Arquitetura e Métodos de Integração de Dados e
Interoperabilidade Aplicados na Saúde Mental

Newton Shydeo Brandão Miyoshi

SERVIÇO DE PÓS-GRADUAÇÃO DA FMRP-USP

Data de Depósito:

Assinatura: _____

Arquitetura e Métodos de Integração de Dados e Interoperabilidade Aplicados na Saúde Mental

Newton Shydeo Brandão Miyoshi

***Orientador:* Prof. Dr. Paulo Mazzoncini de Azevedo Marques**

***Coorientador:* Prof. Dr. João Mazzoncini de Azevedo Marques**

Tese apresentada à Faculdade de Medicina de Ribeirão Preto – FMRP-USP, como parte dos requisitos para obtenção do título de Doutor em Ciências Médicas – Área de Concentração: Clínica Médica. *EXEMPLAR DE DEFESA.*

**USP – Ribeirão Preto
de 2018**

Ficha catalográfica

M634a

Miyoshi, Newton Shydeo Brandão

Arquitetura e métodos de integração de dados e interoperabilidade aplicados na saúde mental / Newton Shydeo Brandão Miyoshi; orientador Paulo Mazzoncinide Azevedo Marques; coorientador João Mazzoncini deAzevedo Marques. - Ribeirão Preto - SP, 2018.

126 p.

Tese (Doutorado - Programa de Pós-Graduação em Clínica Médica) - Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, 2018.

1. Sistemas de Informação em Saúde.
2. Interoperabilidade. 3. Saúde Mental. 4. Terminologias em Saúde. 5. Record Linkage. 6. Qualidade de Dados.
I. Marques, Paulo Mazzoncini de Azevedo, orient. II. Marques, João Mazzoncini de Azevedo, coorient. III. Título.

A minha amada família, pelo incentivo, carinho e apoio de sempre.

AGRADECIMENTOS

Agradeço a Deus por mais esta oportunidade de crescimento.

Às minhas famílias Miyoshi, Brandão, Tarossi e Pollettini pelo apoio e pela alegria com que sempre me recebem em seus corações.

Aos meus queridos pais, Newton e Suely, e minha querida irmã Stephanie, por todo incentivo, amor e sacrifícios que fazem para tornar meus sonhos realidades.

À minha amada esposa Juliana, pelo companheirismo, amor, apoio e incentivo de todas as horas. Também por contribuir de modo essencial com idéias, revisando e auxiliando em todo o processo deste trabalho.

Ao meu orientador Paulo, ao meu coorientador João e ao professor Domingos, pela paciência, pelo incentivo constante, pela oportunidade de trabalho e por me guiarem durante todos esses anos.

Aos amigos do CIIS, Tiago, Lariza, Nati Chiari, André e Nath Crepaldi, que compartilharam os momentos de dificuldades e alegrias no dia a dia do nosso trabalho.

Ao Programa da Clínica Médica em especial ao Émerson, pela atenção e pela presteza, sempre auxiliando aos alunos com muita boa vontade.

*“As invenções são, sobretudo,
o resultado de um trabalho de teimoso.”
(Santos Dumont)*

RESUMO

MIYOSHI, N. S. B.. **Arquitetura e Métodos de Integração de Dados e Interoperabilidade Aplicados na Saúde Mental**. 2018. 126 f. Tese (Doutorado em Ciências Médicas – Área de Concentração: Clínica Médica) – Faculdade de Medicina de Ribeirão Preto (FMRP/USP), Ribeirão Preto – SP.

Contextualização: A disponibilidade e integração das informações em saúde relativas a um mesmo paciente entre diferentes níveis de atenção ou entre diferentes instituições de saúde é normalmente incompleta ou inexistente. Isso acontece principalmente porque os sistemas de informação que oferecem apoio aos profissionais da saúde não são interoperáveis, dificultando também a gestão dos serviços a nível municipal e regional. Essa fragmentação da informação também é desafiadora e preocupante na área da saúde mental, em que normalmente se exige um cuidado prolongado e que integra diferentes tipos de serviços de saúde. Problemas como a baixa qualidade e indisponibilidade de informações, assim como a duplicidade de registros, são importantes aspectos na gestão e no cuidado prolongado ao paciente portador de transtornos mentais. Apesar disso, ainda não existem estudos objetivos demonstrando o impacto efetivo da interoperabilidade e integração de dados na gestão e na qualidade de dados para a área de saúde mental. **Objetivos:** Neste contexto, o projeto tem como objetivo geral propor uma arquitetura de interoperabilidade para a assistência em saúde regionalizada e avaliar a efetividade de técnicas de integração de dados e interoperabilidade para a gestão dos atendimentos e internações em saúde mental na região de Ribeirão Preto, assim como o impacto na melhoria e disponibilidade dos dados por meio de métricas bem definidas. **Métodos:** O framework de interoperabilidade proposto tem como base a arquitetura cliente-servidor em camadas. O modelo de informação de interoperabilidade foi baseado em padrões de saúde internacionais e nacionais. Foi proposto um servidor de terminologias baseado em padrões de informação em saúde. Foram também utilizados algoritmos de Record Linkage para garantir a identificação unívoca do paciente. Para teste e validação da proposta foram utilizados dados de diferentes níveis de atenção à saúde provenientes de atendimentos na rede de atenção psicossocial na região de Ribeirão Preto. Os dados foram extraídos de cinco fontes diferentes: (i) a Unidade Básica de Saúde da Família - I, de Santa Cruz da Esperança; (ii) o Centro de Atenção Integrada à Saúde, de Santa Rita do Passa Quatro; (iii) o Hospital Santa Tereza; (iv) as informações de solicitações de internação contidas no SISAM (Sistema de Informação em Saúde Mental); e (v) dados demográficos do Barramento do Cartão Nacional de Saúde do Ministério da Saúde. As métricas de qualidade de dados utilizadas foram completude, consistência, duplicidade e acurácia. **Resultados:** Como resultado deste trabalho, foi projetado, desenvolvido e testado a plataforma de interoperabilidade em saúde, denominado eHealth-Interop. 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. Foi desenvolvido também um servidor de terminologias, denominado eHealth-

Interop Terminology Server, que pode ser utilizado como um componente independente e em outros contextos médicos. No total foram obtidos dados de 31340 registros de pacientes pelo SISAM, e-SUS AB de Santa Cruz da Esperança, do CAIS de Santa Rita do Passa Quatro, do Hospital Santa Tereza e do Barramento do CNS do Ministério da Saúde. Desse total, 30,47% (9548) registros foram identificados como presente em mais de 1 fonte de informação, possuindo diferentes níveis de acurácia e completude. A análise de qualidade de dados, abrangendo todas os registros integrados, obteve uma melhoria na completude média de 18,40% (de 56,47% para 74,87%) e na acurácia sintática média de 1,08% (de 96,69% para 96,77%). Na análise de consistência houve melhoras em todas as fontes de informação, variando de uma melhoria mínima de 14,4% até o máximo de 51,5%. Com o módulo de *Record Linkage* foi possível quantificar, 1066 duplicidades e, dessas, 226 foram verificadas manualmente. **Conclusões:** A disponibilidade e a qualidade da informação são aspectos importantes para a continuidade do atendimento e gerenciamento de serviços de saúde. A solução proposta neste trabalho visa estabelecer um modelo computacional para preencher essa lacuna. O ambiente de interoperabilidade foi capaz de integrar a informação no caso de uso de saúde mental com o suporte de terminologias clínicas internacionais e nacionais sendo flexível para ser estendido a outros domínios de atenção à saúde.

Palavras-chave: Sistemas de Informação em Saúde, Interoperabilidade, Saúde Mental, Terminologias em Saúde, Record Linkage, Qualidade de Dados.

ABSTRACT

MIYOSHI, N. S. B.. **Investigation of the Effectiveness of Data Integration and Interoperability Methods Applied to Mental Health**. 2018. 126 f. Tese (Doutorado em Ciências Médicas – Área de Concentração: Clínica Médica) – Faculdade de Medicina de Ribeirão Preto (FMRP/USP), Ribeirão Preto – SP.

Contextualization: The availability and integration of health information from the same patient between different care levels or between different health services is usually incomplete or non-existent. This happens especially because the information systems that support health professionals are not interoperable, making it difficult to manage services at the municipal and regional level. This fragmentation of information is also challenging and worrying in the area of mental health, where long-term care is often required and integrates different types of health services and professionals. Problems such as poor quality and unavailability of information, as well as duplicate records, are important aspects in the management and long-term care of patients with mental disorders. Despite this, there are still no objective studies that demonstrate the effective impact of interoperability and data integration on the management and quality of data for the mental health area. **Objectives:** In this context, this project proposes an interoperability architecture for regionalized health care management. It also proposes to evaluate the effectiveness of data integration and interoperability techniques for the management of mental health hospitalizations in the Ribeirão Preto region as well as the improvement in data availability through well-defined metrics. **Methods:** The proposed framework is based on client-service architecture to be deployed in the web. The interoperability information model was based on international and national health standards. It was proposed a terminology server based on health information standards. Record Linkage algorithms were implemented to guarantee the patient identification. In order to test and validate the proposal, we used data from different health care levels provided by the mental health care network in the Ribeirão Preto region. The data were extracted from five different sources: the Family Health Unit I of Santa Cruz da Esperança, the Center for Integrated Health Care of Santa Rita do Passa Quatro, Santa Tereza Hospital, the information on hospitalization requests system in SISAM (Mental Health Information System) and demographic data of the Brazilian Ministry of Health Bus. **Results:** As a result of this work, the health interoperability platform, called eHealth-Interop, was designed, developed and tested. A proposal was adopted for interoperability through web services with a data integration model based on a centralizing database. A terminology server, called eHealth-Interop Terminology Server, has been developed that can be used as an independent component and in other medical contexts. In total, 31340 patient records were obtained from SISAM, eSUS-AB from Santa Cruz da Esperança, from CAIS from Santa Rita do Passa Quatro, from Santa Tereza Hospital and from the CNS Service Bus from the Brazilian Ministry of Health. 47% (9548) records were identified as present in more than 1 information source, having different levels of

accuracy and completeness. The data quality analysis, covering all integrated records, obtained an improvement in the average completeness of 18.40% (from 56.47% to 74.87%) and the mean syntactic accuracy of 1.08% (from 96.69% to 96.77%). In the consistency analysis there were improvements in all information sources, ranging from a minimum improvement of 14.4% to a maximum of 51.5%. With the Record Linkage module it was possible to quantify 1066 duplications, of which 226 were manually verified. **Conclusions:** The information's availability and quality are both important aspects for the continuity of care and health services management. The solution proposed in this work aims to establish a computational model to fill this gap. It has been successfully applied in the mental health care context and is flexible to be extendable to other medical domains.

Key-words: Health Information Systems, Interoperability, Mental Health, Data Integration, Health Terminologies, Record Linkage, Data Quality.

LISTA DE ILUSTRAÇÕES

Figura 1 – Processo de definição do modelo de informação de interoperabilidade	29
Figura 2 – Principais fases de mudança no conjunto dos dados	32
Figura 3 – Arquitetura conceitual do eHealth-Interop	34
Figura 4 – Arquitetura do eHealth-Interop Terminology Server	36
Figura 5 – Diagrama com as principais classes do eHI Terminology Server	37
Figura 6 – Principais atores e grupos de operações relacionados ao Servidor de Terminologia	38
Figura 7 – Fluxo de identificação do paciente com o módulo de Record Linkage	40
Figura 8 – Interface de desenvolvedor do eHealth-Interop	41
Figura 9 – Interface de administração do eHealth-Interop	41
Figura 10 – Distribuição da quantidade de pacientes por número de serviços de saúde diferentes envolvidos no cuidado	42
Figura 11 – Arquitetura do adaptador do SISAM	43
Figura 12 – Arquitetura do adaptador do e-SUS AB	45
Figura 13 – Arquitetura do adaptador para CAIS e HST	45
Figura 14 – Resultado da análise de completude dos registros integrados, calculado para cada conceito demográfico	49
Figura 15 – Resultado da análise de acurácia dos registros integrados, calculado para cada conceito demográfico	50
Figura 16 – Processo de record linkage para identificações de duplicidades	51
Figura 17 – Folha de aprovação do comitê de ética em pesquisa	126

LISTA DE TABELAS

Tabela 1 – Características gerais das fontes de informação utilizadas	29
Tabela 2 – Operações disponíveis pela API do servidor de terminologias	38
Tabela 3 – Quantidade de registros obtidos para cada uma das fontes de dados	46
Tabela 4 – Distribuição da quantidade de registros de paciente integráveis por fonte de dado	46
Tabela 5 – Análise de completude e acurácia no conjunto total de dados integrados . .	47
Tabela 6 – Análise de completude por grupo de atributos	48
Tabela 7 – Análise de acurácia por grupo de atributos	48
Tabela 8 – Análise de consistência por fonte de dados	48
Tabela 9 – Análise de completude dos dados originais	68
Tabela 10 – Melhoria da completude após a integração	69
Tabela 11 – Completude nos registros integráveis	70
Tabela 12 – Análise de acurácia	72
Tabela 13 – Melhoria da acuracia	73
Tabela 14 – Acurácia nos registros integráveis	74
Tabela 15 – Análise de consistência	75
Tabela 16 – Melhoria na consistência	76

LISTA DE ABREVIATURAS E SIGLAS

CTS-2		<i>Common Terminology Services 2</i>
FHIR		<i>Fast Healthcare Interoperability Resources</i>
HIMSS	...	<i>Healthcare Information and Management Systems Society</i>
HITECH	.	<i>Health Information Technology for Economic and Clinical Health</i>
HL7		<i>Health Level 7</i>
IHE		<i>Integrating the Healthcare Enterprise</i>
OMG		<i>Object Management Group</i>
OMS		Organização Mundial de Saúde
RAPS		Rede de Atenção Psicossocial
RAS		Redes de Atenção à Saúde
RL		<i>Record Linkage</i>
SISAM	...	Sistema de Informação em Saúde Mental
SVS		<i>Shared Value Sets</i>
UMLS	...	<i>Unified Medical Language System</i>

SUMÁRIO

1	INTRODUÇÃO	16
1.1	Contextualização	16
1.2	Organização do Documento	18
2	SAÚDE MENTAL E A CONTINUIDADE DO CUIDADO	19
2.1	A Continuidade do Cuidado	19
2.2	A Regionalização da Assistência à Saúde	20
2.3	A Assistência à Saúde Mental no Brasil	20
3	INTEROPERABILIDADE DE DADOS EM SAÚDE	21
3.1	Interoperabilidade	21
3.2	Padrões de Informação em Saúde	22
3.2.1	<i>Common Terminology Services 2</i>	23
3.2.2	<i>Integrating the Healthcare Enterprise</i>	23
3.2.3	<i>Fast Healthcare Interoperability Resources</i>	23
3.3	Record Linkage	24
3.4	Terminologias	25
4	OBJETIVOS	26
5	MATERIAIS E MÉTODOS	27
5.1	Rede de Atenção à Saúde Mental (RAPS) do Departamento Regional de Saúde XIII (DRS-XIII) do Estado de São Paulo	27
5.2	Fontes de Dados	28
5.3	Modelo de Informação de Interoperabilidade	29
5.4	Análise da Necessidade de Integração	30
5.5	Análise de Qualidade de Dados	31
6	RESULTADOS	33
6.1	Arquitetura de Software	33
6.2	Servidor de Terminologias	35
6.2.1	<i>Arquitetura do Servidor de Terminologias</i>	35
6.2.2	<i>Operações definidas pelo WebServices</i>	37
6.2.3	<i>Terminologias Utilizadas</i>	38

6.3	Módulo de Record Linkage	39
6.4	Módulo de Segurança	39
6.5	API de Comunicação	40
6.6	Caso de Uso: Interoperabilidade na Rede de Atenção Psicossocial .	42
6.6.1	<i>Análise da Necessidade de Integração</i>	42
6.6.2	<i>Arquitetura dos Adaptadores</i>	43
6.6.2.1	<i>Adaptador SISAM</i>	43
6.6.2.2	<i>Adaptador e-SUS AB</i>	44
6.6.2.3	<i>Adaptador CAIS-SR e HST</i>	44
6.6.3	<i>Análise de Qualidade de Dados</i>	46
6.6.4	<i>Caso de Exemplo</i>	51
6.7	Resultados Complementares	52
6.7.1	<i>Mapeamento com o SNOMED-CT</i>	52
6.7.2	<i>Incorporação do Conjunto Mínimo de Dados à Saúde</i>	53
6.8	Publicações	53
6.9	Orientações Associadas à Tese	54
7	DISCUSSÃO	55
7.1	Trabalhos Relacionados	55
7.2	Discussão Geral	57
7.2.1	<i>Desenvolvimento da Plataforma eHealth-Interop</i>	57
7.2.2	<i>Definição do Modelo de Informação</i>	58
7.2.3	<i>Definição do Servidor de Terminologias</i>	59
7.2.4	<i>Análise de Qualidade</i>	59
7.2.5	<i>Limitações e Dificuldades</i>	60
8	CONCLUSÕES	61
	REFERÊNCIAS	62
APÊNDICE A	TABELAS COMPLETAS COM ANÁLISES DE QUALIDADE	68
A.1	Análise de Completude	68
A.2	Análise de Acurácia Sintática	71
A.3	Análise de Consistência	75
APÊNDICE B	PÔSTER PARA WHO-FIC NETWORK MEETING 2017	78
APÊNDICE C	ARTIGO PARA O HCIST 2017	80

APÊNDICE D	ARTIGO PARA O JOURNAL OF MEDICAL INTERNET RESEARCH	85
APÊNDICE E	MODELOS DE INFORMAÇÃO PARA INTEROPERABILIDADE	107
APÊNDICE F	MAPEAMENTO DO FORMULÁRIO DEMOGRÁFICO PARA SNOMED-CT	121
ANEXO A	APROVAÇÃO DO COMITÊ DE ÉTICA	126

INTRODUÇÃO

1.1 Contextualização

A Organização Mundial de Saúde (OMS) define saúde mental como uma ampla gama de atividades direta ou indiretamente relacionadas com o componente mental do bem-estar que está incluso na definição de saúde como um todo: "um estado de completo bem-estar físico, mental e social e não meramente a ausência de doença" ([World Health Organization, 2014](#)). Ela está relacionada com a promoção do bem-estar, a prevenção de transtornos mentais, e também com o tratamento e reabilitação de pessoas afetadas por transtornos mentais.

No Brasil, o cuidado público à saúde mental no SUS é realizado por meio da Rede de Atenção Psicossocial (RAPS) instituída na portaria nº 3.088, de 23 de dezembro de 2011 para o cuidado de pessoas com transtornos mentais e com necessidades decorrentes do uso de crack, álcool e outras drogas ([Ministerio da Saude, 2011](#)). Um dos objetivos gerais da RAPS é garantir a articulação e integração dos pontos de atenção das redes de saúde no território, qualificando o cuidado por meio do acolhimento, do acompanhamento contínuo e da atenção às urgências visando assim a integralidade da assistência. Essa necessidade de integração e coordenação dos serviços e dispositivos de apoio fica mais evidente com a definição das Redes de Atenção à Saúde, sendo um dos pontos-chaves a disponibilidade de recursos e informações para toda a rede ([MENDES, 2011](#)).

Na região de Ribeirão Preto, os serviços de saúde que ofereciam apoio na área de saúde mental trabalharam de forma adequada até 2004. A partir desse período a demanda em saúde mental aumentou progressivamente e a coordenação das informações nos diferentes serviços tornou-se problemática. 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. Em 2011, com o estabelecimento da portaria que instituiu a RAPS, os serviços se reorganizaram e foi

iniciado o desenvolvimento do Sistema de Informação em Saúde Mental (SISAM)([YOSHIURA et al., 2017](#)). O principal objetivo do SISAM foi o gerenciamento dos pedidos de internação e dos encaminhamentos após a alta (referência e contra-referência). Embora o SISAM tenha otimizado a coordenação dos serviços e dos processos e, por consequência, organizado a RAPS, ainda assim os sistemas de informação que oferecem apoio nos diferentes níveis de atenção não são integrados.

Uma extensa área de pesquisa das ciências da computação que busca auxiliar no problema de desfragmentação dos sistemas e das informações na área da saúde é a interoperabilidade e integração de dados. A interoperabilidade é definida como a habilidade de diferentes sistemas de informação e aplicações se comunicarem, trocar dados e utilizar de modo transparente esses dados e informações associadas ([HIMSS, 2010](#)).

Essa troca de informações em saúde relativas a um mesmo paciente entre diferentes níveis de atenção e diferentes instituições de saúde é normalmente inexistente ou incompleta. Com isso, a continuidade do cuidado e a comunicação entre os diferentes serviços envolvidos na assistência pode ser prejudicada. Isso se deve, em grande parte, ao fato de os sistemas de informação em saúde que oferecem apoio aos médicos e aos profissionais da saúde não estarem integrados ([BODENHEIMER, 2008](#)).

Coffey et al. ([COFFEY et al., 2008](#)) destacam que a fragmentação do serviços prejudica a atenção à saúde mental criando uma barreira, muitas vezes intransponível, entre diferentes prestadores e o paciente, dificultando o fornecimento de um tratamento efetivo. Segundo Spiranovic et al. ([SPIRANOVIC et al., 2016](#)) a baixa qualidade de dados, inconsistências/discrepâncias das informações e falta de interoperabilidade entre os sistemas de informação em saúde são uma das principais dificuldades para o uso secundário dos dados de registro eletrônicos em saúde mental, incluindo a gestão.

Embora existam diversos estudos qualitativos que evidenciem os benefícios da integração e interoperabilidade em dados de saúde, assim como um senso comum dessa necessidade, ainda faltam estudos mais objetivos sobre quais aspectos clínicos e/ou de gestão são melhor beneficiados por essas técnicas ([HERSH et al., 2015](#))([HINCAPIE; WARHOLAK, 2011](#))([VEST; JASPERSON, 2010](#)). Na área de saúde mental, em específico, não foi encontrado nenhum estudo que realizasse esse tipo de análise.

Neste contexto, o trabalho de pesquisa aqui apresentado tem por finalidade buscar diminuir a lacuna existente na área de interoperabilidade e integração de dados na área de atenção à saúde mental.

1.2 Organização do Documento

A presente tese está organizada da seguinte forma: no [Capítulo 2](#) são apresentados fundamentos teóricos sobre saúde mental e continuidade do cuidado; no [Capítulo 3](#) são discutidos fundamentos computacionais conceituais e tecnológicos sobre interoperabilidade e integração de dados; no [Capítulo 4](#) são listados os objetivos do presente trabalho. No [Capítulo 5](#) a metodologia do trabalho é detalhada; no [Capítulo 6](#) são descritos os resultados obtidos; no [Capítulo 7](#) é apresentada a discussão sobre o trabalho desenvolvido e os resultados obtidos em comparação com outros trabalhos relacionados; e, finalmente, no [Capítulo 8](#) são apresentadas as conclusões do trabalho desenvolvido.

SAÚDE MENTAL E A CONTINUIDADE DO CUIDADO

Neste capítulo são abordados os principais fundamentos teóricos do contexto da área de saúde em que este trabalho está inserido. Os seguintes pontos são detalhados: o problema da continuidade do cuidado, a regionalização da assistência à saúde, em específico a organização dos serviços de saúde no Brasil e, enfim, a assistência à saúde mental.

2.1 A Continuidade do Cuidado

A continuidade do cuidado pode ser definida como a situação em que uma série de diferentes eventos de cuidado à saúde são experienciados de modo coerente, consistente e contínuo com as necessidades e o contexto pessoal do paciente ([HAGGERTY et al., 2003](#)). Em algumas áreas da saúde essa problemática tem maior destaque, como o caso da atenção primária, a saúde mental, doenças crônicas e doenças complexas, pois exigem um acompanhamento longitudinal e multiprofissional.

Haggerty et al. ([HAGGERTY et al., 2003](#)) divide a continuidade no cuidado em três principais partes:

- *Continuidade informacional*: definida como uso da informação como meio de ligação entre diferentes eventos de cuidado ao paciente e entre os diferentes serviços envolvidos. As informações, que incluem as ações de cuidado realizadas, medicamentos, exames, assim como informações pessoais, permitem o cuidado apropriado e individualizado para cada paciente;

- *Continuidade gerencial*: definida como a coordenação entre diferentes serviços de saúde de modo complementar e no tempo correto. Esse gerenciamento inclui planos de cuidado compartilhado e protocolos bem definidos para diferentes situações. Esse aspecto da continuidade do cuidado é particularmente importante no caso de doenças crônicas exigindo tanto consistência quanto flexibilidade.
- *Continuidade relacional*: é o relacionamento terapêutico contínuo entre o paciente e os diferentes profissionais de saúde responsáveis por seu cuidado. Esse é um ponto importante na atenção primária e na saúde mental pois fornece um senso de coerência e previsibilidade nas ações realizadas.

2.2 A Regionalização da Assistência à Saúde

A regionalização da assistência à saúde é uma das diretrizes bases do Sistema Único de Saúde, presente na Constituição Federal do Brasil (CARVALHO *et al.*, 2017). Apesar dessa resolução, houve um atraso na regionalização do SUS, ocasionando, desde 1990, a criação de sistemas locais isolados e independentes, fragmentando a integralidade do cuidado (CARVALHO *et al.*, 2017). Tendo em vista a diversidade geográfica, sócio-econômica, política e técnico-administrativa do cenário brasileiro e, buscando superar a fragmentação dos serviços e ações em saúde, foram definidas as diretrizes das Redes de Atenção à Saúde (RAS) no SUS em 2000. A RAS é definida como uma organização de diversos serviços de saúde objetivando a assistência integral, contínua, humana e qualificada a determinada população (MENDES, 2011).

2.3 A Assistência à Saúde Mental no Brasil

A portaria nº 3.088, de 23 de dezembro de 2011 instituiu no SUS a Rede de Atenção Psicossocial (RAPS) para o cuidado de pessoas com transtornos mentais e com necessidades decorrentes do uso de crack, álcool e outras drogas (Ministerio da Saude, 2010). Conforme a legislação atual da RAPS, um dos aspectos importantes é garantir a articulação e integração dos pontos de atenção das redes de saúde no território, qualificando o cuidado por meio do acolhimento, do acompanhamento contínuo e da atenção às urgências.

INTEROPERABILIDADE DE DADOS EM SAÚDE

Neste capítulo são apresentadas as bases conceituais e tecnológicas que suportam a pesquisa desenvolvida. Primeiramente é discutido o problema da interoperabilidade de dados entre diferentes sistemas de informação. Em seguida são abordados os principais aspectos da interoperabilidade envolvidos na tese como: os padrões de informação em saúde, as terminologias clínicas e as técnicas de *record linkage* para identificação unívoca do paciente.

3.1 Interoperabilidade

A *Healthcare Information and Management Systems Society* (HIMSS) define interoperabilidade como “a capacidade de diferentes sistemas de tecnologia da informação e aplicações de software de comunicar, trocar dados e usar a informação que foi trocada. O esquema e os padrões de troca de informação em saúde devem permitir que os dados sejam compartilhados entre o profissional de saúde, o laboratório, o hospital, a farmácia e o paciente independentemente do fornecedor da aplicação ou do aplicativo.” (HIMSS, 2010).

A interoperabilidade pode ser categorizada em três níveis:

1. *Interoperabilidade Fundamental*: existe a troca de dados entre um sistema de informação e outro porém este não necessariamente é capaz de interpretar o dado recebido. Neste caso o dado é apenas recebido e não recebe nenhum tratamento, por exemplo, na troca de um resultado de exame em pdf ou de um documento digitalizado.

2. *Interoperabilidade Estrutural*: neste nível intermediário os dados são enviados de um sistema para outro em um formato estruturado com uma sintaxe bem definida. Neste caso o objetivo operacional ou clínico da troca de dados é preservado, garantindo que o dado possa ser diretamente lido e quantificado, porém não existe garantia do entendimento da informação associada.
3. *Interoperabilidade Semântica*: este é o nível mais alto de interoperabilidade. Neste nível dois ou mais sistemas são capazes de trocar dados, interpretá-los e utilizar a informação trocada ([Institute of Electrical and Electronics Engineers, 1990](#)). Neste nível, além de os dados trocados seguirem uma sintaxe bem definida, o conteúdo também é codificado utilizando um vocabulário controlado, terminologia ou ontologia.

A partir de 2009, diversos trabalhos relacionados utilizam o termo denominado HIE (*Health Information Exchange*) como sinônimo de interoperabilidade de dados em saúde. Este termo surgiu, principalmente no contexto americano, a partir da lei *Health Information Technology for Economic and Clinical Health* (HITECH) ([AKHLAQ; SHEIKH; PAGLIARI, 2016](#)).

Akhlaq et al. define formalmente HIE como sendo “a mobilização eletrônica de informações clínicas e administrativas dentro ou entre organizações localmente, regionalmente ou até internacionalmente entre vários sistemas de informação, baseado em padrões de informação reconhecidos localmente e/ou nacionalmente, mantendo a autenticidade e precisão da informação sendo trocada, permitindo às instituições envolvidas tomarem decisões para melhorar a qualidade da assistência à saúde de um paciente e da população como um todo” ([AKHLAQ; SHEIKH; PAGLIARI, 2016](#)).

3.2 Padrões de Informação em Saúde

A portaria no 2.073 de agosto de 2011 ([Ministério da Saúde, 2011](#)) regulamenta a utilização dos padrões de informação em saúde e interoperabilidade. Como principais objetivos da portaria, destaca-se: (i) a utilização de ontologias, terminologias e codificações para representação de conceitos na área da saúde visando o compartilhamento das informações em saúde; (ii) permitir a interoperabilidade funcional, sintática e semântica entre sistemas de informação em saúde; (iii) identificação unívoca do usuário nos diferentes sistemas de informação em saúde.

A portaria também orienta quanto aos padrões de informação a serem utilizados: OpenEHR, para definição do Registro Eletrônico em Saúde (RES); HL7 como padrão de interoperabilidade com foco na troca de resultados e solicitação de exames; SNOMED-CT, para codificação de termos clínicos e mapeamento de terminologias; TISS, para troca de informação em saúde suplementar; HL7 CDA, para definição da arquitetura do documento clínico; DICOM, para troca de exames de imagem; LOINC, para codificação de exames laboratoriais; ISBT 128, para identificação das etiquetas de biomateriais, IHE-PIX (Patient Identifier Cross-Referencing) para

cruzamento de pacientes de diferentes sistemas de informação e, por fim, ISO 13606-2 para interoperabilidade de modelos de conhecimento, incluindo arquétipos, templates e metodologia de gestão.

A seguir são detalhados os padrões de informação em saúde que serviram de base para realização deste trabalho.

3.2.1 Common Terminology Services 2

O *Object Management Group* (OMG), fundado em 1989, é um consórcio sem fins lucrativos, formado pela indústria, academia, agências governamentais e usuários finais, que visa desenvolver e manter várias especificações técnicas para uma vasta gama de aplicações (OMG, 2017).

O *Common Terminology Services 2* (CTS-2) é uma dessas especificações mantidas pela OMG (OMG, 2015). CTS-2 define a estruturação de um serviço de padronização da informação, a partir de um modelo computacional para gerenciar terminologias estruturadas de modo independente de plataforma. O objetivo do padrão é permitir uma implementação modular e também possibilitar que somente a parte das especificações relevantes para as necessidades locais seja implantada.

3.2.2 Integrating the Healthcare Enterprise

Integrating the Healthcare Enterprise (IHE) é uma iniciativa conjunta formada inicialmente pela HIMSS (*Healthcare Information System Management System*) e RSNA (*Radiological Society of North America*), em 1988 (VEGODA, 2002). O IHE tem como objetivo coordenar o uso de padrões de informação de saúde, terminologias e protocolos para permitir uma interoperabilidade “*plug-and-play*” entre sistemas de diferentes fornecedores. Os grupos de trabalho do IHE são organizados em diferentes domínios clínicos e operacionais. Cada domínio produz documentos denominados *Technical Frameworks* (IHE, 2016b).

Cada *Technical Framework* é composto por vários Perfis IHE. Um perfil IHE fornece as definições de como os padrões podem ser implementados para atingir um objetivo específico em determinado domínio. O perfil *Shared Value Sets* (SVS) estabelece os principais atores e transações envolvidas no gerenciamento centralizado dos termos e nomenclaturas que são utilizados por sistemas de informação em saúde, servindo de base para a interoperabilidade semântica (IHE, 2016a).

3.2.3 Fast Healthcare Interoperability Resources

Fast Healthcare Interoperability Resources (FHIR) é um padrão *Health Level 7* (HL7), é uma evolução das versões 2 e 3 mas que busca permitir a construção de interoperabilidade eficaz entre aplicações com o uso de padrões de desenvolvimento web (BENDER; SARTIPI, 2013).

Um dos componentes do FHIR é o Módulo de Terminologias que fornece recursos e operações para permitir o gerenciamento de codificações locais e externas assim como subconjuntos dessas codificações. A especificação terminológica é baseada em dois conceitos principais: (i) *CodeSystem* que representa uma terminologia específica, classificação ou ontologia e seus conceitos e termos correspondentes; (ii) *ValueSet* é um grupo de códigos definidos em um ou mais *CodeSystems* para ser usado em um contexto particular (HL7, 2017b). Os serviços de terminologia também compreendem operações e recursos para representar o mapeamento entre vários sistemas de código ou conjuntos de valores (HL7, 2017a).

3.3 Record Linkage

Record Linkage (RL) pode ser definido como o processo de identificação de registros de diferentes fontes que correspondem à mesma entidade do mundo real (GU *et al.*, 2003). O processo de *record linkage* também é conhecido como resolução de entidade, deduplicação, problema de *merge/purge*, reconciliação de dados, ou, no contexto clínico, como *matching* de pacientes. Essas técnicas são essenciais para garantir a integridade da informação e a qualidade dos dados durante a interoperabilidade e a integração de dados.

Os algoritmos de RL podem ser classificados em: determinísticos e probabilísticos. Algoritmos determinísticos são construídos por meio de um conjunto fixo de regras que definem uma identificação unívoca e inequívoca de um registro (ou paciente). Essas regras normalmente são aplicadas aos atributos da entidade. Os algoritmos probabilísticos definem uma probabilidade de diferentes registros se referirem à mesma entidade. Para o cálculo dessa probabilidade existem diferentes abordagens, normalmente a partir da comparação dos atributos dessas unidades e de pesos a estes associados. Por exemplo, considerando os atributos sexo, nome e nome da mãe, se apenas nome e sexo forem similares, a probabilidade que esses registros sejam referentes a uma mesma pessoa será muito menor que no caso de coincidência também do nome e nome da mãe entre dois registros.

Os sistemas de informação em saúde geralmente implementam o método determinístico de RL baseado em identificadores únicos do paciente como CPF, RG, CNS. No contexto da saúde mental e em outras áreas, tais como emergências médicas, muitas vezes não é possível obter uma identificação única ou informações completas de pacientes. Além disso, a duplicidade pode ocorrer devido a entrada incorreta de dados por usuários que operam sistemas de informação de saúde. Os erros de usuário e a falta de informação podem ser propagados em um cenário de troca de informações de saúde em que vários sistemas estão envolvidos.

3.4 Terminologias

A aplicação de tecnologias de informação na área da saúde cada vez mais se mostra uma ferramenta essencial, tanto para resolução rápida e coordenação de diversos processos, como para análises e tomada de decisão no cuidado individual e coletivo. Um aspecto central neste contexto é a padronização de termos e conceitos e isso se deve a três fatores principais: (i) a natureza altamente heterogênea e por vezes ambígua do linguajar médico e sua constante evolução; (ii) uma quantidade enorme de dados que é gerada constantemente pela automatização de processos e a partir de novas tecnologias que surgem e (iii) necessidade de processar, analisar e tomar decisões com base nessas informações ([KONOPKA, 2015](#)).

Ao longo dos anos, diversas soluções foram desenvolvidas para formalizar e representar de modo padronizado o conhecimento médico, tais como: terminologias, codificações, vocabulários controlados, taxonomias, thesaurus e ontologias ([BODENREIDER, 2008](#)), assim como diferentes formatos para representação computacional desses mecanismos. Neste contexto, o servidor de terminologias e ontologias é a peça de software responsável pelo armazenamento e gerenciamento desses artefatos linguísticos computacionais. O servidor de terminologias e ontologias atua como um repositório terminológico e pode agregar outras funções e ferramentas auxiliares para busca de termos e conceitos.

OBJETIVOS

Considerando o caráter descentralizado do SUS, que torna os processos e instrumentos informatizados utilizados na assistência fragmentados, o presente trabalho estuda a hipótese de que a aplicação de técnicas de integração de dados e interoperabilidade pode auxiliar na melhoria dos indicadores de qualidade da informação e, conseqüentemente, nos indicadores de gestão em saúde.

Neste contexto, o projeto tem como objetivo geral propor uma arquitetura de interoperabilidade para a assistência em saúde regionalizada e avaliar a efetividade de técnicas de integração de dados e interoperabilidade para a gestão dos atendimentos e internações em saúde mental na região de Ribeirão Preto assim como o impacto na melhoria e disponibilidade dos dados.

Como objetivos específicos pretende-se:

- (i) Desenvolver uma plataforma computacional para interoperabilidade e integração de dados em saúde mental;
- (ii) Definir quais informações a serem interoperadas entre as diferentes instituições e níveis de atenção no contexto da saúde mental;
- (iii) Definir mecanismos para lidar com o problema da diferença terminológica em saúde visando a interoperabilidade e compartilhamento da informação;
- (iv) Avaliar como a integração de dados altera o grau de qualidade das informações cadastradas nos sistemas de informação em saúde;

MATERIAIS E MÉTODOS

Neste capítulo são explanados os métodos e ferramentas utilizados para elaboração do trabalho. Inicialmente é explanado o contexto regional e de saúde mental em que a tese está inserida. As fontes de dados utilizadas na elaboração do projeto são detalhas. Em seguida, é descrito como o modelo de informação para interoperabilidade foi definido. Posteriormente são apresentadas as medidas de qualidade de dados e indicadores utilizados para análise.

Essa pesquisa envolve o levantamento de dados secundários e foi aprovada pelo Comitê de Ética em Pesquisa do Centro de Saúde Escola da Faculdade de Medicina de Ribeirão Preto (CSE/FMRP/USP) (CAAE: 59603416.7.0000.5414) ([Apêndice A](#)), obtendo a dispensa do Termo de Consentimento Livre e Esclarecido (TCLE). Foi obtido o Termo de Consentimento de Uso de Dados (TCUD) para cada uma das instituições envolvidas. O projeto também foi apresentado e aprovado pela Comissão Intergestores Regional de Abrangência Macrorregional do Departamento Regional de Saúde XIII.

5.1 Rede de Atenção à Saúde Mental (RAPS) do Departamento Regional de Saúde XIII (DRS-XIII) do Estado de São Paulo

No Brasil, a atenção pública em saúde mental deve ser realizada nas redes regionais de atenção à saúde mental, cujo objetivo é assegurar a continuidade e a integralidade do cuidado por meio de diferentes serviços e níveis de atenção ([TRAPÉ; CAMPOS, 2017](#)).

O DRS XIII abrange mais de 26 municípios com aproximadamente 1.400.000 habitantes, sendo Ribeirão Preto a principal cidade. O desenvolvimento e características atuais da RAPS de Ribeirão Preto foram descritos em pormenores em outros trabalhos ([YOSHIURA *et al.*, 2017](#); [BARROS *et al.*, 2016](#); [BARROS *et al.*, 2010](#)). Resumidamente, compreende: 1) Atenção primária

à saúde, comunidades terapêuticas e serviços de emergência geridos pelos 26 municípios; 2) Leitos psiquiátricos de internação aguda, ambulatorios especializados em saúde mental e um hospital-dia localizado em dois hospitais psiquiátricos estaduais e hospital universitário geral estadual; 3) Serviços de saúde mental residenciais gerenciados pelos municípios e pelo estado.

Devido à forte descentralização e pouca integração regional do SUS, cada um desses serviços estaduais e municípios possui seus próprios sistemas de informação. No caso dos serviços de atenção primária existe um sistema informatizado nacional de informação denominado e-SUS AB ([Departamento de Atenção à Saúde, 2016](#)), ainda em fase inicial de implementação no país. As internações psiquiátricas na região são articuladas ao nível regional por meio de um sistema de informação chamado SISAM ([YOSHIURA *et al.*, 2017](#)). Por meio do SISAM também é possível o registro de atendimentos em ambulatorios de saúde mental e registro de nomeações em unidades comunitárias de saúde mental.

Embora diferentes serviços da RAPS sejam coordenados para apoiar os cuidados de saúde mental, do ponto de vista computacional e informacional, sistemas de suporte informatizados não são integrados entre si, o que pode levar à perda de informações, duplicidade de dados, falta de consistência e retrabalho.

5.2 Fontes de Dados

Para este estudo foram consideradas 5 fontes de dados de diferentes sistemas de informação que dão suporte à RAPS: o SISAM (Sistema de Informação em Saúde Mental) responsável pelo controle das solicitações de internação nos serviços hospitalares especializados; o e-SUS AB da Unidade Básica de Saúde 1 de Santa Cruz da Esperança, que gerencia os atendimentos na atenção básica; o sistema de informação do Hospital Santa Tereza de Ribeirão Preto e o sistema de informação do Centro de Atenção Integrada à Saúde de Santa Rita do Passa Quatro, que são os serviços hospitalares especializados de referência para região; e, finalmente, também foi obtido acesso ao Barramento de Serviços do Ministério da Saúde para integração com dados demográficos dos pacientes por meio do Cartão Nacional de Saúde.

Foram realizadas visitas em cada uma das unidades participantes para detalhamento dos dados a serem integrados. As diferenças estão resumidas na [Tabela 1](#). Todas as instituições concordaram em participar do estudo, assinando o Termo de Consentimento de Uso de Dados. O banco de dados do SISAM possui dados de agosto de 2012 a março de 2016. O banco de dados do e-SUS AB de SCE possui dados de setembro de 2015 a março de 2016. Os dados do CAIS de Santa Rita são a partir de 2011 até 2016. No caso do Hospital Santa Tereza os dados disponibilizados são a partir de 2010. O Barramento do CNS do Ministério de Saúde é um meio de acesso aos dados demográficos de todos os usuários do SUS cadastrados, ou seja,

que possuem pelo menos um CNS gerado. Para este trabalho foi construído um banco de dados consultando o Barramento do MS, a partir dos CNS (ou CPFs) de todos os pacientes presentes nas outras fontes de dados.

Tabela 1 – Características gerais das fontes de informação utilizadas

Unidade/Sistema	Abrangência	Qtd de Usuários	Arquitetura	Disponibilidade
e-SUS AB SCE	municipal	pequeno (< 5)	sistema web	público
Hospital Santa Tereza	institucional	médio	sistema desktop	próprio
CAIS Santa Rita	institucional	médio	sistema de planilhas	próprio
Barramento MS	nacional	n/a	serviço web	público
SISAM	regional	alto (> 200)	sistema web	próprio

5.3 Modelo de Informação de Interoperabilidade

O modelo de informação para interoperabilidade corresponde ao conjunto de informações que devem ser trocadas e se tornarem disponíveis aos profissionais de saúde nos diferentes níveis de atenção. Para definição desse conjunto de dados foi realizada a análise dos principais processos envolvidos no cuidado do paciente na RAPS sendo destacados 4 momentos diferentes: (i) a entrada do paciente na rede a partir de novo cadastro em uma das portas de entrada; (ii) o atendimento na atenção básica; (iii) a internação em serviço hospitalar psiquiátrico e; (iv) a alta destes serviços. Para cada um dos processos foi definido um conjunto de elementos a partir da análise de documentos já existentes e em seguida esses elementos propostos foram analisados por especialistas da área que puderam propor a adição ou remoção de alguma informação. Este processo está resumido na [Figura 1](#).

Figura 1 – Processo de definição do modelo de informação de interoperabilidade

Os documentos analisados incluem fichas utilizadas nos instrumentos logísticos da RAPS, normas nacionais e documentos internacionais. As fichas utilizadas incluem: Ficha de Alta presente no SISAM e o Resumo da Alta do Hospital Santa Tereza de Ribeirão Preto. Nacionalmente

foram analisados o Conjunto Mínimo de Dados em Saúde sendo desenvolvido pelo Ministério da Saúde, o Sumário de Alta Hospitalar proposto pelo grupo técnico da ABNT de Informática em Saúde, a Ficha de Atendimento Médico presente do e-SUS AB. Internacionalmente, foram estudados documentos do Reino Unido, Canadá e Austrália. Da Austrália foram analisados os documentos: *National Minimum Data Set – Admitted Patient Mental Health Care* ([Australian Institute of Health and Welfare, 2015a](#)), *National Minimum Data Set – Residential Mental Health Care* ([Australian Institute of Health and Welfare, 2015c](#)), *National Minimum Data Set – Community Mental Health Care* ([Australian Institute of Health and Welfare, 2015b](#)). Do Reino Unido foi analisado o *Mental Health and Learning Disabilities Data Set* (MHLDDS) ([Health and Social Care Information Center, 2015](#)) e o documento analisado do Canadá foi o *Common Data Set - Mental Health* ([Ministry of Health and Long-Term Care, 2013](#)). Também foram consideradas as informações destacadas em trabalho anterior descrevendo dados mínimos para a continuidade dos cuidados ([DELFINI; MIYOSHI; ALVES, 2015](#)).

Para cada um dos principais processos envolvidos na assistência à saúde mental na RAPS, foi definido um modelo computacional de dados: (i) Formulário Demográfico, com informações demográficas e de identificação do paciente; (ii) Formulário de Atendimento na Atenção Primária à Saúde, com informações clínicas do atendimento médico no nível primário; (iii) Formulário de Admissão Hospitalar, com informações sobre o pedido de internação e o momento de admissão; (iv) Sumário da Alta Hospitalar, com dados de todo o processo de internação e alta. Esses modelos de dados foram revisados, reformulados e aprovados por um grupo de representantes profissionais dos diversos serviços envolvidos na RAPS. Esse grupo compreende cerca de 20 pessoas incluindo, representantes dos hospitais especializados em saúde mental, articuladores em saúde mental municipais e regionais, responsáveis pelo SISAM e por alguns representantes de serviços de saúde extra-hospitalares especializados. Os modelos estão descritos em detalhes no [Apêndice E](#).

5.4 Análise da Necessidade de Integração

Rinner et al. ([RINNER et al., 2016](#)) propuseram um método para quantificar a necessidade da criação de um ambiente para troca de informação em saúde. Este método foi aplicado aos serviços envolvidos no cuidado à pacientes com *diabetes mellitus* (DM) na Austrália. Uma das formas de aferir essa necessidade de interoperabilidade foi a partir da definição do grau de fragmentação do cuidado em DM. Quanto mais fragmentado é o cuidado, do ponto de vista informacional, mais necessário se faz o estabelecimento de um ambiente de troca de informação em saúde.

O grau de fragmentação foi calculado a partir da quantidade de diferentes serviços de saúde envolvidos no cuidado ao paciente com DM. Foram consideradas unidades de saúde os seguintes tipos de estabelecimentos: serviços da atenção primária à saúde, serviços especializados extra-hospitalares, hospitais e farmácias. Para esta população de estudo, foi observado que 95% dos pacientes com D.M. passaram por mais 1 serviços de saúde.

O método utilizado por Rinner et al. foi adaptado para o contexto da saúde mental e para o conjunto de dados disponível para este projeto. Esta adaptação busca aferir o impacto que um ambiente de interoperabilidade de dados em saúde na RAPS de Ribeirão Preto. A partir dos dados atualizados do SISAM, como sistema de regulação de diferentes serviços especializados no cuidado à saúde mental, foi calculado a quantidade serviços envolvidos, em média, no cuidado de um paciente na RAPS.

5.5 Análise de Qualidade de Dados

A análise do impacto das técnicas de integração de dados e interoperabilidade é realizada por meio da aferição de métricas de qualidade de dados e de indicadores de gestão antes e após a utilização da plataforma de interoperabilidade e integração (eHealth-Interop). As medidas de qualidade de dados avaliadas são as seguintes:

- **Completude:** representa o grau em que uma determinada coleção de dados possui o dado que deveria representar. Permite aferir os valores faltantes em uma coleção de dados de interesse;
- **Consistência:** mede o grau em que um determinado conjunto de dados em diferentes fontes de informação reflete a mesma informação estando sincronizadas entre si, ou seja, garantindo uma consistência da informação entre as fontes;
- **Duplicidade:** representa o grau em que um determinado dado está duplicado de forma desnecessária dentro de um conjunto de dados de interesse;
- **Acurácia Sintática:** representa o grau em que determinado dado está correto, é confiável e certificado. A acurácia sintática mede o grau em que determinado dado corresponde a um valor possível no conjunto de dados, por exemplo, idade deve ser um valor entre 0 e 120 anos

Para o cálculo das métricas de qualidade foram analisadas as fontes de dados em três momentos diferentes no processo de integração de dados que estão resumidos na [Figura 2](#). Na Fase I, os dados estão como representados originalmente nas fontes de dados. Estes dados estão no formato que foram recebidos e não sofreram nenhuma alteração. Na Fase II os dados passam por um processo de normalização, no qual os dados são limpos de possíveis erros de digitação

mais comuns e previsíveis, assim como os diferentes sinônimos são tratados. Na Fase III ocorre a integração propriamente dita dos dados. O registro do paciente é identificado nas outras fontes de dados e os valores dos atributos são integrados em um registro único de referência.

Os conjuntos de dados da Fase I e Fase III foram utilizados para o cálculo da acurácia e completude. Com essa análise pode-se averiguar como a integração dos dados melhorou a disponibilidade da informação e a qualidade dos dados. No caso da consistência, foram analisados os dados na Fase I e II pois na Fase III todos eles já estão integrados e, portanto, possuem os mesmos valores de atributos. Neste caso aferiu-se como a etapa de normalização contribui para melhoria na consistência. Para o cálculo da capacidade de detectar duplicidades, foram considerados somente os registros na Fase II, etapa em que os algoritmos de *record linkage* ainda não foram aplicados para integração.

Figura 2 – Principais fases de mudança no conjunto dos dados

RESULTADOS

Neste capítulo estão detalhados os resultados obtidos. O principal resultado da tese é o projeto e desenvolvimento da plataforma de interoperabilidade eHealth-Interop. Inicialmente é discutido a arquitetura do eHealth-Interop. Depois são discutidos os principais componentes: o servidor de terminologias, o módulo de *record linkage* e aspectos de segurança. A API de comunicação, por meio da qual é feita a troca de informação, é apresentada junto com as ferramentas de acesso desenvolvidas. Os resultados do caso de uso em saúde mental são demonstrados com os resultados das análises de qualidade obtidos até o momento. Finalmente, são apresentadas as publicações científicas decorrentes do trabalho.

6.1 Arquitetura de Software

A arquitetura de um sistema computacional é uma metáfora à arquitetura de uma construção física (PERRY; WOLF, 1992) pois busca definir as estruturas gerais tais como os seus principais elementos, os relacionamentos entre esses elementos e as propriedades dos elementos e desses relacionamentos (CLEMENTS *et al.*, 2010). Segundo Soni et al (SONI; NORD; HOFMEISTER, 1995) a arquitetura conceitual de um sistema define uma estrutura de alto nível e é independente das formas de codificação.

No domínio médico, para o contexto de integração em Redes de Atenção à Saúde, foi proposto a arquitetura conceitual do framework de interoperabilidade demonstrada na Figura 3. A arquitetura segue um modelo com 3 camadas principais: Camada de Dados, onde são representados e armazenados os dados integrados; Camada Semântica, na qual são descritos os esquemas mediadores e o mapeamento com ontologias; e, por fim, a Camada de Comunicação, onde são gerenciados aspectos de segurança e comunicação com as fontes de dados.

Figura 3 – Arquitetura conceitual do eHealth-Interop

Cada camada é uma abstração lógica para melhor organização das estruturas do eHealth-Interop e é composta por componentes de software específicos. Na camada de dados, as informações são armazenadas de acordo com o modelo dual, sendo divididas em dois grupos: demográfico e clínico. Esta distinção procura estabelecer uma base sólida para os algoritmos de identificação do paciente e, ao mesmo tempo, tornar a base de dados clínica flexível para ser estendida a diferentes contextos de aplicação no campo médico (BEALE *et al.*, 2008). O módulo de *Record Linkage* (RL) é responsável por garantir que não haja duplicidade do registro do paciente por meio da informação fornecida no Banco de Dados Demográficos.

A Camada Semântica é composta pelo servidor de terminologias (eHealth-Interop Terminology Server - descrito em detalhes adiante), responsável pelo armazenamento de codificações, terminologias e ontologias, bem como suas relações. Compreende também os conceitos clínicos e sociodemográficos que compõem os modelos de informação e o módulo de gestão de esquemas que liga o modelo de informação com as terminologias e faz a comunicação com as bases de dados.

A Camada de Comunicação é responsável pela geração da API de Serviço a partir dos modelos de informação representados na Camada Semântica e pelo Módulo de Segurança que verifica a pessoa responsável pelo processo de autenticação e autorização dos dados sendo enviados

e recuperados. A troca de informações da fonte de dados para a plataforma de interoperabilidade é feita por meio de adaptadores, que são agentes de software que traduzem o esquema de origem para o modelo global fornecido pela interface de serviço.

No estabelecimento de uma rede computacional para troca de informação em saúde, a plataforma eHealth-Interop atua como um repositório central dos dados trocados. Neste cenário, os modelos de informação de interoperabilidade definidos, servem de base conceitual para a troca dos dados na rede. Toda comunicação e troca de dados é realizada por meio de serviços web. Cada fonte de dados, que são os sistemas de informação em saúde nos diferentes serviços, envia e consulta dados dos seus pacientes por meio de clientes de serviço web, que se comunicam em tempo real com o eHealth-Interop.

O processo de envio dos dados do paciente pode ser comparado com o passos de Extração-Transformação-Carga (ETL) em um processo de integração de dados clássico(VASSILIADIS; SIMITSIS; SKIADOPOULOS, 2002). Primeiramente os dados de interesse são extraídos das bases de dados locais, no formato que é próprio de cada fonte. Em seguida, os dados são transformados para se conformar com o modelo de dados de interoperabilidade. Após os dados são carregados no eHealth-Interop, por meio do envio destes usando *endpoints* específicos fornecidos pela plataforma. Em seguida os dados passam por uma análise sintática e de duplicidade, visando melhorar a qualidade das informações. Caso o registro do paciente já se encontre cadastrado no eHealth-Interop, as novas informações serão agregadas tornando um registro mais completo e com mais informações.

6.2 Servidor de Terminologias

Como componente importante da plataforma eHealth-Interop foi projetado e desenvolvido um servidor de terminologias clínicas denominado eHealth-Interop Terminology Server (eHI-TS). O eHI-TS foi projetado de maneira modular para atuar como parte da plataforma integradora (componente integrante da Camada Semântica), mas também para que possa ser usado de modo independente ou adaptado para outros contextos. Este servidor foi desenvolvido para atuar como o ator “*Value Set Repository*”, definido no perfil SVS do IHE, assim como o módulo de terminologias do HL7 FHIR. A seguir são detalhados os principais componentes que compõe a arquitetura do servidor de terminologias, os principais atores e as operações que são disponibilizadas pela API *Restful* desenvolvida, assim como as terminologias clínicas definidas para apoio à interoperabilidade de dados em saúde mental.

6.2.1 Arquitetura do Servidor de Terminologias

A arquitetura do eHealth-Interop Terminology Server possui os seguintes componentes principais (Figura 4): BD Terminologia, um banco de dados baseado em documentos, foi utilizado o MongoDB como gerenciador de banco de dados; Modelos Base, conjunto de classes que

representam os principais modelos de dados utilizados para gerenciar terminologias, ontologias, conceitos, *value sets* e mapeamentos; Módulo de Mapeamento é responsável pelas funcionalidades de alto nível relacionadas ao gerenciamento dos mapeamentos entre conceitos de diferentes terminologias; Módulo de Conhecimento é responsável pelas funcionalidades de alto nível para busca de conceitos, terminologias e *value sets*; RESTful API é o conjunto de funcionalidades que são expostas a aplicações cliente do servidor por meio de web services, a partir dos quais é possível a construção de APIs específicas de uma determinada linguagem de programação, tais como Java, Android, Angular entre outras. O eHealth Interop Terminology Server foi construído utilizando o Framework Loopback (versão 3) (Strongloop, 2017).

Figura 4 – Arquitetura do eHealth-Interop Terminology Server

Seguindo a recomendação do IHE e do HL7 FHIR, as funcionalidades de autenticação e segurança são opcionais mas estão representadas nos seguintes componentes: BD Auth, banco de dados para cadastro dos usuários, papéis e aplicações clientes, foi utilizado o banco de dados relacional MySQL; Modelos Auth, são as classes que permitem o gerenciamento do modelo de dados armazenado no AuthDB; e Módulo Auth, que possui a lógica de negócios para gerenciamentos das funcionalidades de autenticação e autorização.

Na Figura 5 mostramos as principais classes que compõe o núcleo do eHealth-Interop Terminology Server:

- *CodeSystem* representa uma determinada terminologia, ontologia ou vocabulário controlado externo (LOINC, SNOMED, CID-10, CIF). Além disso, o *CodeSystem* também pode representar uma codificação criada localmente, como códigos para representar orientação sexual ou estado civil, por exemplo.

- *ConceptDefinition* representa um determinado conceito ou termo pertencente a um *CodeSystem*, por exemplo, o termo “F03 - Demência não especificada” para o CID10.
- *ValueSet* é um agrupamento de conceitos que devem ser utilizados em conjunto, para determinado contexto, pertencentes a um ou mais *CodeSystem*. No caso da saúde mental, o capítulo V do CID10, “Transtornos Mentais e Comportamentais”, pode ser considerado um *ValueSet* com os códigos abrangendo do F00 até F99.
- *ConceptMap* representa o mapeamento entre dois *CodeSystems*. Um exemplo de *ConceptMap* pode ser o mapeamento entre o CID10 e o SNOMED, ou o mapeamento entre uma codificação local de estado civil para o SNOMED.
- *Mapping* representa um único mapeamento semântico, entre dois conceitos, assim como o tipo (equivalência) entre esses conceitos. Um *Mapping* representaria, por exemplo, a equivalência entre o conceito F03 - “Demência não especificada” do ICD10 com o conceito Dementia Disorder (C0497327) do SNOMED.

Figura 5 – Diagrama com as principais classes do eHI Terminology Server

6.2.2 Operações definidas pelo WebServices

A API do eHealth-Interop Terminology Server é a camada de comunicação com as aplicações cliente, por onde estão disponíveis as principais funcionalidades. A [Figura 6](#) mostra os principais atores envolvidos nas chamadas à API e como se relacionam por meio de grupos de operações. O eHI Terminology Server, atua como o ator *Value Set Repository* definido pelo perfil *Sharing Value Sets* do IHE. O *Terminology Consumer* é uma representação das aplicações

clientes e corresponde ao *Value Set Consumer* definido pelo IHE. O *Terminology Admin* é um ator importante responsável por algumas operações especiais de gerenciamento das próprias terminologias do *Terminology Server* e não tem um correspondente no perfil SVS do IHE.

Figura 6 – Principais atores e grupos de operações relacionados ao Servidor de Terminologia

As principais operações estão descritas na [Tabela 2](#). As operações são divididas em dois grupos: as relacionadas ao ator *Terminology Consumer* e as específicas do ator *Terminology Admin*. As operações do *Terminology Admin* são basicamente responsáveis pela configuração e manutenção do *Terminology Server*. Também são de responsabilidade do *Terminology Admin* operações que envolvem a configuração de segurança como papéis e controle de usuários. As operações do *Terminology Consumer* são implementadas pelos Módulo de Mapeamento e de Conhecimento e são as comumente utilizadas pelas aplicações clientes.

Tabela 2 – Operações disponíveis pela API do servidor de terminologias

Ator	Operação	Descrição
Consumer	Retrieve Coding System	Buscar e recuperar um coding system
Consumer	Retrieve Value Set	Buscar e recupera um value set
Consumer	Concept Lookup	Buscar e recuperar uma definição específica de um conceito
Consumer	Retrieve Mapping	Buscar e recuperar um mapeamento
Admin	Add, Delete Coding Systems	Adicionar ou deletar um coding system e seus conceitos relacionados
Admin	Add, Delete Value Sets	Adicionar ou deletar value sets e suas referências à conceitos
Admin	Add, Delete Mappings	Adicionar ou deletar mapeamento entre conceitos

6.2.3 Terminologias Utilizadas

Para dar suporte ao modelo de informação definido para o cuidado continuado na RAPS foram utilizados 5 terminologias ou codificações:

- CID-10 - International Statistical Classification of Diseases and Related Health Problems ([World Health Organization, 2017a](#))
- CIF - Classificação Internacional de Funcionalidade, Incapacidade e Saúde([World Health Organization, 2017b](#))
- CIAP-2 - Classificação Internação da Atenção Primária versão 2([World Health Organization, 2012](#))

- Tabela de Procedimentos, Medicamentos e Órteses, Próteses e Materiais do SUS
- Classificação Brasileira de Ocupações

Além disso foram criadas codificações locais para representar diferentes conceitos tais como: sexo, estado civil, orientação sexual, entre outras. Essas codificações, em sua maioria, estão de acordo com aquelas utilizadas no e-SUS AB ou baseadas no modelos de informação do sumário de alta hospitalar.

6.3 Módulo de Record Linkage

Para solucionar o problema de diversos registros para um mesmo paciente foi proposto um módulo de *Record Linkage* que utiliza uma verificação de dois estágios para assegurar a unicidade do registo do paciente, este processo está resumido na [Figura 7](#). O primeiro passo é extrair identificadores do paciente, tais como CPF, RG, CNS ou registro do paciente e, então, aplicar o algoritmo de *linkage* determinístico usando esses identificadores. Se um paciente não for encontrado, os algoritmos de *linkage* probabilístico são aplicados usando informações demográficas. No caso de ser capaz de detectar uma duplicidade, este registro é sinalizado e deve passar por um processo de auditoria humana em que um especialista analisa as informações e decide se realmente ocorreu uma duplicidade ou não. Para implementar o módulo de *record linkage*, foi utilizada a biblioteca em python Dedupe ([GREGG; EDER, 2016](#)).

6.4 Módulo de Segurança

A segurança e a privacidade das informações são aspectos importantes em qualquer sistema de informação. Na área médica normalmente se torna uma barreira importante quando se trata da troca de informações em saúde entre diferentes instituições ([EDEN et al., 2016](#)). A preocupação com a segurança e privacidade das informações foi adotada desde o design da arquitetura da plataforma de integração, no mecanismo de troca de dados e também na definição da hierarquia de acesso aos dados por meio dos usuários.

A arquitetura da camada de dados dividida em repositórios diferentes que armazenam a informação demográfica do paciente e a informação clínica permite que estes repositórios sejam disponibilizados em servidores diferentes, com métodos de acessos diversos, aumentando a segurança em um nível baixo de acesso. Todo acesso aos dados é feito por meio de serviços web que utilizam o protocolo HTTP over TLS, estabelecendo um canal seguro de comunicação entre o cliente e o servidor, onde os dados são criptografados ([The Internet Society, 2000](#)). Toda aplicação para se integrar ao eHealth-Interop deve ser previamente cadastrada assim como os usuários que irão acessar os dados por meio dessa aplicação. A autenticação das aplicações cliente é realizada por meio de chave criptografada, além da autenticação do usuário acessando. A

Figura 7 – Fluxo de identificação do paciente com o módulo de Record Linkage

privacidade e compartilhamento dos dados dos pacientes é realizada seguindo o modelo adotado pelo Ministério da Saúde do Brasil para a atenção básica que é o modelo *opt-out*, em que o paciente deve optar por não compartilhar os dados. Todas as ações realizadas são armazenadas permitindo estabelecer uma trilha de auditoria com o registro dos eventos.

6.5 API de Comunicação

Toda comunicação entre as diferentes fontes de informação e a plataforma eHealth-Interop é feita por meio de serviços web. A API desenvolvida disponibiliza uma série de serviços para cadastro e busca de pacientes na plataforma, assim como de toda informação clínica associada a estes. A informação clínica está representada no modelo de informação definido *à priori* em diferentes documentos clínicos que são interoperados durante diferentes momentos do processo de cuidado do paciente. Também foi desenvolvida uma ferramenta web para auxiliar na construção de adaptadores mostrando todos os serviços disponíveis no eHealth-Interop, assim como a forma de acessá-los (Figura 8).

Figura 8 – Interface de desenvolvedor do eHealth-Interop

Além da ferramenta de auxílio na construção de adaptadores, também foi construída uma aplicação web de administração do eHealth Interop (Figura 9). Por meio dessa aplicação web, é possível gerenciar o acesso ao eHealth-Interop, assim como o gerenciamento do servidor de terminologias e a auditoria das possíveis duplicidades de registros encontrados pelo módulo de *Record Linkage*.

Figura 9 – Interface de administração do eHealth-Interop

6.6 Caso de Uso: Interoperabilidade na Rede de Atenção Psicossocial

6.6.1 Análise da Necessidade de Integração

Para análise da necessidade de integração, baseado no método de Rinner et al. (RINNER *et al.*, 2016), foram considerados os dados atualizados do SISAM dos anos de 2012 até 2017. O objetivo desta análise foi quantificar por quantos serviços diferentes um paciente têm contato na RAPS.

Neste cálculo foram considerados os seguintes tipos de serviços: unidades da atenção primária à saúde (que possuem acesso ao SISAM), unidades extra-hospitalares especializadas como CAPS e ambulatórios de saúde mental, UPAs, a AMA (Associação de Amigos do Autista), comunidades terapêuticas e hospitais especializados em saúde mental. A identificação de quais unidades participaram do cuidado ao paciente se deu a partir da análise dos seguintes processos no banco de dados do SISAM: (i) cadastro do paciente, (ii) solicitações de avaliação ambulatorial, (iii) solicitações de internação, (iv) internações, (v) encaminhamentos para a AMA, (vi) encaminhamentos para comunidades terapêuticas, (vii) agendamento de consultas especializadas e (viii) ordens judiciais ligadas à saúde mental.

Nestes 5 anos, foram analisados 9984 pacientes no total. Um percentual de 82,33% destes pacientes tiveram, ao menos 2 serviços diferentes de saúde envolvidos no cuidado. Na Figura 10 está apresentada a distribuição da quantidade de pacientes por número de serviços envolvidos. A média de serviços de saúde por paciente foi de 2,52.

Figura 10 – Distribuição da quantidade de pacientes por número de serviços de saúde diferentes envolvidos no cuidado

6.6.2 Arquitetura dos Adaptadores

Os adaptadores são os componentes de software responsáveis por extrair e transformar os dados do esquema das fontes para o esquema alvo correspondente aos modelos de informação de interoperabilidade já definidos. Para cada uma das fontes de informação foram projetados e implementados adaptadores seguindo um modelo geral que foi adaptado para o nível de acesso possível àquela fonte específica. O modelo geral tem 3 principais subcomponentes: conversor, que converte o modelo de dados base da fonte para o modelo de dados do eHealth-Interop dependendo do tipo de informação a ser trocada; cliente de serviços web, que faz a comunicação com o eHealth-Interop; controlador, responsável por receber as chamadas para troca de dados e comunicar com os outros componentes durante todo o processo.

6.6.2.1 Adaptador SISAM

No caso do SISAM, foi obtido acesso ao código fonte do sistema e ao banco de dados, permitindo a construção de um adaptador que atuasse para permitir o maior nível de interoperabilidade, a interoperabilidade semântica. Na [Figura 11](#) é apresentada a arquitetura geral do adaptador do SISAM.

Toda a troca de informação é iniciada a partir de gatilhos que são acionados nos seguintes processos: (i) no cadastro, busca ou visualização dos dados do paciente, (ii) no momento em que é feito um pedido de internação, e (iii) no momento que é informada a alta de um paciente. Dependendo do gatilho acionado, o controlador acessa o banco de dados do SISAM obtendo as informações relevantes e aciona o conversor que realiza a transformação dos dados. Em seguida, a informação é enviada ao eHealth-Interop pelo cliente web-services. Quando ocorre uma busca ou visualização de dados do paciente o processo é inverso e a informação é obtida no eHealth-Interop e apresentada ao usuário pelo SISAM.

Figura 11 – Arquitetura do adaptador do SISAM

6.6.2.2 Adaptador e-SUS AB

O e-SUS AB possui um modelo de integração que é somente de uma via, ou seja, permite somente a inserção de dados e não a recuperação. Seu modelo de integração utiliza o Apache Thrift, que permite, a partir de um documento de especificação, a geração de código em diferentes linguagens. Com o uso do Thrift a interação entre dois sistemas é realizada por meio de chamadas remotas de procedimento (do inglês RPC - *Remote Procedure Call*).

Isso foi planejado para que municípios que já possuíssem sistemas próprios pudessem adaptar seus sistemas para enviar os dados de atendimento na atenção primária para o DATA-SUS sem a obrigação de utilizar o próprio e-SUS AB. Apesar disso, o e-SUS AB é instalado localmente, na unidade que vai utilizá-lo ou de forma centralizada para todas as unidades do município. Em ambos os casos, o banco de dados permanece no próprio local de instalação e foi a partir dessa premissa que o adaptador do e-SUS AB foi planejado e desenvolvido.

Com a permissão dos gestores do município e do serviço para o acesso ao banco de dados do e-SUS AB, foi realizada a engenharia reversa do banco de dados. A partir dessa engenharia reversa foi possível identificar como a informação é armazenada, representada e gerenciada pelo e-SUS AB. A partir desse conhecimento foi construído o Adaptador do e-SUS AB

A [Figura 12](#) apresenta a arquitetura do adaptador do e-SUS AB. Devido à limitação de acesso não foi possível a construção de gatilhos como no adaptador do SISAM, porém, o adaptador do e-SUS AB foi planejado para atuar diretamente no banco de dados do sistema. Um novo componente foi incluído, denominado Sincronizador, que verifica em um período determinado de tempo, a existência de novos registros referentes a processos de interesse. Neste caso os processos monitorados são: cadastro de um novo paciente e atendimento na atenção primária à saúde. Uma vez que o sincronizador encontra um novo registro desses processos, ele aciona a ação correspondente do controlador que obtém o modelo de dados traduzido pelo conversor e em seguida encaminha os dados para o eHealth-Interop por meio do cliente de web services. Um outro componente importante neste modelo é o banco de dados do adaptador, que é um pequeno banco local em que são armazenados todos os registros já processados, assim somente os novos registros são enviados ao eHealth-Interop.

6.6.2.3 Adaptador CAIS-SR e HST

O CAIS-SR não possui um sistema de informação hospitalar geral com dados do prontuário do paciente, porém, para dar suporte à tomada de decisão, os dados são organizados internamente por meio de planilhas, documentos eletrônicos e sistemas de informação federais, que são em maioria utilizados para faturamento. O HST, embora possua um sistema de informação hospitalar, é gerenciado por uma empresa privada, não sendo possível acesso direto ao sistema ou ao banco de dados. Este cenário é, possivelmente, o mais comum encontrado na realidade dos diferentes serviços em saúde no Brasil.

Figura 12 – Arquitetura do adaptador do e-SUS AB

Nestes casos foi planejado um modelo para rápida obtenção, transformação e incorporação dos dados ao eHealth-Interop tornando-os disponíveis para toda a RAPS. Os dados de CAIS-SR e do HST foram disponibilizados em arquivos do tipo CSV (*Comma-Separated Values*). A arquitetura do adaptador é demonstrada na Figura 13 e segue um fluxo muito parecido com a do adaptador do e-SUS AB. A diferença é que, neste caso, os dados são incorporados por meio dos arquivos CSV ao invés da busca direta no banco de dados. Neste modelo, a atualização pode ser determinada com a frequência com que estes arquivos possam ser disponibilizados de forma automatizada ou não.

Figura 13 – Arquitetura do adaptador para CAIS e HST

6.6.3 Análise de Qualidade de Dados

A seguir são descritos os resultados quanto às métricas completude, acurácia, consistência e duplicidade avaliadas para os conjuntos de dados do estudo: SISAM, e-SUS AB de Santa Cruz da Esperança, CAIS de Santa Rita, Hospital Santa Tereza e os dados obtidos via o Barramento do Ministério da Saúde. Na [Tabela 3](#), estão descritos os totais de registros de paciente obtidos para cada uma das fontes de dados.

Tabela 3 – Quantidade de registros obtidos para cada uma das fontes de dados

Fonte de Dados	Total de Registros de Paciente
SISAM	8109
e-SUS AB SCE	821
CAIS Santa Rita	342
Hospital Santa Tereza	18081
Barramento MS	3987

O número de registros de pacientes obtidos com as fontes de dados em saúde mental é de 27353, somando com os registros obtidos pelo Barramento do CNS, totaliza 31340 registros. Desses, 9548 (30,47%) são registros que podem ser integrados, ou seja, o mesmo paciente está presente em pelo menos mais de 1 fonte de informação. A quantidade total de registros de pacientes únicos, após a integração, é de 4252, ou seja, aproximadamente 2,25 registros por paciente em cada fonte de dado. Na [Tabela 4](#) estão descritas a quantidade de registros únicos e a possibilidade de integração nas cinco fontes de dados possíveis.

Tabela 4 – Distribuição da quantidade de registros de paciente integráveis por fonte de dado

Qtd de fontes de dados que possuem o mesmo registro de um paciente	Qtd de registros únicos do paciente	Qtd total de registros integrados
2	3247	6494
3	968	2904
4	35	140
5	2	10
Total	4252	9548

As análises de qualidade foram realizadas a partir do conjunto de informações demográficas por serem estas, as informações em comum entre todas as fontes de informação integradas. Como descrito no [Apêndice E](#), o Formulário Demográfico é composto por 26 diferentes conceitos (atributos ou variáveis).

Tanto as análises foram realizadas a nível de atributo demográfico para cada registro de cada fonte de informação. Na completude, um determinado atributo de um registro em uma fonte de dado, pode receber o valor 1 ou 0. O valor 1 de completude representa que existe alguma informação para aquele atributo e 0 que não existe. De forma análoga, um atributo (exemplo: data de nascimento do paciente), pode receber o valor de 1 ou 0 para acurácia sintática.

Para determinar esse valor, foram definidas regras de análise sintática, por exemplo, a data de nascimento precisa ser uma data válida, considerando que idade máxima é de 120 anos. Cada valor de atributo é testado segundo a sua respectiva regra sintática, por exemplo, se o paciente tem a data de nascimento como 01-02-1970, então é aplicado a regra para verificar se este valor é uma data válida e sua idade é menor que 120 anos. Caso passe no teste, então é atribuído o valor de acurácia sintática 1, caso contrário, 0.

Neste exemplo, se a data de nascimento possui completude 76% e acurácia de 83%, para o conjunto total de dados, significa que em 76% dos registro de paciente analisados existe alguma informação referente à data de nascimento. Porém, somente 83% desses valores podem ser considerados sintaticamente corretos.

A [Tabela 5](#) descreve as análises de completude e acurácia para o conjunto de registros integrados. O conjunto de registros integrados corresponde a todos os registros de pacientes que estavam em presentes em ao menos duas fontes de informação diferentes. As análises foram realizadas levando em conta o conjunto total de informações demográficas. Tanto a completude, acurácia e consistência foram calculadas para cada atributo demográfico (nome de paciente, data de nascimento, etc). A partir desses valores, para cada análise, foram calculados o valor mínimo, máximo, a média e desvio padrão. O resultado de cada métrica, para cada atributo, assim como o cálculo das estatísticas descritivas, foram aplicados antes e após os dados serem integrados no eHealth-Interop. O valor entre parênteses é a diferença observada com o uso da ferramenta (comparação dos dados antes e depois).

Tabela 5 – Análise de completude e acurácia no conjunto total de dados integrados

	Completude (%)		Acurácia (%)	
	Não integrado	Integrado (dif)	Não integrado	Integrado (dif)
min	0,00	0,00 (=)	55,42	68,67 (+13,25)
max	100,00	100,00 (=)	100,00	100,00 (=)
média	56,47	74,87 (+18,40)	95,69	96,77 (+1,08)
d.p.	33,82	35,14 (+1,32)	10,08	7,25 (-2,83)

Para compreender quais conceitos demográficos foram mais afetados pelo compartilhamento de informação, estes foram divididos em 4 grupos diferentes:

- Geral: informações demográficas gerais, incluindo nome do paciente, nome da mãe, nome do pai, sexo, data de nascimento, município de nascimento, idade aproximada, ocupação, raça/cor, nome social, estado civil;
- Identificação: informações que podem identificar univocamente o paciente como CPF, CNS, RG e órgão emissor do RG;
- Contato: informações para contato do paciente ou responsável como telefone residencial, celular, telefone comercial, e-mail;

- Endereço: informações do endereço de residência do paciente abrangendo logradouro, bairro, complemento, número, estado, município e CEP;

A análise de completude realizada por grupo está descrita na [Tabela 6](#), enquanto na [Tabela 7](#) está descrita a análise de acurácia por grupo. Para cada grupo de informações, foram calculadas as métricas antes e após a integração das fontes de informação no eHealth-Interop, assim como a diferença observada.

Tabela 6 – Análise de completude por grupo de atributos

Completo.	Geral	Geral Integrado	Ident.	Ident. Integrado	Contato	Contato Integrado	Endereço	Endereço Integrado
min	0,00	1,25 (+1,25)	0,00	64,09 (+64,09)	0,00	3,28 (+3,28)	5,64	19,34 (+13,70)
max	100,00	100,00 (=)	94,57	98,39 (+3,82)	72,9	85,18 (+12,28)	80,93	99,46 (+18,53)
média	62,57	76,02 (+13,45)	57,11	84,98 (+27,87)	43,9	53,50 (+9,60)	59,99	87,14 (+27,15)
d.p.	35,20	35,41 (+0,21)	40,24	14,65 (-25,59)	38,66	43,98 (+5,32)	26,44	29,92 (+3,48)

Tabela 7 – Análise de acurácia por grupo de atributos

Acurácia	Geral	Geral Integrado	Ident.	Ident. Integrado	Contato	Contato Integrado	Endereço	Endereço Integrado
min	55,42	68,67 (+13,15)	95,65	98,49 (+2,84)	99,82	99,78 (-0,04)	83,28	86,62 (3,34)
max	100,00	100,00 (=)	100,00	100,00 (=)	99,90	100,00 (=)	100,00	100,00 (=)
média	93,83	95,21 (+1,38)	97,96	99,30 (+1,34)	99,86	99,90 (0,04)	95,55	96,43 (0,88)
d.p.	13,85	9,88 (-3,97)	2,19	0,81 (-1,38)	0,06	0,11 (0,05)	7,00	5,79 (-1,21)

Na [Figura 14](#) são apresentados os valores de completude calculados para cada conceito demográfico no conjunto total de registros integrados. Os valores estão em termos percentuais. A porção à direita em cada barra (na cor azul) representa a melhoria obtida após a integração.

A análise de acurácia para cada conceito demográfico está demonstrada na [Figura 15](#). A análise também foi realizada para o conjunto total de registros integrados. A porção à direita em cada barra (na cor azul) representa a melhoria obtida após a integração.

A análise de consistência foi realizada para cada fonte de informação, comparando os dados entre si, antes e após o processo de normalização. Na [Tabela 8](#) são apresentados os resultados de consistência mínima, máxima, média e desvio padrão para cada uma das fontes de informação. Estes valores foram calculados a partir da consistência individual para cada conceito demográfico.

Tabela 8 – Análise de consistência por fonte de dados

Consist.	SISAM	SISAM Norm.	HST	HST Norm.	CAIS-SR	CAIS-SR Norm.	eSUS-AB SCE	eSUS-AB SCE Norm.	Barramento MS	Barramento MS Norm.
min	0,0	0,0 (=)	0,0	0,0 (=)	0,0	0,0 (=)	0,0	0,0 (=)	0,0	24,4 (+24,4)
max	96,5	98,4 (+1,9)	93,5	98,8 (+5,4)	100,0	100,0 (=)	100,0	100,0 (=)	96,9	100,0 (+3,2)
média	29,3	48,4 (+19,1)	30,6	45,0 (+14,4)	39,8	66,5 (+26,7)	29,1	45,4 (+16,3)	26,0	77,5 (+51,5)
d.p.	32,5	33,3 (+0,7)	31,3	33,7 (+2,3)	39,2	38,4 (-0,8)	36,1	41,8 (+5,7)	28,5	18,5 (-10,0)

Além das análises apresentadas anteriormente, também foram realizadas análises complementares mais detalhadas, em nível de atributo, disponível no [Apêndice A](#). Na [Tabela 11](#) está descrita a completude para cada um dos conceitos demográficos no conjunto de dados que são integráveis. Da mesma forma, na [Tabela 14](#), está o resultado da análise de acurácia sintática

Figura 14 – Resultado da análise de completude dos registros integrados, calculado para cada conceito demográfico

para cada conceito em detalhe. Em ambas tabelas são mostrados valores antes e após a integração, assim como a diferença. Para análise de acurácia, na [Tabela 12](#) são apresentados os valores calculados para cada atributo antes da integração, enquanto na [Tabela 13](#) são apresentados os valores de acurácia após a integração, assim como a diferença correspondente. De modo semelhante foi realizada a análise para completude. Na [Tabela 9](#) são apresentados os valores de completude antes da integração, enquanto na [Tabela 10](#) é apresentado o cálculo da métrica após a integração. O resultado da análise de consistência para cada conceito demográfico pode ser visto na [Tabela 15](#). Na [Tabela 16](#) são apresentados os valores de consistência após a etapa de normalização.

Figura 15 – Resultado da análise de acurácia dos registros integrados, calculado para cada conceito demográfico

Para análise de duplicidade foi adotada uma estratégia semi-automática: inicialmente o módulo de *record linkage* foi utilizado para realizar uma pré-seleção de possíveis duplicidades e, em seguida, cada resultado foi verificado manualmente para a determinação de um conjunto curado de duplicidades (Figura 16). O algoritmo de *record linkage* foi configurado para levar em consideração somente o nome do paciente e nome da mãe, pois possuem 100% de percentual de completude e o limiar adotado de 0.75 buscando uma alta sensibilidade, ou seja, quantidade alta de verdadeiros positivos à custa de também um grande número de falsos positivos. De um total de 6427 registros, foram pré-identificadas 1066 duplicidades e verificadas 226.

Figura 16 – Processo de record linkage para identificações de duplicidades

6.6.4 Caso de Exemplo

Para exemplificar como a plataforma pode ser aplicada na prática, será discutido o exemplo de um paciente real. Neste exemplo, os dados de identificação unívoca e outros dados demográficos foram alterados para garantir a anonimidade, porém mantendo as informações gerais mais relevantes ao contexto.

A paciente Maria da Silva, natural de João Pessoa, 60 anos, residente em Santa Cruz da Esperança era acompanhada regularmente pela UBS municipal. Em 23 de julho de 2015 necessitou de uma avaliação psiquiátrica de urgência e foi realizado um pedido via SISAM para a Unidade de Emergência do Hospital das Clínicas da FMRP. O médico psiquiatra na UE verifica a falta de informação e solicita mais dados via SISAM em 16 de agosto. Após 10 dias, a paciente é avaliada e internada na UE. Em seguida uma solicitação de transferência é feita para o CAIS de Santa Rita. A transferência é cancelada e a paciente é indicada para o Hospital Santa Tereza. O médico psiquiatra do HST solicita maiores informações do histórico da paciente. Após 2 dias do pedido, a paciente é internada no HST onde permanece em tratamento por 14 dias, obtendo alta clínica e seguindo acompanhada pela UBS de origem.

Do ponto de vista computacional, diversos sistemas de informação em saúde estão envolvidos no processo de internação em saúde mental da paciente Maria da Silva: o e-SUS AB, onde foi feito o cadastro inicial, juntamente com o prontuário e as fichas de atendimento na atenção primária à saúde; O SISAM, em que foi feito o cadastro da paciente e foi realizado o pedido de internação inicial para a UE; Por meio do SISAM também foi feita a comunicação com os serviços especializados e o encaminhamento para o HST; O sistema da UE e o sistema do HST em que novamente foi realizado o cadastro da paciente, assim como o cadastro da internação e alta.

Analisando as informações demográficas obtidas da paciente, no caso do SISAM, e-SUS AB e HST, foi verificado que nenhum dos registros está completo. Nos registros presentes no SISAM e no e-SUS AB estão faltando informações relevantes como bairro, logradouro e número de residência, nome do pai e ocupação. No e-SUS AB faltam o RG, CPF, estado civil, telefone residencial e o CEP da residência, totalizando 25% de completude. No SISAM falta ainda o município de nascimento, totalizando 54% de completude. Os registros no HST e o obtido no

Barramento do Ministério da Saúde estão mais completos, ambos com 64% de completude. No HST estão faltando informações de raça/cor, o RG e bairro de residência. No Barramento do MS não constam informações de estado civil, ocupação e telefone residencial.

Com os dados todos centralizados no eHealth-Interop após o processo de interoperabilidade e integração, foi verificado que o registro demográfico final está mais completo, contendo a informação conjugada de todas as fontes de informação, totalizando 75% de completude. Além disso, se o eHealth-Interop estivesse operacional na época da internação dessa paciente, seria evitado retrabalho em todos os sistemas para o recadastro da paciente. Possivelmente, informações que os médicos solicitaram, a respeito do histórico, poderiam estar disponíveis por meio do acesso aos atendimentos na atenção primária à saúde. Um outro ponto a se considerar é, no momento em que a paciente tiver alta dos hospitais, os profissionais da APS podem realizar o acompanhamento das informações hospitalares tanto pelo SISAM quanto pelo sumário de alta integrado dos sistemas de informação hospitalares.

6.7 Resultados Complementares

Além dos resultados apresentados anteriormente, dois outros resultados complementares foram desenvolvidos, com o objetivo de tornar o eHealth-Interop uma plataforma mais generalizável e alinhada com a estratégia de eSaúde do Brasil. Foi realizado mapeamento semântico entre os conceitos do modelo de informação do Formulário Demográfico com a terminologia SNOMED-CT. Além disso, foi incorporado ao eHealth-Interop o modelo de informação do Conjunto Mínimo de Dados da Atenção à Saúde (CMD).

6.7.1 Mapeamento com o SNOMED-CT

A SNOMED CT (Systematized Nomenclature of Medicine - Clinical Terms), é uma terminologia clínica internacional presente em mais de 50 países. A sua história se inicia com a criação do SNOP (Systematized Nomenclature of Pathology) pelo Colégio Americano de Patologia. Posteriormente ela foi expandida para abranger outros contextos médicos tornando-se a SNOMED. Em 2002 ela foi agregada com outra terminologia, a CTV3 (Clinical Terms version 3), surgindo a SNOMED CT([SNOMED International, 2017a](#)).

A terminologia SNOMED-CT foi adotada como ontologia de referência para mapeamento semântico dos conceitos dos modelos de informação. A escolha do SNOMED-CT se deu devido a três fatores principais: (i) a terminologia foi definida, pelo Ministério da Saúde do Brasil, na portaria 2.073 de agosto de 2011, como a ferramenta para codificação de termos clínicos e mapeamento para outras terminologias; (ii) é uma terminologia bastante completa, possuindo mais 300.000 conceitos, em sua maioria relacionados à saúde, porém, abrangendo também outros domínios periféricos; (iii) é amplamente utilizada internacionalmente em diversos

contextos, já possuindo mapeamentos para muitas outras ontologias. Atualmente os conceitos do SNOMED-CT, por meio do BioPortal, estão mapeados para cerca de 47.021 conceitos em 438 ontologias diferentes([National Center for Biomedical Ontology, 2017](#)).

O mapeamento do Formulário Clínico para o SNOMED-CT foi feito a partir da busca dos conceitos na ferramenta *International SNOMED Browser*([SNOMED International, 2017b](#)). Os conceitos mapeados foram apresentados e discutidos juntamente com um profissional da saúde especialista em saúde mental. O mapeamento está apresentado no [Apêndice F](#).

Do total de 26 conceitos demográficos e 2 conceitos relacionados à unidade de saúde, 18 (69%) possuíam uma correspondência direta. Os conceitos relacionados a identificação do paciente (RG, Órgão Emissor do RG, CPF e CNS) foram todos mapeados para o conceito *Patient-related Identification code* (422549004). Esse conceito é conceito-pai de alguns identificadores específicos do EUA como o *social-security number*. Para o conceito “nome social” foram encontrados dois possíveis mapeamentos: *Patient alias name* (429086007) e *Patient preferred name* (408676007). Para três atributos do endereço de residência do paciente (número da casa, bairro e complemento) não foi encontrado conceitos equivalentes no SNOMED então foram associados ao conceito Address (734004005) do tipo *property*, em que pode-se definir uma propriedade associada ao endereço do paciente.

6.7.2 Incorporação do Conjunto Mínimo de Dados à Saúde

O Conjunto Mínimo de Dados é um documento padronizado cujo objetivo é proporcionar informações essenciais, aos setores envolvidos na tomada de decisão e processos administrativos oferecendo suporte à diversas atividades incluindo pesquisa, avaliação, monitoramento e faturamento. Internacionalmente, diversos países já definiram os seus CMDs. No Brasil, em 2016, o CMD foi instituído por meio da Resolução CIT n06, de 26 de agosto de 2016.

Em trabalho anterior à formalização do CMD no Brasil, foram estudados diversos documentos de dados mínimos ligados à saúde mental. Estes resultados estão apresentados em ([DELFINI; MIYOSHI; ALVES, 2015](#)) e ([DELFINI, 2015](#)).

Posteriormente, após a definição do CMD no Brasil, foi adotado o modelo de informação do CMD como um dos modelos disponibilizados para interoperabilidade pelo eHealth-Interop. O modelo de informação do CMD utilizado foi a versão de agosto de 2016. Também foi realizado um estudo de mapeamento das informações geradas na atenção primária à saúde, por meio do sistema eSUS-AB PEC e CDS, para o CMD e sua incorporação no eHealth-Interop. Esses resultados são detalhados em ([NEVES, 2016](#)).

6.8 Publicações

Foram escritos três publicações com os resultados deste trabalho:

- "Health Terminology Server for Portuguese-speaking Communities": apresentado no WHO-FIC 2017 ([Apêndice B](#)).
- "Development of a Lightweight and Adaptable Clinical Terminology Server": submetido ao HCIST 2017 (International Conference on Health and Social Care Information Systems and Technologies) ([Apêndice C](#))
- "eHealth-interop: an Interoperability Platform developed Mental Care Integration": artigo submetido ao periódico *Journal of Medical Internet Research* ([Apêndice D](#))

6.9 Orientações Associadas à Tese

Dois trabalhos de conclusão de curso de alunos do curso de Graduação em Informática Biomédica da Universidade de São Paulo, estiveram envolvidos ao projeto descrito nesta tese de doutorado. O autor da presente tese atuou como orientador e co-orientador dos respectivos alunos:

- Milena Gomes Delfini, que apresentou a monografia intitulada "A Continuidade do Cuidado na Saúde Mental: Estudo dos Padrões para Informatização do Sistema de Referência e Contra-Referência", no ano de 2015 ([DELFINI, 2015](#)).
- Vinícius Bertaco Neves, que apresentou a monografia intitulada "Estudo e Projeto de um Modelo para Integração de dados de Atendimento da Atenção Básica ao Conjunto Mínimo de Dados Brasileiro", no ano de 2016([NEVES, 2016](#))

DISCUSSÃO

Nesta seção são inicialmente apresentados os trabalhos relacionados. A busca por trabalhos relacionados envolveu arquiteturas de interoperabilidade em saúde e também a aplicação e avaliação de tecnologias da informação para o contexto da saúde mental. Em seguida, os resultados, achados e experiências adquiridas são discutidos assim como as dificuldades, limitações e possíveis trabalhos futuros decorrentes da tese apresentada.

7.1 Trabalhos Relacionados

A busca na literatura pelos trabalhos relacionados foi baseada nos seguintes termos de referência: interoperabilidade, integração ou troca de dados em saúde, cuidado integrado em saúde mental, sistemas de informação em saúde mental e repositório ou servidor de terminologias e ontologias em saúde.

Cifuentes et al. (CIFUENTES *et al.*, 2015) realizaram um estudo observacional de práticas de assistência cujo objetivo é o cuidado integrado em saúde. O estudo buscou levantar os desafios ligados a tecnologia da informação na atenção primária e à saúde mental. Os dados foram coletados a partir de formulários e o relatório informativo foi analisado. No estudo participaram 8 equipes de saúde primária e 3 centros comunitários em saúde mental que fazem parte de uma iniciativa denominada “Avançando no Cuidado Coordenado em Colorado nos Estados Unidos”. As três principais dificuldades encontradas no cuidado integrado foram: documentar e acompanhar informações relevantes à saúde mental, permitir comunicação e coordenação entre diferentes equipes e, por fim, a falta de interoperabilidade entre os sistemas de informação em saúde e outros dispositivos. Os autores concluem ressaltando a importância que desenvolvedores de sistemas de registro eletrônico em saúde possam projetar os sistemas para ter funcionalidades que permitam o cuidado integrado e ao longo do tempo.

Yuksel et al. (YUKSEL *et al.*, 2016) desenvolveram a plataforma SALUS (*Scalable, Standard based Interoperability Framework for Sustainable Proactive Post Market Safety Studies*), um framework de interoperabilidade baseado em ontologias. A proposta do SALUS é lidar com as diferentes representações das fontes de informação por meio de um mediador semântico que permite a troca de dados realizando a tradução baseada em regras a partir de representações formais de cada uma das fontes integradas. A arquitetura do SALUS foi projetada para permitir a realização de diferentes tipos de estudos observacionais a partir do levantamento de dados demográficos e clínicos em diversas fontes de informação diferentes.

Moraes et al. (MORAES *et al.*, 2016) propuseram uma metodologia para troca de informação por meio de sistemas multi-agentes baseados no padrão de informação em saúde OpenEHR. O caso de estudo da metodologia foi a implementação de agentes de software para auxílio no planejamento de cirurgias cardíacas. Toda a troca de informação é feita por meio de uma linguagem comum, a ACL (*Agent Communication Language*), proposta pela FIPA (*Foundation for Intelligent Physical Agents*). Para construção dos agentes de software, os autores propõe o uso do framework JADE (*Java Agent DEvelopment Framework*). Na metodologia proposta, o OpenEHR é utilizado em dois momentos: para construção da base de conhecimento composta por arquétipos e as mensagens trocadas entre os agentes de software tem como conteúdo “*EHR Extracts*” que são uma estrutura de dados proposta pelo OpenEHR para representar registro eletrônico em saúde.

Rac-Albu et al. (RAC-ALBU *et al.*, 2016) propuseram uma metodologia para troca de documentos médicos buscando a interoperabilidade de dados em saúde para contexto da assistência à saúde na Romênia. Toda troca de documentos médicos é realizada por mensagens HL7 v2 e por meio de um middleware de integração baseado na biblioteca de software JCAPS (*Java Composite Application Platform Suite*). Os documentos médicos são esquemas XML, construídos a partir de questionários com a ajuda de um profissional médico. Os autores não implementaram toda a proposta, construíram o documento clínico para pacientes com taquicardia ventricular mas a solução de interoperabilidade não foi desenvolvida.

Diversos trabalhos têm sido propostos cujo objetivo é estabelecer um repositório de termos ou conceitos para área da saúde. Um dos mais famosos é o *Unified Medical Language System* (UMLS) (BODENREIDER, 2004), um repositório de terminologias biomédicas gerenciado pela *US National Library of Medicine*. O UMLS é composto por: *Metathesaurus*, um banco de dados de conceitos biomédicos e dos seus respectivos relacionamentos; a *Semantic Network* que são um conjunto de categorias e subcategorias que classificam os conceitos do *Metathesaurus* e o *SPECIALIST Lexicon* e *Lexical Tools* que são ferramentas de processamento de linguagem natural. O Bioportal (WHETZEL *et al.*, 2011) é um portal web, gerenciado pelo NCBO (*National Center for Biomedical Ontology*), que contém um repositório de ontologias e terminologias biomédicas assim como o mapeamento entre essas terminologias. O Bioportal também permite que a comunidade participe auxiliando na evolução e avaliação do conteúdo.

O LexEVS ([National Cancer Institute, 2016](#)), desenvolvido pela *Mayo Clinic* e o *US National Cancer Institute*, é um servidor centralizado de terminologias que serve como base de padrão terminológico para dados em diversos institutos do NIH e agências federais dos Estados Unidos assim como também pela comunidade biomédica e de pesquisa em câncer. O LexEVS inclui uma API em JAVA e implementa o protocolo CTS-2. Peterson et al ([PETERSON et al., 2016](#)) propôs um protótipo de um servidor de terminologias que segue uma implementação baseada em componentes e implementa um conjunto de serviços baseados no CTS-2. Este protótipo foi projetado em torno do processo de ETL (*Extract Transformation and Loading*) onde o usuário especialista, por meio de uma interface gráfica, auxilia para que as terminologias possam ser corretamente carregadas e mapeadas.

Em geral, os trabalhos analisados no contexto computacional de terminologias em saúde podem ser categorizados em dois tipos: (i) trabalhos em que o servidor de terminologias é a base para estabelecer um portal terminológico de contexto geral ou específico, onde usuários podem consultar termos, obter terminologias, estudar definições entre outras funcionalidades, como o UMLS e BioPortal; (ii) trabalhos em que um servidor de terminologias atua como um componente de software, parte de uma aplicação maior, geralmente envolvendo o registro eletrônico em saúde ou outras aplicações. Neste segundo caso se encontra o trabalho de Peterson et al e o LexEVS. Embora todos possuam funcionalidades semelhantes, os servidores de terminologia são normalmente ligados a um determinado contexto de aplicação ou a uma linguagem específica de programação o que dificulta sua reutilização.

7.2 Discussão Geral

7.2.1 Desenvolvimento da Plataforma eHealth-Interop

Uma das características de um sistema de saúde descentralizado é a autonomia relativa de diferentes serviços e atores, envolvidos no processo de cuidado ao paciente, na implantação de sistemas de informação em saúde. Essa autonomia permite a implantação de sistemas bastante específicos, que se especializam em resolver as necessidades locais do ambiente em que estão implantados. Essa flexibilidade é interessante na resolução e mapeamento dos processos individuais porém, cria um contexto em que uma multiplicidade de sistemas de informação são utilizados para resolução de problemas específicos. Cada sistema, desenvolvido independentemente, em diferentes linguagens de programação, arquiteturas, com diferentes terminologias e codificações. Esse contexto dificulta a criação de um ambiente de troca de informação e, consequentemente, o apoio no cuidado continuado ao paciente.

As características desse ambiente serviram como plano diretor para todo o desenvolvimento deste projeto, desde a definição da arquitetura, escolha das tecnologias, modelo de interoperabilidade e terminologias. A hipótese deste trabalho é que para o estabelecimento de rede de atenção à saúde informatizada e interoperável, neste contexto de um sistema de saúde

descentralizado, é necessário um componente de software adicional que seja responsável por centralizar, sincronizar e normalizar as informações trocadas. A alternativa a esse modelo seria um ambiente de troca de informação *peer-to-peer* onde cada serviço da rede possa interoperar diretamente com outro ator sem a necessidade de um intermediário. A abordagem *peer-to-peer*, embora robusta, causa sobrecarga maior sobre os serviços quando comparada ao modelo centralizador.

O modelo de interoperabilidade escolhido, baseado em serviços web, permite a troca de dados de forma independente de arquitetura, linguagem de programação e localização física, sendo necessária somente uma conexão com a internet ou por meio do estabelecimento de uma rede privada. O uso do framework Loopback, como base tecnológica para a construção do eHealth-Interop, permite que sejam geradas APIs específicas para determinadas linguagens de programação como: Javascript, Angular, Android, IOS, Java, Xamarin, PHP, entre outros. Essas APIs encapsulam todo o processo de conexão e autenticação, facilitando a criação dos adaptadores necessários para troca de informação.

No eHealth-Interop as informações sócio-demográficas e clínicas são divididas em dois bancos de dados distintos, visando tanto a segurança da informação quanto a sua extensibilidade. Diferente de outros trabalhos relacionados, o eHealth-Interop permite a troca de dados de pacientes sem identificadores únicos obrigatoriamente definidos, situação que ocorre principalmente no caso da saúde mental e emergências médicas.

7.2.2 Definição do Modelo de Informação

Um ponto importante em todo o processo foi o uso da abordagem sócio-técnica (GHAFARIAN, 2011) durante, principalmente, a definição dos modelos de informação para interoperabilidade. Nesta abordagem, potenciais usuários do sistema participaram ativamente auxiliando na identificação dos principais processos, definição das propriedades a serem interoperadas e na forma com que elas possam ser apresentadas.

Na definição do modelo de informação, em conversa conjunta com os profissionais da saúde, foi decidido contemplar o máximo possível dos padrões definidos pela ABNT e pelo Ministério da Saúde, mas também levando em consideração os outros documentos já utilizados pelos serviços localmente e os documentos internacionais. Essas decisões foram motivadas levando em consideração a estratégia de e-Saúde do Ministério da Saúde. Nessa estratégia, o e-SUS AB tem um papel organizador definindo o prontuário eletrônico do cidadão. No contexto hospitalar, verificou-se que o Sumário de Alta Hospitalar, contemplava as informações já utilizadas localmente pelos serviços de saúde. Alguns conceitos estavam mais distantes do contexto da saúde mental como, por exemplo, procedimentos cirúrgicos realizados, alergias, incapacidades via codificação CIF. Apesar disso, foi adotada a estratégia de manter essas informações como opcionais.

7.2.3 Definição do Servidor de Terminologias

O problema terminológico em saúde é crucial para uma série de questões: desde permitir análise secundária dos dados, comparações com outros bancos de dados, até a interoperabilidade semântica. Algumas terminologias já fazem parte do cotidiano médico, como o CID-10 para codificação de diagnóstico, porém, uma grande quantidade de informações normalmente possui codificações locais e raramente fazem referência a uma ontologia de alto nível.

Para permitir a interoperabilidade semântica, foi projetado um servidor de terminologias baseado no perfil IHE SVS e no serviço de terminologia do padrão HL7 FHIR e, por consequência, no padrão da OMG CTS-2. O design modular, seguindo a arquitetura orientada a serviços, permite que o servidor de terminologias possa ser reutilizado como um componente de software a parte para outros contextos e projetos. Um aspecto importante da solução adotada é permitir o mapeamento entre ontologias e terminologias. Esse mapeamento pode permitir a tradução automática entre diferentes terminologias.

7.2.4 Análise de Qualidade

Para verificar qual o impacto na qualidade dos dados a partir do uso de um ambiente de troca de dados em saúde foram realizadas algumas métricas. Foram averiguadas a completude, a acurácia, consistência e duplicidade. Quanto a completude, a partir da análise dos resultados presentes nas Tabelas 9 e 10, houve uma melhora na maioria das propriedades para todas as fontes de informação envolvidas. Algumas bases de dados não possuíam alguns tipos de informações como idade aproximada ou nome social e passaram a ter esses dados após a inserção no eHealth-Interop. Também houve uma melhora nas informações relacionadas ao contato e localização do paciente, que no caso da saúde mental são aspectos bastante importantes.

Já na análise de acurácia, demonstrada nas Tabelas 12 e 13, houve um comportamento diferente. Como a diferença era pouca quanto a acurácia, na Tabela 13 está demonstrada a acurácia final após a integração e, entre parênteses, a porcentagem em melhoria que houve, antes e após a integração dos dados. Em alguns casos a acurácia diminuiu, mas esse comportamento é de se esperar principalmente para as propriedades em que houve um grande aumento da completude. Isso também se deve a estratégia de integração adotada. Foi adotado a priorização da completude em favor da acurácia, ou seja, é preferível ter mais informações disponíveis e deixar que os usuários possam corrigi-las ou descartá-las, se for o caso. Conforme os registros são integrados, são agregadas mais informações, e nem sempre essas informações passam nas regras definidas para análise de acurácia. Mesmo assim, a diminuição na acurácia sintática, em praticamente todos os casos, é menor que 1%.

A análise de consistência foi feita com os dados originais e após o processo de normalização. Assim como na acurácia, na Tabela 16 estão representadas a consistência final e a diferença entre antes e após o processo de normalização dos dados entre parênteses. Pode-se notar que

houve também uma grande melhora na consistência das informações após essa etapa. Isso se deu principalmente porque cada fonte de informação possui codificações próprias e formatos próprios para representação de determinados tipos de dados como data e número.

Para validação do algoritmo de *record linkage* foi realizada a análise de duplicidade. Essa análise foi feita somente no banco do SISAM. Isso porque o SISAM é utilizado por mais de 200 serviços de saúde e não é obrigatório o preenchimento de identificadores como CNS e CPF. Neste caso a probabilidade de existir registros duplicados é alta. Verificamos que o algoritmo foi capaz de identificar duplicidades que posteriormente foram validadas manualmente.

7.2.5 Limitações e Dificuldades

O eHealth-Interop permite a troca de informações entre a atenção primária à saúde e a atenção hospitalar. Não foi possível realizar testes com dados da atenção ambulatorial especializada. Apesar disso, após a definição das informações relevantes e a geração do modelo de informação, a ferramenta é capaz de gerar automaticamente a API de comunicação e sua importação para o banco de dados. Outra limitação do projeto atual na aplicação da saúde mental é não haver um mapeamento entre os conceitos das terminologias adotadas. A ferramenta permite que esse mapeamento seja feito, mas não houve a necessidade de tradução de conceito de uma terminologia para outra.

CONCLUSÕES

Considerando os objetivos definidos para a realização deste trabalho:

- (i) Foi projetada uma arquitetura de uma plataforma computacional para interoperabilidade de dados em saúde. Essa arquitetura foi implementada e testada no contexto da saúde mental com sucesso.
- (ii) A partir de documentos e informações padronizados internacionalmente, nacionalmente e regionalmente foram definidos modelos de informação para representar processos importantes no cuidado ao paciente em saúde mental. Destaca-se a necessidade de criar os modelos a partir de um padrão já estabelecido mas permitir que a realidade local seja representada.
- (iii) Para lidar com a diferença terminológica, foi desenvolvido um servidor de terminologias e ontologias. Esse servidor é parte integrante da plataforma de interoperabilidade, porém foi desenvolvido de modo que possa ser utilizado de modo independente. Foi possível incluir no servidor as terminologias internacionais, nacionais e locais para apoio à execução do projeto.
- (iv) Buscando melhorar a qualidade das informações interoperadas foram desenvolvidos validadores e analisadores. Para validar a solução projetada, foram conduzidas análises nos conjuntos de dados obtidos, aferindo as taxas de completude, acurácia, consistência e duplicidade. Foram obtidas melhoras gerais nos indicadores de qualidade. Para identificação do paciente, na análise de duplicidade, foi estabelecida uma abordagem que leva em consideração regras determinísticas e também algoritmos probabilísticos de *record linkage*.

REFERÊNCIAS

AKHLAQ, A.; SHEIKH, A.; PAGLIARI, C. Defining Health Information Exchange: Scoping Review of Published Definitions. **Journal of Innovation in Health Informatics**, v. 23, n. 4, p. 684–764, 2016. ISSN 2058-4563. Disponível em: <<https://hijournal.bcs.org/index.php/jhi/article/view/838/960>>. Citado na página 22.

Australian Institute of Health and Welfare. **Admitted Patient Mental Health Care National Minimum Data Set**. Australian Institute of Health and Welfare, 2015. Disponível em: <<http://meteor.aihw.gov.au/content/index.phtml/itemId/588909>>. Citado na página 30.

_____. **Community Mental Health Care National Minimum Data Set**. 2015. Disponível em: <<http://meteor.aihw.gov.au/content/index.phtml/itemId/565694><http://www.webcitation.org/6qdcc9byt>>. Citado na página 30.

_____. **Residential Mental Health Care National Minimum Data Set**. 2015. Disponível em: <<http://meteor.aihw.gov.au/content/index.phtml/itemId/565678><http://www.webcitation.org/6qdcWMMiB>>. Citado na página 30.

BARROS, R. E. M.; MARQUES, J. M. d. A.; CARLOTTI, I. P.; ZUARDI, A. W.; DEL-BEN, C. M. Short admission in an emergency psychiatry unit can prevent prolonged lengths of stay in a psychiatric institution. **Revista Brasileira de Psiquiatria**, Associação Brasileira de Psiquiatria (ABP), v. 32, n. 2, p. 145–151, 6 2010. ISSN 1516-4446. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1516-44462010000200009&lng=en&tlng=en>. Citado na página 27.

BARROS, R. E. M.; MARQUES, J. M. de A.; SANTOS, J. L. F.; ZUARDI, A. W.; DEL-BEN, C. M. Impact of length of stay for first psychiatric admissions on the ratio of readmissions in subsequent years in a large Brazilian catchment area. **Social Psychiatry and Psychiatric Epidemiology**, v. 51, n. 4, p. 575–587, 4 2016. ISSN 0933-7954. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/26801498><http://link.springer.com/10.1007/s00127-016-1175-x>>. Citado na página 27.

BEALE, T.; HEARD, S.; KALRA, D.; LLOYD, D. **Demographic Information Model**. 2008. 25 p. Disponível em: <http://www.openehr.org/releases/1.0.2/architecture/rm/demographic_im.pdf>. Citado na página 34.

BENDER, D.; SARTIPI, K. HL7 FHIR: An Agile and RESTful approach to healthcare information exchange. In: **Proceedings of the 26th IEEE International Symposium on Computer-Based Medical Systems**. IEEE, 2013. p. 326–331. ISBN 978-1-4799-1053-3. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6627810>>. Citado na página 23.

BODENHEIMER, T. Coordinating care—a perilous journey through the health care system. **The New England journal of medicine**, Massachusetts Medical Society, v. 358, n. 10, p. 1064–71, 3 2008. ISSN 1533-4406. Disponível em: <<http://www.nejm.org/doi/full/10.1056/NEJMp0706165>>. Citado na página 17.

BODENREIDER, O. The Unified Medical Language System (UMLS): integrating biomedical terminology. **Nucleic Acids Research**, Oxford University Press, v. 32, n. 90001, p. 267D–270, 1 2004. ISSN 1362-4962. Disponível em: <<https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkh061>>. Citado na página 56.

_____. Biomedical ontologies in action: role in knowledge management, data integration and decision support. **Yearbook of medical informatics**, NIH Public Access, p. 67–79, 2008. ISSN 0943-4747. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/18660879http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2592252>>. Citado na página 25.

CARVALHO, A. L. B. d.; JESUS, W. L. A. d.; SENRA, I. M. V. B.; CARVALHO, A. L. B. d.; JESUS, W. L. A. d.; SENRA, I. M. V. B. Regionalização no SUS: processo de implementação, desafios e perspectivas na visão crítica de gestores do sistema. **Ciência & Saúde Coletiva**, ABRASCO - Associação Brasileira de Saúde Coletiva, v. 22, n. 4, p. 1155–1164, 4 2017. ISSN 1413-8123. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-81232017002401155&lng=pt&tlng=pt>. Citado na página 20.

CIFUENTES, M.; DAVIS, M.; FERNALD, D.; GUNN, R.; DICKINSON, P.; COHEN, D. J. Electronic Health Record Challenges, Workarounds, and Solutions Observed in Practices Integrating Behavioral Health and Primary Care. **The Journal of the American Board of Family Medicine**, v. 28, n. Supplement 1, p. S63–S72, 9 2015. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/26359473>>. Citado na página 55.

CLEMENTS, P.; BACHMANN, F.; BASS, L.; GARLAN, D.; IVERS, J.; LITTLE, R.; MERSON, P.; NORD, R.; STAFFORD, J. **Documenting Software Architectures: Views and Beyond**. 2 ed. ed. [S.l.]: Addison-Wesley Professional, 2010. 592 p. ISBN 978-0321552686. Citado na página 33.

COFFEY, R. M.; BUCK, J. A.; KASSED, C. A.; DILONARDO, J.; FORHAN, C.; MARDER, W. D.; VANDIVORT-WARREN, R. Transforming mental health and substance abuse data systems in the United States. **Psychiatric services (Washington, D.C.)**, v. 59, n. 11, p. 1257–63, 11 2008. ISSN 1557-9700. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/18971401http://ps.psychiatryonline.org/article.aspx?articleID=99882>>. Citado na página 17.

DELFINI, M. G. **A Continuidade do Cuidado na Saúde Mental: Estudo dos Padrões para Informatização do Sistema de Referência e Contra-Referência**. 85 p. Tese (Monografia) — Universidade de São Paulo, 2015. Citado 2 vezes nas páginas 53 e 54.

DELFINI, M. G.; MIYOSHI, N. S. B.; ALVES, D. Minimum Data Consensus: Essential Information to Continuing Healthcare. In: **2015 IEEE 28th International Symposium on Computer-Based Medical Systems**. IEEE, 2015. p. 205–207. ISBN 978-1-4673-6775-2. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7167487>>. Citado 2 vezes nas páginas 30 e 53.

Departamento de Atenção à Saúde. **e-SUS Atenção Básica**. 2016. Disponível em: <<http://dab.saude.gov.br/portaldab/esus.php>>. Citado na página 28.

EDEN, K. B.; TOTTEN, A. M.; KASSAKIAN, S. Z.; GORMAN, P. N.; MCDONAGH, M. S.; DEVINE, B.; PAPPAS, M.; DAEGES, M.; WOODS, S.; HERSH, W. R. Barriers and facilitators to exchanging health information: a systematic review. **International Journal of Medical Informatics**, v. 88, p. 44–51, 2016. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1386505616300041>>. Citado na página 39.

GHAFFARIAN, V. The new stream of socio-technical approach and main stream information systems research. **Procedia Computer Science**, v. 3, p. 1499–1511, 2011. ISSN 18770509. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S1877050911000408>>. Citado na página 58.

GREGG, F.; EDER, D. **Dedupe**. 2016. Disponível em: <<https://github.com/dedupeio/dedupehttp://www.webcitation.org/6qftp1G0X>>. Citado na página 39.

GU, L.; BAXTER, R.; VICKERS, D.; RAINSFORD, C. **Record Linkage: Current Practice and Future Directions**. Canberra - Australia, 2003. 32 p. Disponível em: <<http://datamining.csiro.au>>. Citado na página 24.

HAGGERTY, J. L.; REID, R. J.; FREEMAN, G. K.; STARFIELD, B. H.; ADAIR, C. E.; MCKENDRY, R. Continuity of care: a multidisciplinary review. **BMJ (Clinical research ed.)**, BMJ Publishing Group, v. 327, n. 7425, p. 1219–21, 11 2003. ISSN 1756-1833. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/14630762http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC274066>>. Citado na página 19.

Health and Social Care Information Center. **Mental Health and Learning Disabilities Data Set**. 2015. Disponível em: <<http://www.hscic.gov.uk/mhldds>>. Citado na página 30.

HERSH, W. R.; TOTTEN, A. M.; EDEN, K. B.; DEVINE, B.; GORMAN, P.; KASSAKIAN, S. Z.; WOODS, S. S.; DAEGES, M.; PAPPAS, M.; MCDONAGH, M. S. Outcomes From Health Information Exchange: Systematic Review and Future Research Needs. **JMIR medical informatics**, v. 3, n. 4, p. e39, 1 2015. ISSN 2291-9694. Disponível em: <<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4704923&tool=pmcentrez&rendertype=abstract>>. Citado na página 17.

HIMSS. **HIMSS Dictionary of Healthcare Information Technology Terms, Acronyms and Organizations**. 2. ed. [S.l.]: HIMSS Publishing, 2010. 226 p. ISBN 978-1-938904-03-5. Citado 2 vezes nas páginas 17 e 21.

HINCAPIE, A.; WARHOLAK, T. The impact of health information exchange on health outcomes. **Applied clinical informatics**, v. 2, n. 4, p. 499–507, 1 2011. ISSN 1869-0327. Disponível em: <<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3612997&tool=pmcentrez&rendertype=abstract>>. Citado na página 17.

HL7. **Terminology Module - FHIR Release 3 (STU)**. 2017. Disponível em: <<https://www.hl7.org/fhir/terminology-module.htmlhttp://www.webcitation.org/6qekctc7S>>. Citado na página 24.

_____. **Terminology Service - FHIR Release 3 (STU)**. 2017. Disponível em: <<https://www.hl7.org/fhir/terminology-service.htmlhttp://www.webcitation.org/6qekV2p4l>>. Citado na página 24.

IHE. **IHE IT Infrastructure (ITI) Technical Framework Volume 2b Transactions Part B**. [S.l.], 2016. 380 p. Citado na página 23.

_____. **IHE Profiles**. 2016. Disponível em: <<http://www.webcitation.org/6qZle3ELf>>. Citado na página 23.

Institute of Electrical and Electronics Engineers. **IEEE Standard Computer Dictionary: A Compilation of IEEE Standard Computer Glossaries**. New York, NY: [s.n.], 1990. Citado na página 22.

KONOPKA, B. M. Biomedical ontologies—A review. **Biocybernetics and Biomedical Engineering**, v. 35, n. 2, p. 75–86, 2015. ISSN 02085216. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0208521614000503>. Citado na página 25.

MENDES, E. V. **As Redes De Atenção À Saúde**. 2 ed. ed. Brasília: Organização Pan-Americana da Saúde, 2011. ISBN 9788579670756. Citado 2 vezes nas páginas 16 e 20.

Ministerio da Saude. **Portaria 4.279 de 30 de dezembro de 2010**. 2010. Disponível em: http://bvsms.saude.gov.br/bvs/saudelegis/gm/2010/prt4279_30_12_2010.html. Citado na página 20.

Ministério da Saúde. **Portaria 2.073 de 31 de agosto de 2011**. Ministério da Saúde, 2011. Disponível em: http://bvsms.saude.gov.br/bvs/saudelegis/gm/2011/prt2073_31_08_2011.html. Citado na página 22.

Ministerio da Saude. **Portaria 3.088, de 23 de Dezembro de 2011**. 2011. Disponível em: http://bvsms.saude.gov.br/bvs/saudelegis/gm/2011/prt3088_23_12_2011_rep.html. Citado na página 16.

Ministry of Health and Long-Term Care. **Common Data Set – Mental Health (CDS-MH) Manual v5**. Toronto/CA, 2013. 172 p. Disponível em: http://www.housingfirsttoolkit.ca/sites/default/files/CDS-MH_Manual_Version5.0FY20132014FINAL.doc. Citado na página 30.

MORAES, J. L. Cardoso de; SOUZA, W. L. de; PIRES, L. F.; PRADO, A. F. do. A methodology based on openEHR archetypes and software agents for developing e-health applications reusing legacy systems. **Computer Methods and Programs in Biomedicine**, v. 134, p. 267–287, 10 2016. ISSN 01692607. Disponível em: <http://linkinghub.elsevier.com/retrieve/pii/S016926071630298X>. Citado na página 56.

National Cancer Institute. **LexEVS**. 2016. Disponível em: <https://wiki.nci.nih.gov/display/LexEVS/LexEVShttp://www.webcitation.org/6qegLPzeK>. Citado na página 57.

National Center for Biomedical Ontology. **SNOMED CT - Mappings | NCBO BioPortal**. 2017. Disponível em: <https://bioportal.bioontology.org/ontologies/SNOMEDCT/?p=mappings>. Citado na página 53.

NEVES, V. B. **Estudo e Projeto de um Modelo para Integração de dados de Atendimento da Atenção Básica ao Conjunto Mínimo de Dados Brasileiro**. Tese (Monografia) — Universidade de São Paulo, 2016. Citado 2 vezes nas páginas 53 e 54.

OMG. **Common Terminology Services 2 (CTS2) Version 1.2**. [S.l.], 2015. 36 p. Disponível em: <http://www.omg.org/spec/CTS2/1.2/>. Citado na página 23.

_____. **About the Object Management Group**. 2017. Disponível em: <http://www.omg.org/gettingstarted/gettingstartedindex.htmhttp://www.webcitation.org/6qem8xCYz>. Citado na página 23.

PERRY, D. E.; WOLF, A. L. Foundations for the study of software architecture. **ACM SIGSOFT Software Engineering Notes**, ACM, v. 17, n. 4, p. 40–52, 10 1992. ISSN 01635948. Disponível em: <http://dl.acm.org/citation.cfm?id=141874.141884>. Citado na página 33.

PETERSON, K. J.; JIANG, G.; BRUE, S. M.; LIU, H. Leveraging Terminology Services for Extract-Transform-Load Processes: A User-Centered Approach. **AMIA ... Annual Symposium proceedings. AMIA Symposium**, American Medical Informatics Association, v. 2016, p. 1010–1019, 2016. ISSN 1942-597X. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/28269898http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5333225>>. Citado na página 57.

RAC-ALBU, E. M.; CIOBANU, V.; RAC-ALBU, M.; POPESCU, N. Interoperability of Medical Data Through e-Health Service in Romania. In: **Lecture Notes in Business Information Processing**, vol 247. Springer, Cham, 2016. p. 683–692. Disponível em: <http://link.springer.com/10.1007/978-3-319-32689-4_52>. Citado na página 56.

RINNER, C.; SAUTER, S. K.; ENDEL, G.; HEINZE, G.; THURNER, S.; KLIMEK, P.; DUFTSCHMID, G. Improving the informational continuity of care in diabetes mellitus treatment with a nationwide Shared EHR system: Estimates from Austrian claims data. **International Journal of Medical Informatics**, 5 2016. ISSN 13865056. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1386505616300739>>. Citado 2 vezes nas páginas 30 e 42.

SNOMED International. **History of SNOMED**. 2017. Disponível em: <<https://www.snomed.org/snomed-ct/what-is-snomed-ct/history-of-snomed-ct>>. Citado na página 52.

_____. **International SNOMED CT Browser**. 2017. Disponível em: <<http://browser.ihtsdotools.org/?>> Citado na página 53.

SONI, D.; NORD, R.; HOFMEISTER, C. **Software Architecture in Industrial Applications**. 1995. 196 p. Citado na página 33.

SPIRANOVIC, C.; MATTHEWS, A.; SCANLAN, J.; KIRKBY, K. C. Increasing knowledge of mental illness through secondary research of electronic health records: opportunities and challenges. **Advances in Mental Health**, v. 14, n. 1, p. 14–25, 1 2016. ISSN 1838-7357. Disponível em: <<http://www.tandfonline.com/doi/full/10.1080/18387357.2015.1063635>>. Citado na página 17.

Strongloop. **LoopBack - Node.js framework**. 2017. Disponível em: <<http://www.webcitation.org/6fwAow3zf>>. Citado na página 36.

The Internet Society. **HTTP Over TLS**. [S.l.], 2000. Disponível em: <<https://tools.ietf.org/html/rfc2818http://www.webcitation.org/6qfxyii1I>>. Citado na página 39.

TRAPÉ, T. L.; CAMPOS, R. O. The mental health care model in Brazil: analyses of the funding, governance processes, and mechanisms of assessment. **Revista de saude publica**, Universidade de São Paulo. Faculdade de Saúde Pública., v. 51, n. 0, p. 19, 3 2017. ISSN 1518-8787. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/28355335>>. Citado na página 27.

VASSILIADIS, P.; SIMITSIS, A.; SKIADOPOULOS, S. Conceptual modeling for ETL processes. In: **Proceedings of the 5th ACM international workshop on Data Warehousing and OLAP - DOLAP '02**. New York, New York, USA: ACM Press, 2002. p. 14–21. ISBN 1581135904. Disponível em: <<http://portal.acm.org/citation.cfm?doid=583890.583893>>. Citado na página 35.

VEGODA, P. Introducing the IHE Concept. **Journal of Healthcare Information Management**, v. 16, n. 1, p. 3, 2002. Disponível em: <<http://rdcms-himss.s3.amazonaws.com/files/production/public/HIMSSorg/Content/files/jhim/15-4/column06.pdf>>. Citado na página 23.

VEST, J. R.; JASPERSON, J. What should we measure? Conceptualizing usage in health information exchange. **Journal of the American Medical Informatics Association : JAMIA**, The Oxford University Press, v. 17, n. 3, p. 302–7, 1 2010. ISSN 1527-974X. Disponível em: <<http://jamia.oxfordjournals.org/content/17/3/302.abstract>>. Citado na página 17.

WHETZEL, P. L.; NOY, N. F.; SHAH, N. H.; ALEXANDER, P. R.; NYULAS, C.; TUDORACHE, T.; MUSEN, M. A. BioPortal: enhanced functionality via new Web services from the National Center for Biomedical Ontology to access and use ontologies in software applications. **Nucleic acids research**, v. 39, n. Web Server issue, p. 541–5, 7 2011. ISSN 1362-4962. Disponível em: <<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3125807&tool=pmcentrez&rendertype=abstract>>. Citado na página 56.

World Health Organization. International Classification of Primary Care, Second edition (ICPC-2). **WHO**, World Health Organization, 2012. Disponível em: <<http://www.who.int/classifications/icd/adaptations/icpc2/en/http://www.webcitation.org/6qfAa0Yuc>>. Citado na página 38.

_____. **Mental health: a state of well-being**. World Health Organization, 2014. Disponível em: <http://www.who.int/features/factfiles/mental_health/en/http://www.webcitation.org/6qf9uFPJG>. Citado na página 16.

_____. **International Classification of Diseases**. World Health Organization, 2017. Disponível em: <<http://www.who.int/classifications/icd/en/http://www.webcitation.org/6qfB142cU>>. Citado na página 38.

_____. **International Classification of Functioning, Disability and Health (ICF)**. World Health Organization, 2017. Disponível em: <<http://www.who.int/classifications/icf/en/http://www.webcitation.org/6qfA3oDUU>>. Citado na página 38.

YOSHIURA, V. T.; AZEVEDO-MARQUES, J. M. de; RZEWUSKA, M.; VINCI, A. L. T.; SASSO, A. M.; MIYOSHI, N. S. B.; FUREGATO, A. R. F.; RIJO, R. P. C. L.; DEL-BEN, C. M.; ALVES, D. A web-based information system for a regional public mental healthcare service network in Brazil. **International Journal of Mental Health Systems**, BioMed Central, v. 11, n. 1, p. 1, 12 2017. ISSN 1752-4458. Disponível em: <<http://ijmhs.biomedcentral.com/articles/10.1186/s13033-016-0117-z>>. Citado 3 vezes nas páginas 17, 27 e 28.

YUKSEL, M.; GONUL, S.; ERTURKMEN, G. B. L.; SINACI, A. A.; INVERNIZZI, P.; FACCHINETTI, S.; MIGLIAVACCA, A.; BERGVALL, T.; DEPRAETERE, K.; ROO, J. D. An Interoperability Platform Enabling Reuse of Electronic Health Records for Signal Verification Studies. **BioMed Research International**, Hindawi Publishing Corporation, v. 2016, p. 1–18, 2016. ISSN 2314-6133. Disponível em: <<http://www.hindawi.com/journals/bmri/2016/6741418/>>. Citado na página 56.

TABELAS COMPLETAS COM ANÁLISES DE QUALIDADE

A.1 Análise de Completude

Na [Tabela 9](#) estão descritos em detalhes a completude de cada atributo do modelo de dados demográfico para, cada uma das fontes de dados, em percentuais. Na [Tabela 10](#) está descrita a melhoria de completude (em percentuais) também para cada campo em cada fonte de dados após realizada a integração de dados via o eHealth-Interop.

Tabela 9 – Análise de completude dos dados originais

Propriedade	SISAM	e-SUS AB SCE	CAIS-SR	HST	Barramento MS
Nome completo	100.0%	100.0%	100.0%	99.99%	100.0%
Nome completo da mãe	100.0%	99.63%	100.0%	90.06%	100.0%
Nome completo do pai	14.31%	0.0%	100.0%	84.76%	100.0%
Data de nascimento	94.5%	100.0%	100.0%	69.19%	100.0%
Município de nascimento	0.0%	100.0%	99.86%	47.83%	99.93%
Idade aproximada	20.05%	0.0%	100.0%	0.0%	0.0%
Sexo	100.0%	55.42%	100.0%	100.0%	100.0%
Raça/Cor	47.98%	100.0%	0.0%	33.59%	100.0%
Ocupação	35.91%	0.49%	0.0%	50.74%	0.0%
Nome social	0.0%	0.12%	0.0%	0.0%	1.45%
Estado civil	51.67%	0.0%	0.0%	50.99%	0.0%
Identificação do Paciente					
CNS	53.88%	54.57%	100.0%	16.62%	100.0%
CPF	27.85%	0.0%	99.28%	13.11%	87.59%
Orgão emissor do RG	0.0%	0.0%	0.0%	0.0%	73.54%

CNS	56.78% (2.9%)	54.57% (0.0%)	100.0% (0.0%)	17.12% (0.5%)	100.0% (0.0%)
CPF	42.46% (14.61%)	6.94% (6.94%)	98.54% (0.0%)	14.73% (1.62%)	89.14% (4.38%)
Orgão emissor do RG	26.16% (26.16%)	11.08% (11.08%)	87.43% (87.43%)	7.82% (7.82%)	71.12% (0.07%)
Número do RG	45.63% (11.06%)	11.21% (11.21%)	97.95% (0.29%)	19.16% (0.79%)	84.88% (13.83%)
Endereço de Residência					
Estado de resi- dência	79.15% (6.83%)	18.15% (18.15%)	97.95% (97.95%)	100.0% (0.0%)	99.91% (0.31%)
Município de re- sidência	77.41% (8.25%)	18.15% (18.15%)	97.95% (97.95%)	100.0% (29.56%)	99.89% (0.27%)
CEP de residên- cia	48.85% (25.65%)	17.9% (17.9%)	96.49% (96.49%)	35.57% (0.3%)	97.09% (3.21%)
Bairro	68.16% (12.48%)	18.03% (18.03%)	97.95% (97.95%)	38.81% (0.62%)	99.91% (0.29%)
Logradouro	72.01% (10.01%)	18.03% (18.03%)	97.95% (97.95%)	84.46% (1.04%)	99.89% (0.38%)
Número	53.72% (23.94%)	16.08% (16.08%)	97.66% (97.66%)	41.42% (0.37%)	98.24% (1.23%)
Complemento	10.63% (6.25%)	1.58% (1.58%)	7.02% (7.02%)	3.91% (1.7%)	19.0% (3.55%)
Informações de Contato					
Telefone resi- dencial	60.67% (5.7%)	4.99% (3.78%)	54.68% (54.68%)	28.2% (0.32%)	70.61% (35.09%)
Telefone comer- cial	1.47% (1.47%)	0.85% (0.85%)	1.46% (1.46%)	0.34% (0.34%)	3.98% (0.04%)
Celular	55.8% (5.78%)	3.41% (2.07%)	2.92% (2.92%)	30.37% (1.73%)	58.19% (33.22%)
E-mail	1.02% (0.1%)	0.12% (0.12%)	0.0% (0.0%)	0.08% (0.08%)	0.81% (0.6%)

Na [Tabela 11](#) estão descritas a completude para cada atributo demográfico no conjunto de dados que são integráveis. Nesta tabela também estão descritas a melhoria após a integração e a completude final dos dados já integrados.

Tabela 11 – Completude nos registros integráveis

Propriedade	Demográfico	Melhoria após Integração	Demográfico Integrado
Nome completo	100.0%	0.00%	100.0%
Nome completo da mãe	98.34%	1.33%	99.66%

Nome completo do pai	48.58%	41.08%	89.66%
Data de nascimento	98.75%	1.25%	100.0%
Município de nascimento	43.49%	54.27%	97.76%
Idade aproximada	13.61%	4.09%	17.69%
Sexo	99.91%	0.09%	100.0%
Raça/Cor	65.64%	29.79%	95.43%
Ocupação	52.56%	8.55%	61.1%
Nome social	0.0%	1.25%	1.25%
Estado civil	67.36%	6.29%	73.65%
Identificação do Paciente			
CNS	94.57%	3.83%	98.39%
CPF	66.55%	21.95%	88.5%
Orgão emissor do RG	0.0%	64.09%	64.09%
Número do RG	67.32%	21.59%	88.91%
Endereço de Residência			
Estado de residência	80.93%	18.53%	99.46%
Município de residência	78.61%	20.79%	99.4%
CEP de residência	52.39%	44.16%	96.55%
Bairro	70.85%	27.88%	98.73%
Logradouro	76.11%	22.94%	99.05%
Número	55.41%	42.05%	97.46%
Complemento	5.64%	13.70%	19.34%
Informações de Contato			
Telefone residencial	72.9%	12.28%	85.18%
Celular	58.79%	13.25%	72.04%
Telefone comercial	0.0%	3.28%	3.28%
E-mail	0.0%	0.00%	0.0%

A.2 Análise de Acurácia Sintática

O resultado da análise de acurácia, para cada atributo do modelo de informação e para cada fonte de dado está descrito na [Tabela 12](#). Na [Tabela 13](#) está descrito o percentual de melhoria na acurácia após a integração para cada fonte de dado.

Tabela 12 – Análise de acurácia

Propriedade	SISAM	eSUS-AB SCE	CAIS-SR	HST	Barramento MS
Nome completo	99.78%	100.00%	97.12%	99.06%	99.93%
Nome completo da mãe	93.29%	100.00%	88.20%	99.52%	99.27%
Nome completo do pai	93.48%	NA	85.04%	99.54%	86.07%
Data de nascimento	100.00%	100.00%	100.00%	97.22%	99.96%
Município de nascimento	NA	99.76%	34.01%	99.99%	99.70%
Idade aproximada	70.42%	NA	83.88%	NA	NA
Sexo	99.89%	100.00%	100.00%	100.00%	100.00%
Raça/Cor	100.00%	100.00%	NA	100.00%	84.79%
Ocupação	98.49%	100.00%	NA	97.82%	NA
Nome social	NA	0.00%	NA	NA	82.47%
Estado civil	100.00%	NA	NA	100.00%	NA
Identificação do Paciente					
CNS	100.00%	100.00%	100.00%	100.00%	100.00%
CPF	95.35%	NA	99.28%	85.83%	100.00%
Orgão emissor do RG	NA	NA	NA	NA	2.02%
Número do RG	98.72%	NA	99.42%	83.06%	99.39%
Endereço de Residência					
Estado de residência	100.00%	NA	NA	100.00%	100.00%
Município de residência	99.88%	NA	NA	100.00%	99.95%
CEP de residência	100.00%	NA	NA	95.44%	100.00%
Bairro	98.79%	NA	NA	99.46%	99.53%
Logradouro	68.27%	NA	NA	66.43%	95.30%
Número	96.95%	NA	NA	95.37%	97.45%
Complemento	90.14%	NA	NA	90.25%	80.42%
Informações de Contato					
Telefone residencial	100.00%	100.00%	NA	83.86%	100.00%
Telefone comercial	NA	NA	NA	NA	100.00%
Celular	99.98%	100.00%	NA	99.86%	100.00%
E-mail	NA	NA	NA	NA	NA

Tabela 13 – Melhoria da acurácia

Propriedade		SISAM	eSUS-AB SCE	CAIS-SR	HST	Barramento MS
Nome completo	com-pleto	99.71%(-0.07%)	100.00%(0.00%)	99.14%(2.06%)	99.29%(0.23%)	99.93%(-0.06%)
Nome completo da mãe	com-pleto da mãe	98.48%(5.19%)	100.00%(0.00%)	99.21%(9.44%)	99.16%(-0.36%)	99.77%(0.10%)
Nome completo do pai	com-pleto do pai	96.78%(3.31%)	73.65%(NA)	98.99%(13.02%)	99.33%(-0.21%)	99.38%(13.80%)
Data de nascimento		100.00%(0.00%)	100.00%(0.00%)	100.00%(0.00%)	98.37%(1.15%)	99.96%(0.01%)
Município de nascimento		79.09%(NA)	99.76%(0.00%)	79.09%(49.76%)	97.51%(-2.48%)	99.64%(-0.19%)
Idade aproximada	aproxi-mada	70.33%(-0.09%)	NA(NA)	69.88%(0.00%)	70.33%(NA)	78.48%(NA)
Sexo		100.00%(0.11%)	100.00%(0.00%)	100.00%(0.00%)	100.00%(0.00%)	100.00%(0.00%)
Raça/Cor		100.00%(0.00%)	100.00%(0.00%)	100.00%(NA)	100.00%(0.00%)	100.00%(15.61%)
Ocupação		98.54%(0.05%)	100.00%(0.00%)	100.00%(NA)	97.99%(0.17%)	98.58%(NA)
Nome social		0.00%(NA)	0.00%(0.00%)	0.00%(NA)	0.00%(NA)	88.89%(6.46%)
Estado civil		100.00%(0.00%)	100.00%(NA)	NA(NA)	100.00%(0.00%)	100.00%(NA)
Identificação do Paciente						
CNS		100.00%(0.00%)	100.00%(0.00%)	100.00%(0.00%)	100.00%(0.00%)	100.00%(0.00%)
CPF		95.80%(0.45%)	100.00%(NA)	98.81%(0.00%)	91.15%(5.32%)	99.22%(-0.78%)
Órgão emissor do RG		NA(NA)	0.00%(NA)	NA(NA)	NA(NA)	2.02%(0.13%)
Número do RG	do RG	98.76%(0.04%)	100.00%(NA)	99.10%(0.00%)	90.71%(7.65%)	99.17%(-0.13%)
Endereço de Residência						
Estado de residência		100.00%(0.00%)	100.00%(NA)	NA(NA)	100.00%(0.00%)	100.00%(0.00%)
Município de residência		99.88%(0.00%)	100.00%(NA)	NA(NA)	99.96%(-0.04%)	99.96%(0.01%)
CEP de residência		100.00%(0.00%)	100.00%(NA)	NA(NA)	96.48%(1.04%)	100.00%(0.00%)
Bairro		99.35%(0.57%)	100.00%(NA)	NA(NA)	99.42%(-0.04%)	99.53%(-0.06%)

Logradouro	68.53%(0.26%)	98.65%(NA)	NA(NA)	66.96%(0.53%)	95.27%(-0.26%)
Número	98.50%(1.55%)	98.48%(NA)	NA(NA)	96.20%(0.84%)	97.41%(0.11%)
Complemento	90.14%(0.00%)	80.00%(NA)	NA(NA)	90.20%(-0.05%)	88.54%(8.79%)
Informações de Contato					
Telefone residencial	100.00%(0.00%)	100.00%(0.00%)	100.00%(NA)	91.44%(7.58%)	99.82%(-0.18%)
Telefone comercial	NA(NA)	100.00%(NA)	NA(NA)	NA(NA)	100.00%(0.00%)
Celular	100.00%(0.02%)	100.00%(0.00%)	100.00%(NA)	99.92%(0.06%)	99.93%(-0.07%)
E-mail	NA	NA	NA	NA	NA

Na [Tabela 14](#) é descrita a acurácia no conjunto de registros demográficos que podem ser encontrados em mais de 1 fonte de informação. Assim como na análise de completude, a tabela demonstra a melhoria a partir da integração dos dados.

Tabela 14 – Acurácia nos registros integráveis

Propriedade	Demográfico	Melhoria após Integração	Demográfico Integrado
Nome completo	99.57	0.39	99.96
Nome completo da mãe	97.87	1.98	99.85
Nome completo do pai	97.40	2.08	99.48
Data de nascimento	100.00	0.00	100.00
Município de nascimento	89.79	5.58	95.36
Idade aproximada	55.42	13.25	68.67
Sexo	99.91	0.09	100.00
Raça/Cor	100.00	0.00	100.00
Ocupação	98.38	0.59	98.96
Nome social	nan	nan	85.07
Estado civil	100.00	0.00	100.00
Identificação do Paciente			
CNS	100.00	0.00	100.00
CPF	95.65	3.06	98.71
Orgão emissor do RG	nan	nan	100.00

Número do RG	98.23	0.26	98.49
Endereço de Residência			
Estado de residência	100.00	0.00	100.00
Município de residência	99.93	0.01	99.94
CEP de residência	99.89	0.03	99.92
Bairro	99.50	0.44	99.94
Logradouro	83.28	3.34	86.62
Número	98.52	0.56	99.08
Complemento	87.75	1.73	89.48
Informações de Contato			
Telefone residencial	99.82	-0.04	99.78
Celular	99.90	0.02	99.92
Telefone comercial	nan	nan	100.00
E-mail	nan	nan	nan

A.3 Análise de Consistência

A análise de consistência está descrita na [Tabela 15](#) e a melhoria correspondente após a sua incorporação no eHealth-Interop está descrita na [Tabela 16](#)

Tabela 15 – Análise de consistência

Propriedade	SISAM	e-SUS AB SCE	CAIS-SR	HST	Barramento MS
Nome completo	67.99%	57.05%	59.79%	82.28%	72.60%
Nome completo da mãe	51.24%	44.97%	28.54%	60.14%	54.36%
Nome completo do pai	51.59%	NA	29.28%	46.39%	28.14%
Data de nascimento	0.04%	0.00%	0.15%	0.03%	0.03%
Município de nascimento	NA	51.68%	0.00%	59.82%	30.00%
Idade aproximada	0.00%	NA	0.00%	NA	NA
Sexo	76.35%	0.34%	98.37%	0.01%	58.71%
Raça/Cor	0.00%	66.44%	NA	30.95%	15.22%
Ocupação	6.82%	0.00%	NA	5.14%	NA
Nome social	NA	NA	NA	NA	0.00%
Estado civil	0.00%	NA	NA	0.00%	NA
Identificação do Paciente					
CNS	72.84%	100.00%	44.96%	65.59%	68.01%
CPF	96.52%	NA	100.00%	93.47%	96.85%

Orgão emissor do RG	NA	NA	NA	NA	0.00%
Número do RG	48.19%	NA	76.14%	45.42%	43.24%
Endereço de Residência					
Estado de residência	70.21%	NA	NA	0.00%	42.01%
Município de residência	44.91%	NA	NA	70.07%	47.07%
CEP de residência	0.00%	NA	NA	41.39%	20.91%
Bairro	28.23%	NA	NA	24.49%	17.13%
Logradouro	4.45%	NA	NA	3.74%	1.56%
Número	49.90%	NA	NA	41.15%	27.77%
Complemento	4.58%	NA	NA	2.05%	1.23%
Informações de Contato					
Telefone residencial	0.05%	0.00%	NA	0.00%	0.07%
Celular	0.00%	0.00%	NA	0.00%	0.00%
Telefone comercial	NA	NA	NA	NA	0.00%
E-mail	0.00%	NA	NA	NA	0.00%

Tabela 16 – Melhoria na consistência

Propriedade	SISAM	e-SUS SCE	AB	CAIS-SR	HST	Barramento MS
Nome completo	79.45%(11.47%)	88.19%(28.40%)	85.97%(3.70%)	66.44%(9.40%)	82.08%(9.48%)	
Nome completo da mãe	62.12%(10.88%)	80.77%(52.23%)	70.50%(10.36%)	49.33%(4.36%)	66.25%(11.89%)	
Nome completo do pai	65.83%(14.24%)	79.13%(49.85%)	55.05%(8.66%)	NA(NA)	86.02%(57.88%)	
Data de nascimento	88.57%(88.53%)	96.37%(96.22%)	89.84%(89.82%)	0.00%(0.00%)	91.03%(91.01%)	
Município de nascimento	NA(NA)	0.00%(0.00%)	59.84%(0.01%)	73.49%(21.81%)	76.14%(46.14%)	
Idade aproximada	0.00%(0.00%)	0.00%(0.00%)	NA(NA)	NA(NA)	62.38%(NA)	
Sexo	98.41%(22.06%)	98.91%(0.54%)	98.83%(98.82%)	97.99%(97.65%)	98.36%(39.66%)	
Raça/Cor	54.66%(54.66%)	NA(NA)	43.48%(12.53%)	66.78%(0.34%)	73.71%(58.49%)	

Ocupação	7.56%(0.74%)	NA(NA)	5.58%(0.44%)	0.00%(0.00%)	89.44%(NA)
Nome social	NA(NA)	NA(NA)	NA(NA)	NA(NA)	100.00%(100.00%)
Estado civil	20.56%(20.56%)	NA(NA)	16.76%(16.76%)	NA(NA)	95.05%(NA)
Identificação do Paciente					
CNS	72.84%(0.00%)	44.96%(0.00%)	65.59%(0.00%)	100.00%(0.00%)	60.89%(-7.12%)
CPF	96.52%(0.00%)	100.00%(0.00%)	93.47%(0.00%)	NA(NA)	97.54%(0.69%)
Orgão emissor do RG	NA(NA)	NA(NA)	NA(NA)	NA(NA)	99.88%(99.88%)
Número do RG	48.19%(0.00%)	76.23%(0.09%)	45.42%(0.00%)	NA(NA)	78.23%(34.99%)
Endereço de Residência					
Estado de residência	70.21%(0.00%)	NA(NA)	0.00%(0.00%)	NA(NA)	60.51%(18.50%)
Município de residência	79.86%(34.95%)	NA(NA)	79.09%(9.02%)	NA(NA)	82.50%(35.43%)
CEP de residência	61.92%(61.92%)	NA(NA)	45.95%(4.56%)	NA(NA)	77.70%(56.79%)
Bairro	32.20%(3.97%)	NA(NA)	27.90%(3.40%)	NA(NA)	46.68%(29.55%)
Logradouro	5.53%(1.08%)	NA(NA)	4.69%(0.95%)	NA(NA)	24.41%(22.84%)
Número	50.14%(0.24%)	NA(NA)	41.15%(0.00%)	NA(NA)	67.94%(40.17%)
Complemento	4.58%(0.00%)	NA(NA)	2.05%(0.00%)	NA(NA)	86.32%(85.09%)
Informações de Contato					
Telefone residencial	11.03%(10.98%)	NA(NA)	9.61%(9.61%)	0.00%(0.00%)	64.93%(64.86%)
Celular	6.68%(6.68%)	NA(NA)	3.32%(3.32%)	0.00%(0.00%)	70.53%(70.53%)
Telefone comercial	NA(NA)	NA(NA)	NA(NA)	NA(NA)	100.00%(100.00%)
E-mail	NA(NA)	NA(NA)	NA(NA)	NA(NA)	NA(NA)

PÔSTER PARA WHO-FIC NETWORK MEETING 2017

A Organização Mundial de Saúde estabeleceu um conjunto de Centros Colaboradores em diversos países com o objetivo de contribuir com a Família de Classificações Internacionais da OMS. A este grupo foi dado o nome de *WHO-FIC (World Health Organization - Family of International Classifications) Network*.

Anualmente a WHO-FIC organiza um evento denominado WHO-FIC Network Meeting, onde recebe trabalhos de diversos grupos associados ao uso e disseminação da Família de Classificações Internacionais da OMS.

No ano de 2017, o evento ocorreu dos dias 16 à 21 de outubro, na Cidade do México (México)¹. Como parte do resultado do trabalho da presente tese, foi apresentado um pôster denominado "Health Terminology Server for Portuguese-speaking Communities". Neste pôster, a arquitetura do eHealth-Interop Server foi apresentada junto com uma proposta para agregar tecnologias da Web Semântica.

¹ <http://www.who.int/classifications/network/meeting2017/en/>

M É X I C O

Health Terminology Server for Portuguese-speaking Communities

16-21 October 2017
Mexico City, Mexico

Authors: Domingos Alves¹, Newton S. B. Miyoshi¹, João M. A. Marques¹, Paulo M. A. Marques¹, Rui P. C. L. Rijo²

C710

1 Ribeirão Preto Medical School of the University of São Paulo, Brazil

2 Polytechnic Institute of Leiria, Portugal

Abstract

In this poster we show the design and development of a clinical terminology server based on health information standards. It has two main characteristics: robust and functional as an independent piece of software that can serve as a framework of a knowledge portal; flexible to act as a component and can be easily integrated as part of a legacy information system.

Introduction

This poster presents an international research project that involves the creation, management, diffusion and application of health knowledge in the areas of science and information and decision-making technologies. In particular, this initiative aims to coordinate and make feasible a terminology service in international health capable of establishing a convergence and automated terminology equivalence among standard terminologies in health to be used by Portuguese-speaking countries.

This is done from the specification and development of a clinical terminology server based on health information standards. Its main feature is its adaptability, which allows the service to be easily incorporated as a component of other health information systems. In addition, it also serves as the basis for a knowledge portal for terminologies.

Methods & Materials

To do this, a set of methodologies is being developed that is being developed by the groups of the Health Information and Informatics Center of the Ribeirão Preto Medical School of the University of São Paulo (CIIS-FMRP / USP) and the Center for Research in Technologies and Health Services (CINTESIS - FM-UP, EE-UP and IPEleiria), which establishes the criteria for small-scale analysis, convergence and validation tests of the terminological study process to avoid large-scale conceptual deviations. Thus, these institutions can work together through the Terminology Service, offering several products and tools, such as easy access to terminologies, mappings and convergences between them and the possibility of extracting knowledge by other information systems from marking content semantics.

For this service, the terminologies defined by the International Classification Family of the World Health Organization and associated ones, namely: ICD-10 and 11, CIF and CIAP-2, shall be taken into account, as well as references to terminology SNOMED-CT.

The architecture of the Terminology Service has the following main components (Figure 1):

(i) Terminology Database: a document-based database for storing all terminology and concept data. MongoDB is used as the system manager ;

(ii) Base models: set of classes that represent the main data models that are used to manage terminologies, ontologies, concepts, value sets and mappings;

(iii) Mapping Module: responsible for the high level functionalities related to the management of mappings between concepts of different terminologies. This module provides a tool for mapping between terminologies, in the which a health specialist counts on the use of data mining techniques, more specifically mining of association rules, and textual search as an aid in defining the correct mapping.;

(iv) Knowledge Module: responsible for high level functionality to research concepts, terminologies and sets of values;

(v) RESTful web services are a set of features (APIs) that are exposed to client applications over the HTTP protocol. From these services, external systems can consume and provide data, respecting the security policies implemented by the Terminology Service.

(vi) SPARQL endpoint: the SPARQL endpoint allows queries to be made based on terminology using the SPARQL language. As previously explained, in this query language it is possible to dynamically relate concepts and manipulate data in RDF format;

(vii) Ontologies: set of formal representations (metadata) to describe the main concepts involving the base models for managing the terminology server;

(viii) Semantic Module: module responsible for the semantic manipulation of the terminologies, encodings and ontologies managed by the server. Through this module it is also possible to search using distance and concept matching functions

The eHealth Interop Terminology Server is built using the Loopback 3 framework software library. For the construction of the Semantic Module and the SPARQL endpoint, the use of the D2RQ platform will be considered.

Conclusions

In this service, the terminologies are formalized using ontologies. The main effort is to build an ontology mapping to assure interoperability between those ontologies. With these ontologies and the ontology mapping it's possible to markup semantically the classifications and, in this way, gather data with a unique meaning to support accurate information. Another key advantage is the possibility to offer the classified data in any target ontology, allowing each institution to use their own terminology.

The terminologies, described as formal ontologies will be available through the terminologies service including the reference ontologies and the ontologies mapping. We also intend to enable semantic querying using a SPARQL endpoint and implement algorithms to help the process of semantic markup.

Finally, the whole effort of this joint initiative will serve as reference to support research and development in the area of Science and Information Technology of each country, as well as the intention to formalize the nucleation of an international reference center, as a Collaborating Center of the Family of Classifications (WHO-FIC) in the Portuguese classification.

Figure 1: Proposed architecture for terminology server

ARTIGO PARA O HCIST 2017

O HCist - *International Conference on Health and Social Care Information Systems and Technologies*¹ - é um congresso científico na área de tecnologias de informação e sistemas de informação em saúde.

Foi submetido um artigo denominado "Development of a Lightweight and Adaptable Clinical Terminology Server" com parte dos resultados desta tese. Neste artigo está descrito o servidor de terminologias clínicas que é parte integrante do eHealth-Interop.

¹ <http://hcist.scika.org/>

Development of a Lightweight and Adaptable Clinical Terminology Server

Newton Shydeo Brandão Miyoshi, João Mazzoncini de Azevedo Marques,
Domingos Alves, Paulo Mazzoncini de Azevedo Marques

Ribeirão Preto Medical School, University of São Paulo, Ribeirão Preto, São Paulo, Brazil

Abstract

Health terminologies, vocabularies and ontologies are increasingly becoming indispensable computational artifacts since they allow the standardization and conceptualization of medical vocabulary. Its applications vary from the cleaning and improvement of data quality, learning, and serving as a basis for health information exchange and up to improve data analysis and decision making. In this context an important aspect is the software component responsible for storing and managing these linguistic artifacts often named as terminology server, ontology repository, metadata repository, among other names. In this paper we present the eHealth Interop Terminology Server, a clinical terminology server based on health information standards. The main characteristics is its adaptability and can be easily implemented as a component within a health information system as well as serve as the basis for a terminological portal. It uses service oriented architecture so all the communication is made using a web services API. We also perform a functional test with clinical terminologies in the context of a mental health care network.

Keywords: clinical terminologies; ontologies; web services; health information exchange; concept matching.

1. Introduction

The application of information technologies in the health area is increasingly an essential tool for both rapid resolution and coordination of various processes and for analysis and decision making in individual and collective care. A central aspect in this context is the standardization of terms and concepts and this is due to three main factors: (i) the highly heterogeneous and sometimes ambiguous nature of medical language and its constant evolution; (ii) an enormous amount of data that is generated constant by the automation of processes and from new technologies that arise and (iii) the need to process, analyze and make decisions based on this information¹.

Over the years several artifacts have been created and formalized and standardized medical knowledge such as: terminologies, codifications, controlled vocabularies, taxonomies, thesaurus and ontologies² as well as different formats for the formal representation of these artifacts. In this context, the terminology and ontology server is the piece of software responsible for storing and managing these computational linguistic artifacts. The terminology and ontology server acts as a terminology repository and can add other functions and auxiliary tools to search terms and concepts.

The objective of this research is the design and development of a clinical terminology server based on health information standards and has two main characteristics: (i) it is robust and functional as an independent piece of software that can serve as the basis for the construction of A knowledge portal on a specific topic; (ii) is flexible to act as a software component and can be easily integrated as part of an existing health information system.

2. Methods

To allow the development of a clinical terminology server that can be deployed in several environments we choose a service-based architecture. The service-oriented computing paradigm allows applications to be quickly built and can be easily composed in a distributed and heterogeneous environment³. In this context the terminology server named eHealth-Interop Terminology Server follows a client-server architecture in which all functionality is made available in a set of restful web services. As health information standards, the Common Terminologies Services version 2⁴, the HL7 FHIR Terminology Service⁵ and the IHE Sharing Value Sets Profile⁶ were used.

2.1. Use Case

Although Brazil has a Unified Health System for all the population, the services organization and care delivery is done in a regional context in health care networks. In the region of Ribeirão Preto, in state of São Paulo, the mental health care network comprises about 26 municipalities and approximately 1,400,000 inhabitants. An ongoing project aims to establish a data interoperability platform to support care coordination in this region. It was defined a set of information models to be exchanged in this context. We use the solution presented here as the terminology server that will be the basis for the standardization of mental health care information.

3. Results

A clinical terminology server called eHealth-Interop Terminology Server was designed and developed. This server acts as the Value Set Repository actor defined in the IHE SVS profile as well as the terminology module of the HL7 FHIR. In the next subsections we will detail the main components that comprises the server architecture, the actors and operations that are made available by the API

Restful developed, and the use case of using clinical terminologies to support the mental health care network.

3.1. Software Architecture

The eHealth-Interop Terminology Server architecture has the following main components: Terminology Database, a document-based database to store all terminologies and concepts data, the MongoDB⁷ was used as the database engine; Base Models, set of classes that represent the main data models used to manage terminologies, ontologies, concepts, value sets and mappings; Mapping Module is responsible for the high-level functionalities related to the management of mappings between concepts of different terminologies; Knowledge Module is responsible for high level functionalities for searching concepts, terminologies and value sets; the RESTful Web services is the set of functionalities (API) that are exposed to client applications through the HTTP protocol, from these web services it is possible to build APIs specific to a particular programming language such as Java, among others. The eHealth Interop Terminology Server was built using the Loopback 3 framework⁸.

Following the recommendation of the IHE SVS Profile and HL7 FHIR, the authentication and security features are optional but are represented in the following components: Auth DB, database for user registration, roles and client applications, the MySQL relational database was used as database engine; Auth Models, are the classes that allow the management of the data model stored in Auth DB and finally; Auth Module, which has the business logic for managing the authentication and authorization features.

3.2. RESTful Web Services Operations

The eHealth-Interop Terminology Server API is the communications layer with the client applications, in which the main functionalities are available. The main actors involved are: Terminology Server, Terminology Consumer and Terminology Administrator. The Terminology Server actor, which is the solution described in this paper, acts as the Value Set Repository actor defined by the IHE SVS Profile. Terminology Consumer is a representation of a client application and corresponds to the Value Set Consumer defined by the IHE SVS Profile. Terminology Administrator is an important actor responsible for some special restricted operations for managing all the data in the Terminology Server and does not have a correspondent actor in IHE SVS profile.

3.3. Mental Health Care Network – Use Case

In order to support the information model defined for care coordination in the mental health care network of Ribeirão Preto region, we used 3 international terminologies and 2 national classifications. The international terminologies are: ICD-10 - International Statistical Classification of Diseases and

Related Health Problems, ICF - International Classification of Functioning, Disability and Health, ICPC-2 - International Classification of Primary Care. The national classifications are the Table of Procedures, Medications and Orthotics, Prostheses and Materials of the Unified Health System in Brazil and the Brazilian Classification of Occupations.

In addition, local terminologies were created to represent different concepts such as: sex, marital status, sexual orientation, among others. These codifications, for the most part, are in agreement with those used in the National Information System for primary care or based on information models from national standard for discharge summary.

All these terminologies were instantiated into the eHealth-Interop Terminology Services and their concepts and terms were imported using a client application acting as an “Terminology Admin” actor. This client application parsed the different formats and send them using the web services API.

4. Conclusion and Future Work

In this paper we present the eHealth-Interop Terminology Server as a software component that can be adapted to several health contexts and applications. The use of the service-oriented paradigm through the Loopback framework has proved to be a quick and effective way to develop a solution adaptable to different scenarios. The document-oriented database as the terminology database proved to be a good solution because it allowed the storage of particular properties of each terminology without losing the basic structure of the information models.

References

1. Konopka BM. Biomedical ontologies—A review. *Biocybern Biomed Eng.* 2015;35(2):75-86. doi:10.1016/j.bbe.2014.06.002.
2. Bodenreider O. Biomedical ontologies in action: role in knowledge management, data integration and decision support. *Yearb Med Inform.* 2008;67-79. <http://www.ncbi.nlm.nih.gov/pubmed/18660879>. Accessed May 12, 2017.
3. Papazoglou MP, Traverso P, Dustdar S, Leymann F. Service-Oriented Computing: A Research Roadmap. *Int J Coop Inf Syst.* 2008;17(2):223-255. doi:10.1142/S0218843008001816.
4. Object Management Group. Common Terminology Services 2 (CTS2) Version 1.2.; 2015. <http://www.omg.org/spec/CTS2/1.2/>. Accessed March 10, 2017.
5. HL7. Terminology Service - FHIR Release 3 (STU). <https://www.hl7.org/fhir/terminology-service.html>. Published 2017. Accessed May 2, 2017.
6. IHE. IHE IT Infrastructure (ITI) Technical Framework Volume 1 Integration Profiles. Vol 13.; 2016.
7. MongoDB. MongoDB for GIANT Ideas. <http://www.webcitation.org/6qZIFWE1o>. Published 2017. Accessed May 3, 2017.
8. Strongloop. LoopBack - Node.js framework. <http://www.webcitation.org/6fwAow3zf>. Published 2017. Accessed May 3, 2017.

ARTIGO PARA O JOURNAL OF MEDICAL INTERNET RESEARCH

O *Journal of Medical Internet Research*¹ (JMIR; ISSN 1438-8871, Abreviação no Medline: J Med Internet Res) teve início em 1999 e é uma revista científica *open-access* na área de informática em saúde e políticas de serviços de saúde. JMIR é indexada em mais de 18 bancos de dados bibliográficos incluindo Medline, CINAHL, Scopus, Pubmed Central, entre outros. O fator de impacto do JMIR, calculado pela Clarivate Analytics (antiga Thomson Reuters) para 2016, foi de 5.175 e o ferramenta de classificação de artigos Scimago classificou o JMIR como Q1 na área de Informática em Saúde.

Foi escrito um artigo denominado "eHealth-interop: an Interoperability Platform developed for Mental Care Integration" que foi submetido a uma edição especial de Computação aplicada à Saúde Mental do JMIR. Neste artigo os principais resultados da pesquisa são apresentados: a plataforma de interoperabilidade eHealth-Interop e o caso de uso para saúde mental. O artigo está na fase de revisão pelos pares.

¹ <http://jmir.org>

eHealth-interop: an Interoperability Platform developed for Mental Care Integration

Newton Shydeo Brandão Miyoshi, João Mazzoncini de Azevedo Marques, Domingos Alves,
Paulo Mazzoncini de Azevedo Marques

Ribeirão Preto Medical School, University of São Paulo, Ribeirão Preto, SP, Brazil

Abstract

Background:

The electronic exchange of health-related data can support different professionals and services to act in more coordinated and transparent manner and make the management of health services networks more efficient. Although mental health care is one of the areas that can benefit from a secure health information exchange, since it usually involves long-term and multiprofessional care, there are few published studies on this topic, particularly in low and middle income countries.

Objective: The design, implementation and evaluation of an eHealth platform that allows the technical and informational support of a Brazilian regional network of mental health care. This solution enables the health information exchange, improve data quality and identify and monitor patients over time and in different services.

Methods: The proposed platform is based on client-server architecture to be deployed in the web following a web-services communication model. The interoperability information model was based on international and Brazilian health standards. To test platform usage, we have utilized the case of the mental health care network of the XIII Regional Health Department of the São Paulo state, Brazil. Data were extracted from five different sources, involving 26 municipalities: national demographic data, data from primary health care, data from requests for psychiatric hospitalizations performed by community services, and data obtained with psychiatric hospitals about hospitalizations. Data quality metrics such as accuracy and completeness were evaluated to test the proposed solution.

Results: The eHealth-Interop integration platform was designed, developed and tested. It contains a built-in terminology server and a record linkage module to support patient's identification and deduplication. The proposed interoperability environment was able to integrate information in the mental health care network case with the support of five international and national terminologies. In total, 27353 records containing demographic and clinical data were integrated into eHealth-Interop. Of all these records, 34% (9547) were identified as patients who were present in more than 1 data source with different levels of accuracy and completeness. In general, it was possible to achieve an improvement of 18,4% in completeness and 1,1% in syntactic accuracy over the test data set after integration and deduplication.

Conclusions: The proposed platform established an eHealth solution to fill the gap in the availability and quality of information within a network of health services, in order to improve the continuity of care and the health services management. It has been successfully applied in the context of mental health care and is flexible to be tested in other areas of care.

Key-words (mesh terms): eHealth, mental health; mental health services; health information exchange; health information interoperability; record linkage; continuity of patient care; integrated care;

Introduction

Background

The care of mental disorders frequently require a coordination of efforts between different healthcare professionals and services and attention levels (THORNICROFT and TANSELLA, 2013). There is also a frequent need to integrate the care of mental disorders with the care of non-psychiatric health problems (such as cardiovascular disorders, diabetes, maternal health, HIV / AIDS and cancer) in the same people, since this type of comorbidity is common and there is evidence that an adequate management of behavioral problems can positively impact other health conditions (PATEL et al, 2013)(BOWER et al, 2014).

However, all around the world, an adequate continuity of care - defined as the coordination and integration of different health care events to meet patient care needs in a coherent, connected and consistent manner (HAGGERTY, 2003) - is commonly affected by the lack of consistent information exchange between different services and attention levels (DURBIN et al, 2012).

An extensive computer science research area that seeks to assist in the problem of fragmentation of healthcare information systems is interoperability and data integration. Interoperability is defined by the HIMSS (Healthcare Information and Management Systems Society) as the ability of different information systems and applications to communicate, exchange data and transparently use such data and associated information (HIMSS, 2010).

The exchange of health information relating to the same patient between different attention levels and different health institutions is usually non-existent or incomplete. This can affect the continuity of care and communication between different care providers. Due, in large part, to the fact that health information systems that offer support to physicians and other health professionals are not integrated (BODENHEIMER, T. 2008).

Coffey et al. (COFFEY et al., 2008) point out that the fragmentation of services undermines mental health care by creating a barrier between different providers and the patient, making it difficult to provide effective treatment. According to Spiranovic et al. (SPIRANOVIC, C. et al., 2016), poor data quality, inconsistencies / discrepancies in information, and lack of interoperability among health information systems are one of the main difficulties for the secondary use of electronic health record data including the public health management.

Brazil is a middle-income country with the world's fifth largest population and, since 1988, has been implementing a complex national health system - highly decentralized, with responsibilities in the funding and management for the three federated levels (cities, states and country) and composed of a mix of public and private, profit and nonprofit services (PAIM et al, 2011). As in other countries, the prevalence of mental disorders is high and the

comorbidity between mental disorders and non-psychiatric health problems is common (RZEWUSKA et al, 2017). And, as the regional integration between services is still limited and extremely challenging, there is an urgent need to develop new, feasible integration strategies for mental health services too (GREGORIO et al 2012).

Objective

Considering the need for a computational environment to facilitate the exchange of health information to improve the continuity of care between attention levels, the eHealth-interop computing platform was designed and deployed. The main objective of the eHealth-Interop platform is to provide a service for the exchange of standardized information on mental health, promoting the improvement of the quality of information and its availability in a transparent way for all the regional services responsible for mental health care.

Methods

Context of the study: The Mental Health Care Network (MHCN) of the XIII Regional Health Department (XIII RHD) of the Sao Paulo State, Brazil

In Brazil, public mental health care should be carried out through regional public mental health care networks, whose objectives are to guarantee the continuity and integrality of care through the different services and attention levels (TRAPE and CAMPOS, 2017) .

The XIII RHD covers over 26 municipalities (SECRETARIA DE SAÚDE DO ESTADO DE SÃO PAULO, 2017) with approximately 1,400,000 inhabitants (INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA, 2017), where Ribeirão Preto is the main city. The development and current features of its MHCN were described in detail elsewhere (YOSHIURA et al, 2017; BARROS et al 2016; BARROS et al 2010). Briefly, it comprises: 1) primary health care, community-based mental health and emergency services managed by the 26 municipalities; 2) Acute inpatient psychiatric beds, outpatient specialized clinics and a day-hospital located at two state psychiatric hospitals and one state university general hospital; 3) Long-term residential mental health care facilities managed by municipalities and the state.

Due to the strong decentralization and little integration between services of the public health system in Brazil, ordinarily each of them has its own information systems. Only for the primary health care services there is a national computerized information system called e-SUS AB (DEPARTAMENTO DE ATENÇÃO À SAÚDE, 2016), that still in an initial phase of implementation around the country. The psychiatric referral and counter-referral in the region are articulated at the regional level through an information system called SISAM (Sistema de Informação em Saúde Mental - in english Mental Health Information System) (YOSHIURA V et al, 2017), which has a module for appointments scheduling in community mental health care units too.

Although the different services present in the network has mechanisms to integrate care between services, as recommended by Brazilian health legislation (such as the SISAM itself and periodic meetings between services representatives) (Barros et al 2016, Yoshiura, 2017), from the informational and computational point of view, support systems of the different services are not integrated. So information loss, data duplicity, lack of consistency and rework are common. Therefore, in this work, it was considered the integration of 5

information sources from different information systems that support MHCN in the various levels of attention: SISAM (responsible for the control of all hospitalization requests of the community services and discharges of the specialized hospital services); e-SUS AB (an information system provided by the Ministry of Health for primary health care facilities); data from two state psychiatric hospitals; and access to the Service Bus of the Brazilian Ministry of Health for integration with patients demographic data through the National Health Card.

Interoperability Information Model

In the eHealth-Interop platform, the interoperability data model corresponds to the information set that is exchanged and made available to health professionals at different attention levels. To define this set of information we analyzed the main processes involved in patient care. Four different events were highlighted: (i) patient entry into the network care; (ii) Assessment in primary care; (iii) The admission in psychiatric hospital service and (iv) the discharge of these services. For each process, a set of data elements were defined from the analysis of already existing documents. Finally these proposed data elements were reviewed by mental health specialists, for the addition or removal of some information (Figure 1).

The documents analyzed include records used in MHCN logistic instruments, national standards and international documents. The records used are as follows: SISAM Demographic Form and the Hospital Discharge Summary from a regional mental health specialized hospital. Furthermore, the Minimal Health Data Set developed by the Brazilian Ministry of Health was analyzed (SECRETARIA DE ATENÇÃO À SAÚDE, 2015), together the Discharge Summary proposed by the ABNT (Associação Brasileira de Normas Técnicas - in english Brazilian Association of Technical Standards) technical group on health informatics (ABNT CEE-78, 2016).

Besides this set of national documents, we have also analyzed documents from the UK, Canada and Australia. From Australia the following documents were analyzed: National Minimum Data Set (AUSTRALIAN INSTITUTE OF HEALTH AND WELFARE, 2015a), National Minimum Data Set (AUSTRALIAN INSTITUTE OF HEALTH AND WELFARE, 2015b), National Minimum Data Set - Community Mental Health Care (AUSTRALIAN INSTITUTE OF HEALTH AND WELFARE, 2015c). From the United Kingdom, the Mental Health and Learning Disabilities Data Set (MHLDDS) was analyzed and the document analyzed from Canada was the Common Data Set - Mental Health (MINISTRY OF HEALTH AND LONG TERM CARE, 2013). We also reviewed the information highlighted in a paper describing minimum data to continuity of care (DELFINI, et al. 2015).

Figure 1. Process generating the Interoperability Information Model

As a result of the described process, a set of information was defined for each of the four events: Demographic Form, with patient demographic and identification information; Primary Care Assessment, with clinical information from primary care; Admission Form, with information of the hospitalization request and the admission moment and the Discharge Summary, with data of the whole hospitalization process. These information sets were reviewed, reformulated and approved by a group of around 20 professionals, representative of the several health services and municipalities involved in MHCN.

Interoperability

The defined interoperability data models serve as a conceptual basis for the health information exchange. The eHealth-Interop platform acts as a central repository of the data exchanged. All communication is performed through web services available at the eHealth-Interop. Each data source, which is the information systems in the different health services, sends and queries data from its patients through web service clients that communicate with eHealth-Interop.

The process of sending patient data can be compared with the Extraction-Transformation-Load (ETL) steps in a classical data integration process. First the data of interest is extracted from the local databases, in the format that is own of each information system. Then, the data is transformed to conform to the interoperability data model previously defined. This transformed data is then uploaded to eHealth-Interop, it is sent using specific endpoints provided by the platform web services API. Then the data undergo a syntactic and duplicity analysis, aiming to improve the quality of the information. If the patient's record is already recorded in eHealth-Interop, the new information will be aggregated making a more complete record.

Data Quality Analysis

The analysis of the integrated final data set obtained with the proposed interoperability solution was performed through the measurement of data quality metrics. The measured data quality metrics are as follows:

- **Completeness:** represents the degree to which a given collection of data has the data it should represent. Corresponds to missing values in a collection of data of interest;
- **Syntactic Accuracy:** represents the degree to which a given data is correct and reliable. Syntactic accuracy measures the degree to which a given data corresponds to a possible value in the data set, for example, age must be between 0 and 120 years;
- **Duplicity:** represents the degree to which a given data is unnecessarily duplicated within a data set of interest, e.g. the two or more records belonging to a same record in the same data source;

The data quality metrics of completeness and syntactic accuracy were calculated considering the data sources in three different moments during the data interoperability process. In the first moment, the data is available based on as it is originally represented in the different data sources. In a second moment the data is available considering the normalization process into the data source adapters. During this process, each value of each attribute is transformed (if necessary) to conform to the defined interoperability data model. In a third moment, different records belonging to the same patient, which are originally in the different sources, are integrated into a single record and then the quality metrics are applied over them.

The duplication analysis was performed to verify the use of the record linkage algorithms and was performed using a set of records in the second stage that did not have deterministic identifiers.

Results

eHealth-Interop Architecture

The architecture of a computational system is a metaphor for the architecture of a physical construction (PERRY and WOLF, 1992) because it seeks to define the general structures such as its main elements, the relationships between these elements and the properties of the elements and these relationships (CLEMENTS et al., 2010). According to Soni et al (SONI; NORD; HOFMEISTER, 1995) the conceptual architecture of a system defines a high level structure and is independent of the implementation and technical decisions.

Taking into account that regional care delivery is made up of different actors located at different points in a given region, the web environment was defined as the basis for the integration platform, following a client-server architecture. The proposed conceptual architecture of eHealth-Interop is shown in Figure 2. The architecture follows a model with 3 main layers: Data Layer, Semantic Layer and Communication Layer. The Data Layer is where the data to be integrated are represented and stored; Semantic Layer, in which the mediating schemas, terminologies, ontologies, local codifications and mappings between them are described; Finally, the Communication Layer, where security aspects and communication with the data sources are managed.

Figure 2. Conceptual architecture of eHealth-Interop

Each layer has specific components. Functionalities of each component are described below.

In the Data Layer, data are stored following a dual model, ensuring the differentiation between demographic and clinical information. This distinction seeks to establish a solid foundation for patient identification algorithms and at the same time make the database flexible to be extended to different application contexts in the medical field (Beale et al., 2008). Another important component of the Data Layer is the Record Linkage (RL) module. This module is responsible for ensuring that there is no duplication of patient registration through the information provided in the Demographic Database.

The Semantic Layer is mainly composed by a terminology server, responsible for the storage of codifications, terminologies and ontologies, as well as their relationships. It also comprises the clinical and sociodemographic concepts that compose the information models and the schema management module that connects the information model with the terminologies and communicates with databases.

The Communication Layer is responsible for generating the Service API from the information models represented in the Semantic Layer and for promote a secure communication by the Security Module that verifies the authentication and authorization of the person responsible for the process of data sending and retrieval. The exchange of information from the data source to the interoperability platform is done through adapters that are software agents that translate the source schema into the global model provided by the service interface.

The implementation of this conceptual architecture was made using the LoopBack framework (STRONGLOOP, 2017). As database we have used MySQL (ORACLE, 2018) to store demographic information and MongoDB (MONGODB, 2017) for the clinical information.

Terminology Server

Medical vocabulary is highly heterogeneous and several information artifacts are constantly emerging aiming to standardize not only medical concepts but relationships between them. These artifacts are usually named codifications, terminologies, controlled vocabularies, ontologies, among others (BODENREIDER, 2010). The terminology server is a software component that provides several services for metadata management related to medical concepts. In health information exchange, the terminology server is important to ensure not only the standardization of terms but the information quality being integrated. The eHealth-Interop Terminology Server was designed based on the terminology specification of the HL7 FHIR, and has part of its design based on the OMG Common Terminology Services 2 (OBJECT MANAGEMENT GROUP, 2014) specification.

The main classes implemented are: CodeSystem, ConceptDefinition, ConceptMap and Mapping (Figure 3). CodeSystem represents a particular ontology or terminology, it represents the total set of codes and their meanings. Concepts represents a particular concept or term in a ontology or terminology. ConceptMap represents a mapping between two CodeSystem or part of it. A Mapping represents a relationship between two Concepts from different CodeSystems.

In the medical context, or specific mental health, CodeSystem represent from more complex ontologies such as SNOMED, terminologies such as ICD-10 and even local codifications such as to represent sexual orientation. ConceptDefinition, are specific concepts, for example, the diagnosis of ICD10, F03 - Dementia. Mapping can represent the mapping between two ontologies, such as CID10 and SNOMED, or the mapping between a local encoding (sexual orientation) and SNOMED, for example. ConceptMap are the individual concept-to-concept mapping, for example, Dementia (F03) of the ICD10 that can be mapped as equivalent to the concept of Dementia Disorder (C0497327).

Figure 3. Terminology Server Components

In order to support the information model defined for continued care in the MHCN, five terminologies were used:

- ICD-10 - International Statistical Classification of Diseases and Related Health Problems
- ICF - International Classification of Functioning, Disability and Health
- ICPC-2 - International Classification of Primary Care
- Table of Procedures, Medications and Orthotics, Prostheses and Materials of the Unified Health System in Brazil
- Brazilian Classification of Occupations

In addition, local terminologies were created to represent different concepts such as: sex, marital status, sexual orientation, among others. These codifications, for the most part, are in agreement with those used in the National Information System for primary care or based on information models from a national standard for discharge summary.

Record Linkage Module

Record Linkage can be defined generally as the process of identifying records from different sources that corresponds to the real world entity (GU, et al 2003). The record linkage process is also known as entity resolution, deduplication, entity matching, merge / purge problem, data reconciliation or in the a clinical context as patient matching, among others.

The record linkage process is essential to ensure the integrity of information and data quality during health information exchange and data integration.

Health information system usually implements the record linkage based on a deterministic method in which certain precise rules are defined in order to guarantee the uniqueness of a patient's record. In mental health context and in other areas such as medical emergencies, it is often not possible to obtain an unique identification or reliable information from patients. So, to use a deterministic method for patient record linkage is not always possible. In addition, duplicity may occur from erroneous manual registration by users operating health information systems. These errors and lack of information can be propagated in a scenario of health information exchange in which several systems are involved. To solve this kind of inconsistency , eHealth-Interop comprises a Record Linkage module that uses a two-stage verification to ensure the uniqueness of the patient record. This process is summarized in Figure 4. The first step is to extract patient identifiers, e.g. patient's national identifier or health record number, and apply the deterministic linkage using those identifiers. If a patient is not found, probabilistic linkage algorithms are applied using demographic information. When probabilistic linkage algorithms find a possible duplicity of identity, this duplicated record is assigned to a human audit process in which an expert analyzes the information and decides if a duplication really occurred or not. To implement the record linkage module we have used the software API Dedupe (GREGG, 2016).

Figura 4. Information flow during patient registry in eHealth-Interop

Security Server

Information security and privacy are important aspects of any information system. In the medical field it usually becomes a major barrier for the exchange of health information between different institutions (EDEN, 2016). The concern with information security and privacy was considered in different stages during the eHealth-Interop development: in the conceptual design of the integration platform architecture, for data exchange mechanism adoption and also for the definition of the user data access hierarchy.

The division of the data layer architecture in two different repositories, that store separately patient demographic information and clinical information, allows these repositories to be made available on different physical locations with different access methods, increasing security at system level. All data access is done through web services that use the HTTP over TLS protocol, establishing a secure communication channel between the client and the server, where the data is encrypted (NETWORK WORKING GROUP, 2000). Every client application to interoperate with eHealth-Interop platform must be previously registered as well as the users who will access the data through this application. The access is provided based on the OAuth2 protocol (IETF, 2012). The privacy guarantee when sharing patient data is performed following the model adopted by the Brazilian Ministry of Health for primary care, that is the opt-out model, in which the patient should choose not to share the data. All actions are stored in a logging system, allowing the establishment of an audit trail.

Communication API

All communication between the different information sources and the eHealth-Interop platform is done through web services. The web services provides a series of functionalities for registration and query of patient demographic data in the platform as well as all the clinical information associated with them. The webservice application was built using the OpenAPI Specification Standard (OPEN API, 2017). Clinical information is represented on information model defined *a priori* in clinical documents that are interoperated during different moments of the patient care process. A web tool was also developed to assist in the construction of adapters displaying all the services available in eHealth-Interop as well as how to access them (Figure 5).

Figure 5. eHealth-Interop web services API description tool

In addition, a web administration application was also built (Figure 6). Through this web application it is possible to manage all users and applications that have access to eHealth-Interop as well as the terminology server and to audit a possible duplicate records found by the Record Linkage module.

Figure 6. eHealth-Interop Administration System

Integrated Care Web Application

In order to allow the analysis of the integrated information, a web application was created that acts as client of the eHealth-Interop server. This tool is coupled with SISAM, allowing access to all professionals who work in the mental health in the region, based on the reuse of pre-defined and agreed access levels. In this way, access for patient demographic information and clinical information about appointments in the different attention levels are provided.

This web application also provides access to a timeline of the care events that a patient has participated (Figure 7).

Figura 7. Web application tool to access integrated healthcare information

Use Case Example in Mental Health Care Network

A total of 27353 patient health records were obtained from the five data sources used in this work, for a period from 2011 to 2016. From this total, 9548 (34.91%) records have been integrated in eHealth-Interop platform. In these records the same patient is present in at least two different source of information with the same identification data. The total number of unique patient records after integration is 4252, that means approximately 2.25 records per patient in the original data set. The number of data sources having the same patient's record are presented in Table 1.

Table 1. Distribution of the number of integrable patient records per data source

Number of data sources having	Number of unique patient's	Total of integrated records
-------------------------------	----------------------------	-----------------------------

the same patient's record	records	
2	3247 (76,36%)	6494 (68,02%)
3	968 (22,77%)	2904 (30,41%)
4	35 (0,82%)	140 (1,47%)
5	2 (0,05%)	10 (0,10%)
Total:	4252	9548

Completeness and syntactic accuracy analyzes were performed over patients demographic data. In total 26 different attributes have been analyzed. For each attribute of each record a value of 1 or 0 of completeness was assigned, based on the presence or absence of any data. Then the syntactic accuracy analysis were performed for those attributes with completeness value equals to 1. The syntactic accuracy analysis is done at attribute level for each patient record. For each attribute were defined rules of syntactic analysis, for example, date of birth, must be a valid date, considering that maximum age is 120 years. Then each attribute value is tested according to the defined rule of the respective attribute, for example, if the patient has the birth date as 01-02-1970, then the rule is applied to verify that this value is a valid birth date, which in this exemple, it is valid. If the value passes in the test, then it has a value of syntactic accuracy equals to 1, otherwise 0. Table 2 shows analysis for all demographic dataset, before and after all data is integrated in eHealth-Interop.

Table 2. Completeness and accuracy analysis in the whole dataset of integrated records

	Completeness (%)		Accuracy (%)	
	Non Integrated	Integrated (diff)	Non Integrated	Integrated (diff)
min	0.00	0.00 (=)	55.42	68.67 (+13.25)
max	100.00	100.00 (=)	100.00	100.00 (=)
mean	56.47	74.87 (+18.40)	95.69	96.77 (+1.08)
std	33.82	35.14 (+1.32)	10.08	7.25 (-2.83)

The results from Table 2 refer to the total set of integrated demographic information. The analyzes were redone by dividing demographic information into subgroups to better understand how eHealth-Interoperability affected data quality. Each subgroup consists of a set of related information. . Four subgroups were defined, they were named: General, Identification, Contact and Address. "General" subgroup consisting of general demographic information such as name, mother's name, father's name, gender, date of birth; "Identification" subgroup is composed by key attributes for patient unique identification, including identifiers used at national and health level; subgroup "Contact" which includes residential telephone, cell phone number, email; and finally subgroup "Address" including the residence address information of the patient such as, city name, street name, number. Table 3 and 4 shows the completeness and accuracy analysis, respectively, for each subgroup before and after the data integration through eHealth-Interop.

Table 3. Completeness analysis of demographic information divided in subgroups

Compleat.	General	General	Ident.	Ident.	Contact	Contact	Address	Address
-----------	---------	---------	--------	--------	---------	---------	---------	---------

(%)		Integrated		Integrated		Integrated		Integrated
min	0.00	1.25 (+1.25)	0.00	64.09 (+64.09)	0.00	3.28 (+3.28)	5.64	19.34 (+13.7)
max	100.00	100.00 (=)	94.57	98.39 (+3.82)	72.90	85.18 (+12.28)	80.93	99.46 (+18.53)
mean	62.57	76.02 (+13.45)	57.11	84.98 (+27.87)	43.90	53.50 (+9.6)	59.99	87.14 (+27.15)
std	35.20	35.41 (+0.21)	40.24	14.65 (-25.59)	38.66	43.98 (+5.32)	26.44	29.92 (+3.48)

Table 4. Accuracy analysis of demographic information

Accuracy (%)	General	General Integrated	Ident.	Ident. Integrated	Contact	Contact Integrated	Address	Address Integrated
min	55.42	68.67 (+13.15)	95.65	98.49 (+2.84)	99.82	99.78 (-0.04)	83.28	86.62 (3.34)
max	100.00	100.00 (=)	100.00	100.00 (=)	99.90	100.00 (=)	100.00	100.00 (=)
mean	93.83	95.21 (+1.38)	97.96	99.30 (+1.34)	99.86	99.90 (0.04)	95.55	96.43 (0.88)
std	13.85	9.88 (-3.97)	2.19	0.81 (-1.38)	0.06	0.11 (0.05)	7.00	5.79 (-1.21)

For duplicity analysis, a cured dataset was organized with records that did not have unique identifiers. Then a semi-automatic strategy was conducted: initially the record linkage module was used to perform a pre-selection of possible duplications and then, in the next step, each duplication result was manually verified for the determination of a true positive set of duplicates.

The record linkage algorithm has been configured to take into account the patient's name and mother's name, because they have 100% of completeness after data integration. The adopted threshold was 0.75 (referring to the probability of finding a duplicity) seeking for a high sensitivity for duplications detection. From a total of 6427 records tested, 1066 duplications were automatically pre-identified and, from those, 226 were manually confirmed.

Discussion

Principal Results

Coordination and continuity of care at different attention levels has been a challenge in many health systems. Major barriers are the lack of information and support information mechanisms to ensure this continuity. Health information systems are important tools to overcome these barriers but its fragmentation and heterogeneity, particularly prominent in the Brazilian national health system, does not allow to achieve the greatest advantages that their use can provide.

To solve this kind of limitation, an interoperable health care platform was designed and developed to provide the computational support to establish a health information exchange environment. In eHealth-Interop, the socio-demographic and clinical information are splitted into two different repositories, aiming to provide data security and extensibility.

In the definition of the interoperability information model, in a joint decision with health professionals, it was decided to consider as much as possible of the defined national standards, but also taking into account the other documents already used by the services locally and the international documents. Some concepts were further from the context of mental health such as surgical procedures performed, allergies. Despite this, the strategy of keeping this information as optional in the data model was adopted. The standardization and use of international documents may also allow a comparison with other services and contexts.

To enable semantic interoperability, a terminology server was developed based on the HL7 FHIR standard terminology service. It's modular design, following a service-oriented architecture, allows the terminology server to be reused as a separate software component for other contexts and projects.

Unlike other related work, eHealth-Interop allows for the exchange of patient data without unique, fully-defined identifiers that end up occurring in mental health and emergencies.

Limitations

In the mental health use case described, the eHealth-Interop support health information exchange between primary health care and mental hospital care, but it was not possible to conduct tests with specialized ambulatory care data. Although, after defining the relevant important information and generating the information model the tool is able to automatically generate the web services API and the database schema to store this information.

Another limitation of the current project, in the mental health use case, there is no mapping between the concepts of the adopted terminologies. The eHealth-Interop supports inter-terminologies mapping or this mapping to be done, however there was no need to translate concept from one terminology to another.

Conclusions and Future Directions

The eHealth-Interop is a computing platform designed for health information system interoperability. This proposed was successfully implemented and tested in the context of mental health care. This platform has been built to be: 1) flexible, so that it can be applied in other scenarios and clinical domains; 2) robust, so that records of patients with little information can be integrated and completed.

In order to measure and improve the quality of interoperable information data quality metrics have been applied. To validate the projected solution, we conduct analyzes on the obtained data sets, assessing the completeness, syntactic accuracy and duplicity metrics. We found that we achieved overall improvements in quality indicators. To prevent record duplication and unique identify the patient's records, an approach was established that takes into account deterministic rules and also probabilistic record linkage algorithms.

As future work we intend to support the creation of alerts and automatic warnings based on specific process, e.g. when a patient is hospitalized, a warning is generated for a certain health service warning about the event. We also intend to add similarity functions to the terminology server allowing a broad approach in concept searching.

Conflicts of Interest

None declared.

Acknowledgements

NSBM, PMAM and DA were responsible for the platform technical design, methods and computational techniques involved in the work while JMAM offered medical and mental health basis for the project's development. NSBM implemented the designed solution and conducted the quality tests and assessments. All authors participated in writing and reviewing the manuscript for publication. We would also like to thank the professionals of the Regional Health Department XIII, of São Paulo State, Brazil, responsible for the coordination of the regional mental health care services.

Abbreviations

FHIR - Fast Health Information Resource

HL7 - Health Level 7

MHCN - Mental Health Care Network

OMG - Object Management Group

RHD - Regional Health Department

References

THORNICROFT, G.; TANSELLA, M. The balanced care model for global mental health. **Psychological Medicine**, v. 43, n. 4, p. 849–863, 11 abr. 2013. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/22785067>>. Acesso em: 18 abr. 2017.

PATEL, V.; BELKIN, G. S.; CHOCKALINGAM, A.; COOPER, J.; SAXENA, S.; UNÜTZER, J. Grand Challenges: Integrating Mental Health Services into Priority Health Care Platforms. **PLoS Medicine**, v. 10, n. 5, p. e1001448, 28 maio 2013. Disponível em: <<http://dx.plos.org/10.1371/journal.pmed.1001448>>. Acesso em: 18 abr. 2017.

HAGGERTY, J. L.; REID, R. J.; FREEMAN, G. K.; STARFIELD, B. H.; ADAIR, C. E.; MCKENDRY, R. Continuity of care: a multidisciplinary review. **BMJ**, v. 327, n. 7425, 2003.

DURBIN, J.; BARNESLEY, J.; FINLAYSON, B.; JAAKKIMAINEN, L.; LIN, E.; BERTA, W.; MCMURRAY, J. Quality of Communication Between Primary Health Care and Mental Health Care: An Examination of Referral and Discharge Letters. **The Journal of Behavioral Health**

Services & Research, v. 39, n. 4, p. 445–461, 2 out. 2012. Disponível em: <<http://link.springer.com/10.1007/s11414-012-9288-9>>. Acesso em: 14 dez. 2016.

HIMSS. **Himss dictionary of healthcare information technology terms, acronyms and organizations**. 2. ed. [s.l.] HIMSS Publishing, 2010.

CHARLES, D.; SWAIN, M.; PATEL, V. **Interoperability among U.S. Non - federal Acute Care Hospitals , 2014**. [s.l.: s.n.].

COFFEY, R. M.; BUCK, J. A.; KASSED, C. A.; DILONARDO, J.; FORHAN, C.; MARDER, W. D.; VANDIVORT-WARREN, R. Transforming mental health and substance abuse data systems in the United States. **Psychiatric services (Washington, D.C.)**, v. 59, n. 11, p. 1257–63, nov. 2008. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/18971401>>. Acesso em: 1 abr. 2016.

SPIRANOVIC, C.; MATTHEWS, A.; SCANLAN, J.; KIRKBY, K. C. Increasing knowledge of mental illness through secondary research of electronic health records: opportunities and challenges. **Advances in Mental Health**, v. 14, n. 1, p. 14–25, 2 jan. 2016. Disponível em: <https://apps.webofknowledge.com/full_record.do?product=UA&search_mode=GeneralSearch&qid=3&SID=1EIZUvGx8i5rLE9HXiF&page=1&doc=1&cacheurlFromRightClick=no>. Acesso em: 16 maio. 2016.

PAIM, J.; TRAVASSOS, C.; ALMEIDA, C.; BAHIA, L.; MACINKO, J.; SILVA, J. da; AL., et. The Brazilian health system: history, advances, and challenges. **The Lancet**, v. 377, n. 9779, p. 1778–1797, 21 maio 2011. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/21561655>>. Acesso em: 18 abr. 2017.

RZEWUSKA, M.; DE AZEVEDO-MARQUES, J. M.; COXON, D.; ZANETTI, M. L.; ZANETTI, A. C. G.; FRANCO, L. J.; SANTOS, J. L. F. Epidemiology of multimorbidity within the Brazilian adult general population: Evidence from the 2013 National Health Survey (PNS 2013). **PLOS ONE**, v. 12, n. 2, p. e0171813, 9 fev. 2017. Disponível em: <<http://dx.plos.org/10.1371/journal.pone.0171813>>. Acesso em: 18 abr. 2017.

GREGÓRIO, G.; TOMLINSON, M.; GEROLIN, J.; KIELING, C.; MOREIRA, H. C.; RAZZOUK, D.; MARI, J. de J. Setting Priorities for Mental Health Research in Brazil. **Revista Brasileira de Psiquiatria**, v. 34, n. 4, p. 434–439, dez. 2012. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S1516444612000360>>. Acesso em: 18 abr. 2017.

TRAPÉ, T. L.; CAMPOS, R. O. The mental health care model in Brazil: analyses of the funding, governance processes, and mechanisms of assessment. **Revista de saúde pública**, v. 51, n. 0, p. 19, 23 mar. 2017. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/28355335>>. Acesso em: 18 abr. 2017.

SECRETARIA DA SAÚDE DO ESTADO DE SÃO PAULO. **Regionais de Saúde**. Disponível em:

<<http://www.saude.sp.gov.br/ses/institucional/departamentos-regionais-de-saude/regionais-de-saude>>. Acesso em: 18 abr. 2017.

STRONGLOOP. LoopBack - Node.js framework. 2017. Disponível em: <<http://www.webcitation.org/6fwAow3zf>>.

MONGODB. MongoDB for GIANT Ideas. <http://www.webcitation.org/6qZIFWE1o>. Published 2017. Accessed May 3, 2017.

OPEN API. Open API Initiative. <https://www.openapis.org/>. Published 2017. Accessed July 5, 2017.

ORACLE. Mysql. <https://www.mysql.com/>. Published 2018. Accessed January 5, 2018.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Cidades**. Disponível em: <<http://cidades.ibge.gov.br/v3/cidades/home-cidades>>. Acesso em: 18 abr. 2017.

DEPARTAMENTO DE ATENÇÃO À SAÚDE. **e-SUS Atenção Básica**. Disponível em: <<http://dab.saude.gov.br/portaldab/esus.php>>. Acesso em: 18 abr. 2017.

YOSHIURA, V. T.; DE AZEVEDO-MARQUES, J. M.; RZEWSKA, M.; VINCI, A. L. T.; SASSO, A. M.; MIYOSHI, N. S. B.; FUREGATO, A. R. F.; RIJO, R. P. C. L.; DEL-BEN, C. M.; ALVES, D. A web-based information system for a regional public mental healthcare service network in Brazil. **International Journal of Mental Health Systems**, v. 11, n. 1, p. 1, 3 dez. 2017. Disponível em: <<http://ijmhs.biomedcentral.com/articles/10.1186/s13033-016-0117-z>>. Acesso em: 24 fev. 2017.

DELFINI, M. G.; MIYOSHI, N. S. B.; ALVES, D. Minimum Data Consensus: Essential Information to Continuing Healthcare. In: 2015 IEEE 28th International Symposium on Computer-Based Medical Systems. IEEE, jun. 2015. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7167487>>. Acesso em: 11 out. 2015.

GHAFFARIAN, V. The new stream of socio-technical approach and main stream information systems research. **Procedia Computer Science**, v. 3, p. 1499–1511, 2011. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S1877050911000408>>. Acesso em: 30 maio. 2017.

MODELOS DE INFORMAÇÃO PARA INTEROPERABILIDADE

Os modelos de informação para interoperabilidade são descritos a seguir em detalhes. O formato para descrição do modelo de informação foi baseado no mesmo adotado pela Comissão de Estudos Especiais 78 (CEE-78) da Associação Brasileira de Normas Técnicas.

Os modelos estão descritos em formato tabular onde:

- **Coluna 1 - Nível:** descreve o nível do elemento no modelo de informação
- **Coluna 2 - Ocorrência:** descreve o número de vezes que o elemento pode aparecer de acordo com as seguintes regras
 - 0..1 – indica que o elemento não é obrigatório e, se ocorrer, só deve aparecer uma vez;
 - 1..1 – indica que o elemento deve estar presente pelo menos uma vez;
 - 0..n – indica que o elemento pode ocorrer várias vezes;
 - 1..n – indica que o elemento deve aparecer uma vez e pode ocorrer várias vezes;
- **Coluna 3 - Elemento:** Nome ou descrição do elemento no modelo de informação
- **Coluna 4 - Tipo de Dado:** formato esperado para representar o elemento

Modelo de Informação - Formulário Demográfico

Conteúdo

Nível	Ocorrência	Elemento	Tipo de Dado
1	[1..1]	Informações Sócio- Demográficas	
2	[1..1]	Nome completo	Sequência de caracteres alfanuméricos
2	[1..1]	Nome completo da mãe	Sequência de caracteres alfanuméricos
2	[0..1]	Nome completo do pai	Sequência de caracteres alfanuméricos
2	[0..1]	Data de nascimento	Data válida
2	[0..1]	Idade aproximada*	Número
2	[0..1]	Município de nascimento	Código IBGE do município
2	[1..1]	Sexo	Texto codificado
2	[0..1]	Raça/Cor	Texto codificado
2	[0..1]	Ocupação	CBO - Cadastro Brasileiro de Ocupação
2	[0..1]	Nome social	Sequência de caracteres alfanuméricos
2	[0..1]	Estado civil	Texto codificado
1	[0..1]	Identificação do Paciente	
2	[0..1]	Número do CNS	CNS válido
2	[0..1]	Número do RG	RG válido com dígito verificador
2	[0..1]	Orgão emissor do RG	Sigla do órgão emissor do RG
2	[0..1]	Número do CPF	CPF válido
1	[0..1]	Endereço de Residência	
2	[0..1]	Estado de residência	Sigla do estado
2	[0..1]	Município de residência	Código IBGE válido
2	[0..1]	CEP de residência	CEP válido
2	[0..1]	Logradouro	Sequência de caracteres alfanuméricos
2	[0..1]	Número da casa	Sequência de caracteres alfanuméricos
2	[0..1]	Bairro	Sequência de caracteres alfanuméricos
2	[0..1]	Complemento	Sequência de caracteres alfanuméricos

1	[0..1]	Contato do Paciente	
2	[0..1]	Telefone residencial	Sequência de caracteres alfanuméricos
2	[0..1]	Telefone comercial	Sequência de caracteres alfanuméricos
2	[0..1]	Celular	Sequência de caracteres alfanuméricos

Modelo de Informação - Formulário de Admissão Hospitalar

Nível	Ocorrência	Campo	Tipo de Dado
1	[1..1]	Identificação do Paciente	
2	[1..1]	Nome do paciente	Sequência de caracteres alfanuméricos
1	[1..1]	Identificação do Profissional Responsável	
2	[1..1]	Nome do profissional responsável pela internação	Sequência de caracteres alfanuméricos
1	[1..1]	Informações da Internação	
2	[1..1]	Data e hora da internação	Data válida
2	[1..1]	Tipo de internação (Caráter do atendimento)	Texto codificado
2	[0..1]	Motivo da internação	Sequência de caracteres alfanuméricos
2	[1..1]	Diagnóstico principal	Código válido CID10
2	[0..n]	Comorbidades	Código válido CID10
2	[0..n]	Incapacidades	Código válido CIF 2004
2	[0..1]	História clínica	Sequência de caracteres alfanuméricos
1	[1..1]	Medicação Inicial	
2	[0..N]	Prescrição da inicial	Texto livre
2	[0..1]	Lista de medicamentos iniciais (estruturada)	Prescrição estruturada – ou a forma em texto livre ou a estruturada deve estar preenchida
3	[1..N]	<i>Linha de prescrição</i>	
4	[1..1]	Nome do medicamento	(É informado o nome do medicamento, a concentração e forma farmacêutica)

			Exemplo: Paracetamol 500 mg comprimidos) Texto livre
4	[1..1]	Código do medicamento	Código da Tabela de Medicamentos ANVISA
4	[1..1]	Dose	Quantidade
4	[1..1]	Via de administração	Texto codificado por terminologia externa conforme vocabulário controlado de formas farmacêuticas, vias de administração e embalagens de medicamentos da Anvisa
4	[1..1]	Frequência de uso do medicamento	Intervalo de tempo
4	[1..1]	Duração de uso do medicamento	Duração
4	[0..1]	Orientação sobre o uso do medicamento	Texto livre

Modelo de Informação - Sumário de Alta Hospitalar

Nível	Ocorrência	Elemento	Tipo de Dado
1	[1..1]	Identificação do paciente e caracterização da internação	
2	[1..1]	Nome completo	Sequência de caracteres alfanuméricos
2	[1..1]	Nome completo da mãe	Sequência de caracteres alfanuméricos
2	[1..1]	Número do CNS	CNS com dígito verificador válido
2	[1..1]	Número do prontuário	Sequência de caracteres alfanuméricos
2	[1..1]	Caráter da internação	Texto codificado
2	[1..1]	Data e hora da internação	Data e hora válida
2	[1..1]	Data e hora da saída da internação	Data e hora válida
2	[1..1]	Modalidade de atendimento	Texto codificado
1	[1..1]	Motivo da admissão, diagnósticos relevantes e patologias associadas desenvolvidas na internação	
2	[1..1]	Código CID 10 do diagnóstico principal	Texto codificado por terminologia CID 10
2	[0..N]	Código CID 10 de outros	Texto codificado por terminologia CID

		diagnósticos relacionados ao principal	10
1	[0..N]	Classificação internacional de funcionalidade e incapacidade em saúde (CIF)	
2	[1..N]	Código CIF-2004	Texto codificado por terminologia externa CIF
1	[0..N]	Lista de Problemas	
2	[1..1]	Código do problema	CID 10
2	[1..1]	Status do problema	Texto codificado
1	[1..N]	Procedimentos diagnósticos realizados	
2	[0..1]	Data do resultado do procedimento	Data válida
2	[1..1]	Código do procedimento diagnóstico realizado na internação	Tabela de Procedimentos do SUS
2	[1..1]	Procedimento diagnóstico realizado na internação	Texto livre
2	[0..1]	Resultado e/ou observações do procedimento diagnóstico realizado	Texto livre
1	[1..N]	Procedimentos terapêuticos realizados	
2	[1..1]	Código do procedimento terapêutico realizado na internação	Tabela de Procedimentos do SUS
2	[1..1]	Texto do procedimento terapêutico realizado na internação	Texto livre
2	[0..1]	Resultado e/ou observações do procedimento terapêutico realizado	Texto livre
2	[0..1]	Data do resultado do procedimento	Data válida
1	[1..N]	Procedimentos cirúrgicos realizados	
2	[1..1]	Código do procedimento cirúrgico realizado na internação	Tabela de Procedimentos do SUS
2	[1..1]	Texto do procedimento cirúrgico	Texto livre
2	[0..1]	Descrições e/ou observações do procedimento cirúrgico	Texto livre

		realizado	
2	[0..1]	Data de realização do procedimento	Data válida
1	[1..1]	Resumo da internação (histórico da internação)	
2	[1..1]	Descrição dos principais eventos da internação incluindo, se necessário, medicamentos prescritos e suspensos	Texto livre
1	[0..N]	Alergias e reações adversas na internação	
2	[1..1]	Categoria do agente causador da alergia/reação adversa	Texto codificado
2	[1..1]	Agente/Substância específica	Texto livre
2	[0..1]	Tipo da alergia/reação	Texto codificado
2	[0..1]	Gravidade	Texto codificado
2	[0..1]	Data da instalação da reação adversa	Data/hora válidos
2	[0..1]	Evolução da alergia/reação adversa	Texto livre
1	[1..1]	Prescrição e orientação da alta	
2	[0..N]	Prescrição da alta (texto livre)	Texto livre
2	[0..1]	Lista de medicamentos da alta (estruturada)	Prescrição estruturada – ou a forma em texto livre ou a estruturada deve estar preenchida
3	[1..N]	<i>Linha de prescrição</i>	
4	[1..1]	Nome do medicamento	(É informado o nome do medicamento, a concentração e forma farmacêutica Exemplo: Paracetamol 500 mg comprimidos) Texto livre
4	[1..1]	Código do medicamento	Código da Tabela de Medicamentos ANVISA
4	[1..1]	Dose	Quantidade
4	[1..1]	Via de administração	Texto codificado por terminologia externa conforme vocabulário controlado de formas farmacêuticas, vias de administração e embalagens de medicamentos da Anvisa
4	[1..1]	Frequência de uso do	Intervalo de tempo

		medicamento	
4	[1..1]	Duração de uso do medicamento	Duração
4	[0..1]	Orientação sobre o uso do medicamento	Texto livre
1	[0..N]	Instruções; orientações; recomendações da equipe assistencial	
2	[1..1]	Profissional responsável pelo evento da internação	Texto livre
3	[1..1]	Nome do profissional responsável pela orientação/recomendação	Texto livre
3	[1..1]	Conselho profissional do profissional responsável pela orientação/recomendação	Conselho Profissional
3	[1..1]	UF do conselho profissional responsável pela instrução, orientação/recomendação	Texto codificado por terminologia IBGE
1	[0..1]	Informações da alta hospitalar	
2	[1..1]	Condição do paciente	Texto codificado
2	[1..1]	Desfecho da internação	Texto codificado
3	[0..1]	<i>Local para acompanhamento pós-alta</i>	
4	[1..1]	Nome do estabelecimento de saúde/ profissional para encaminhamento pós-alta	Texto livre
3	[1..1]	<i>Profissional responsável pelo evento da internação</i>	
4	[1..1]	Nome do profissional responsável pelo evento da internação	CRM ou COREN
4	[1..1]	Conselho profissional do profissional responsável pelo evento da internação	Texto codificado por terminologia IBGE
4	[1..1]	UF do conselho profissional responsável pelo evento da internação	Sigla do estado
1	[0..N]	Anexos	
2	[1..1]	Descrição dos anexos	Texto livre

Modelo de Informação - Formulário de Atendimento na Atenção Primária à Saúde

Nível	Ocorrência	Elemento	Tipo de Dado
1	[1..1]	Identificação	
2	[1..1]	CNS do Paciente	CNS válido
2	[1..1]	CNS do Profissional	CNS válido
2	[0..1]	Data do atendimento	Data válida
2	[0..1]	Prontuário	Número do prontuário
2	[0..1]	Local do atendimento	Sequência de caracteres alfanuméricos
1	[1..1]	Atendimento	
2	[0..1]	Sexo	Texto codificado
2	[1..1]	Data de nascimento	Data válida
2	[0..1]	Tipo de atendimento	Sequência de caracteres alfanuméricos
2	[0..1]	Peso	Número
2	[0..1]	Altura	Número
2	[0..1]	Vacinação em dia	Valor booleano
2	[0..n]	Exames Solicitados/Avaliados	Sequência de caracteres alfanuméricos
2	[0..1]	Conduta	Sequência de caracteres alfanuméricos
1	[0..1]	Gestante	
2	[0..1]	DUM	Data válida
2	[0..1]	Gravidez planejada	Valor booleano
2	[0..1]	Idade gestacional	Idade válida
2	[0..1]	Gestas prévias	Valor numérico
2	[0..1]	Partos	Valor numérico
1	[0..1]	Problema/Condição Avaliada	
2	[0..n]	CIAP	Código CIAP válido
2	[0..n]	CID 10	Código CID válido
1	[0..1]	Exames Solicitados/Avaliados	

2	[0..n]	Conduta	Sequência de caracteres alfanuméricos
2	[0..1]	Tipo de Encaminhamento	Sequência de caracteres alfanuméricos

Codificações e Terminologias

Nesta seção serão descritas as codificações locais e terminologias externas utilizadas para os elementos em que o tipo de dado é um texto codificado:

- Sexo
- Ocupação
- Estado Civil
- Raça/Cor
- Caráter da internação
- Modalidade do atendimento
- Status do problema
- Categoria do agente causador da alergia/reação adversa
- Gravidade
- Condição do paciente
- Desfecho da internação
- Município de nascimento
- Código CID 10 do diagnóstico principal / Código CID 10 de outros diagnósticos relacionados ao principal / Código do Problema
- Código CIF-2004
- Código do procedimento diagnóstico realizado na internação / Código do procedimento cirúrgico realizado na internação / Código do procedimento cirúrgico realizado na internação
- Código do medicamento

Sexo

Fonte: Local - Sumário de Alta Hospitalar ABNT

Código	Descrição
masculino	Sexo masculino
feminino	Sexo feminino
indeterminado	Sexo indeterminado

Ocupação

Fonte: Cadastro Brasileiro de Ocupações

Este elemento segue a terminologia externa CBO (Cadastro Brasileiro de Ocupações) v4.0.0.

Link: <http://www.mtecbo.gov.br/>

Estado civil

Fonte: eSUS-AB

Código	Descrição
solteiro	Solteiro
casado	Casado
viuvo	Viúvo
separado	Separado Judicialmente
uniao	União consensual

Raça/Cor

Fonte: eSUS-AB

Código	Descrição
branca	Cor branca
preta	Cor preta
parda	Cor parda
amarela	Cor amarela
indigena	Raça indígena

Caráter da internação

Fonte: Local - Sumário de Alta Hospitalar ABNT

Código	Descrição
eletiva	Internação eletiva
urgencia	Internação de urgência

Modalidade do atendimento

Fonte: Local - Sumário de Alta Hospitalar ABNT

Código	Descrição
hospitalar	Hospitalar
hospital-dia	Hospital DIA
domiciliar	Domiciliar

Status do problema

Fonte: Local - Sumário de Alta Hospitalar ABNT

Código	Descrição
ativo	Ativo
inativo	Inativo

Categoria do agente causador da alergia/reação adversa

Fonte: Local - Sumário de Alta Hospitalar ABNT

Código	Descrição
alimento	alimento, animal
ingrediente	ingrediente não ativo do medicamento
farmaco	fármaco(s) presente(s) no medicamento ou contraste radiológico

ambiental	produto ambiental
químico	outras substâncias ou produtos químicos
outros	outros

Gravidade

Fonte: Local - Sumário de Alta Hospitalar ABNT

Código	Descrição
baixo	Baixo risco – potencial impacto clínico de uma nova reação é considerado baixo. Existe contraindicação relativa de uso futuro da substância
alto	Alto risco – potencial impacto clínico de uma nova reação é considerado alto. Existe contraindicação ABSOLUTA de uso futuro da substância

Condição do paciente

Fonte: Local - Sumário de Alta Hospitalar ABNT

*Não se aplica no caso de evasão

Código	Descrição
curado	Paciente obteve cura
melhorado	Paciente melhorou em relação ao momento da internação
piorado	Paciente piorou em relação ao momento da internação
inalterado	Paciente permanece na mesma condição clínica que quanto ao momento em que internou

Desfecho da internação

Fonte: Local - Sumário de Alta Hospitalar ABNT

*Não se aplica no caso de evasão

Código	Descrição
ambulatorial	Paciente obteve alta para continuar tratamento ambulatorial
transferencia	Paciente teve transferência para outra unidade hospitalar
hospital-dia	Paciente teve alta para acompanhamento em hospital dia
domiciliar	Paciente obteve alta para acompanhamento domiciliar

Município de nascimento / Município de Residência

Fonte: IBGE

Este elemento é codificado conforme o código de município do IBGE

Link: <http://www.ibge.gov.br/home/geociencias/areaterritorial/area.shtm>

Código CID 10 do diagnóstico principal / Código CID 10 de outros diagnósticos relacionados ao principal / Código do Problema

Fonte: Classificação Estatística Internacional de Doenças e Problemas Relacionados à Saúde - CID-10

Estes elementos seguem a terminologia externa CID-10

Link: <http://www.datasus.gov.br/cid10/V2008/cid10.htm>

Código CIF-2004

Fonte: Classificação Internacional de Funcionalidade, Incapacidade e Saúde - versão 2004

Link: http://www.inr.pt/uploads/docs/cif/CIF_port_%202004.pdf

Código do procedimento diagnóstico realizado na internação / Código do procedimento cirúrgico realizado na internação / Código do procedimento cirúrgico realizado na internação

Fonte: Tabela de Procedimentos, Medicamentos e OPM do SUS

Link: <http://sigtap.datasus.gov.br>

Código do medicamento

Fonte: Tabela de medicamentos da Anvisa

Link: <http://portal.anvisa.gov.br/listas-de-precos>

MAPEAMENTO DO FORMULÁRIO DEMOGRÁFICO PARA SNOMED-CT

Foi realizado o mapeamento de cada conceito do Formulário Demográfico para o SNOMED-CT.

Na tabela que descreve o Formulário Demográfico, foi adicionada uma terceira coluna. Nesta coluna estão descritos os códigos dos conceitos na SNOMED assim como a nomenclatura oficial correspondente em inglês.

Formulário Demográfico / SNOMED

Nível	Ocorrência	Elemento	Snomed
1	[1..1]	Informações da Unidade de Saúde	
2	[1..1]	Número do prontuário	398225001 Medical record number (observable entity)
2	[1..1]	CNES	43741000 Site of care (environment)
1	[1..1]	Informações Sócio Demográficas	
2	[1..1]	Nome completo	371484003 Patient name (observable entity)
2	[1..1]	Nome completo da mãe	444301002 Mother of subject (person)
		Nome completo do pai	444295003 Father of subject (person)
2	[0..1]	Data de nascimento	184099003 Date of birth (observable entity)
2	[0..1]	Idade aproximada*	424144002 Current chronological age (observable entity)
2	[0..1]	Município de nascimento	433178008 City of residence (observable entity)
2	[1..1]	Sexo	184100006 Patient sex (observable entity)
2	[0..1]	Raça/Cor	103579009 Race (observable entity)
2	[0..1]	Ocupação	184104002 Patient occupation (observable entity)
2	[0..1]	Nome social*	429086007 Patient alias name (observable entity) 408676007 Patient preferred name (observable entity)
2	[0..1]	Estado civil	125680007 Marital status (observable entity)
1	[0..1]	Identificação do Paciente	
2	[0..1]	Número do CNS	422549004 Patient-related Identification code (observable entity)
2	[0..1]	Número do RG	422549004 Patient-related Identification code (observable entity)
2	[0..1]	Orgão emissor do RG	422549004 Patient-related Identification code (observable entity)
2	[0..1]	Número do CPF	422549004 Patient-related Identification

			code (observable entity)
1	[0..1]	Endereço de Residência - 184097001 Patient address (observable entity)	
2	[0..1]	Estado de residência	398070004 State (environment)
2	[0..1]	Município de residência	433178008 City of residence (observable entity)
2	[0..1]	CEP de residência	184102003 Patient postal code (observable entity)
2	[0..1]	Logradouro	398099009 Street address (observable entity)
2	[0..1]	Número da casa	734004005 Address (property) (qualifier value)
2	[0..1]	Bairro	734004005 Address (property) (qualifier value)
2	[0..1]	Complemento	734004005 Address (property) (qualifier value)
1	[0..1]	Contato do Paciente	
2	[0..1]	Telefone comercial	428843000 Patient work telephone number (observable entity)
2	[0..1]	Telefone residencial	429697006 Patient home telephone number (observable entity)
2	[0..1]	Celular	428481002 Patient mobile telephone number (observable entity)
2	[0..1]	E-mail	424966008 Patient - email address (observable entity)

APROVAÇÃO DO COMITÊ DE ÉTICA

**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: CSECUABA@FMRP-USP.BR

Figura 17 – Folha de aprovação do comitê de ética em pesquisa