEREZINA, Nº 690 – SUMAREZINHO – CEP 14055-380 – RIBEIRÃO PRETO – SP FONE: (16) 3315-0000/0001 FAX: (16) 3315-0029
EMAIL: CSECUIABA@FMRP.USP.BR