

A CONSTRUÇÃO DE REPOSITÓRIOS DIGITAIS COMO FERRAMENTAS DE PRESERVAÇÃO E APROPRIAÇÃO DAS INFORMAÇÕES DE SAÚDE

Hélio Salomão Silva Cordoeira

ECA-USP

RESUMO: A pandemia do covid-19 demonstrou a capacidade catalisadora e disruptiva que a oferta de dados, informações e conhecimento em saúde tem tanto para o sistema de saúde em si quanto para a academia e o mercado. Construir políticas públicas cada vez mais engajadas em preservar informações em saúde através da criação e manutenção de repositórios digitais de informação em saúde se mostra uma solução cada vez mais tecnicamente viável e eficiente.

PALAVRAS-CHAVE: Informações em saúde, Repositórios digitais, preservação.

THE CONSTRUCTION OF DIGITAL REPOSITORIES AS TOOLS FOR THE PRESERVATION AND APPROPRIATION OF HEALTH INFORMATION

ABSTRACT: The covid-19 pandemic demonstrated the catalytic and disruptive capacity that the provision of health data, information and knowledge has both for the health system itself and for the academy and the market. Building public policies increasingly engaged in preserving health information through the creation and maintenance of digital health information repositories proves to be an increasingly technically feasible and efficient solution.

KEY WORDS: Health information, Digital repositories, preservation.

INTRODUÇÃO

A preservação de documentos digitais se mostra desafiadora tanto pela velocidade de produção, quanto pelas inseguranças que este suporte traz consigo. Hoje bilhões de terabytes de informação são produzidos em um ritmo nunca experimentado pela humanidade. Produzimos e armazenamos informação em escala e velocidades muito superiores a qualquer outro tempo, contudo o fazemos ainda sem garantias de que futuras gerações terão acesso a essas informações. É possível que uma parte considerável da história do início do século XXI se perca nos próximos anos.

Não estamos falando só de tweets, blogs e sites genéricos. Estamos falando de registros médicos, parte da cultura, registros de ações do estado. Tudo corre um grande risco de não existir ou existir sem uma forma viável de acesso. Um grande marco na preservação das informações digitais foi a perda dos registros das sondas Viking I e Viking II, esta perda fora o estopim para a construção do modelo OAIS e para o crescimento sobre

o Tema em todo mundo.

No campo da saúde a informação desempenha papel fundamental, médicos e profissionais da saúde estão cotidianamente tentando coletar informações processáveis e tomar decisões. Assim, no final do processo, quando da cura do paciente, aquele atendimento produziu e coletou um largo conjunto de informações e conhecimentos novos que podem ser reutilizados no atendimento de outros pacientes. O registro desse processo é útil não somente como história e memória, mas é capaz de alavancar o próprio processo de tomada de decisão em nível tático (Diagnostico) ou estratégico (Gestão, tanto da unidade quanto do sistema de saúde nos três níveis federativos).

Este artigo assim busca demonstrar como a construção de repositórios digitais como parte de uma política de preservação que garanta a comunicação através do tempo (Ferreira 2006), o aumento dos níveis de transparência e accountability, a apropriação das informações de saúde bem como possibilita a produção de produtos de informação que constituam melhorias de processos tornando-os mais eficientes, rápidos e seguros.

Durante a pandemia da Covid-19 foram construídas algumas soluções informacionais para a comunicação do andamento da pandemia para população, busca de criar ou melhorar os tratamentos para doença e no subsídio da pesquisa científica. Uma dessas soluções é o repositório DataSharing/BR ainda, fora profícua a produção de programas que utilizando inteligência artificial buscavam auxiliar os esforços do sistema de saúde na mitigação da pandemia.

Este trabalho busca juntar-se a outros na defesa de uma política de preservação de documentos de saúde que mais do que criar arquivos em unidades de saúde constitua um sistema capaz de dar conta das necessidades informacionais do Sistema de saúde Brasileiro. A construção de repositórios digitais com dados ou informações sobre saúde poderiam auxiliar tanto a tomada de decisões, quanto novas investigações. Impactar tanto a produção científica, quanto a gestão das unidades de saúde em nível local ou estratégico.

DESENVOLVIMENTO

Algumas definições passam a ser importantes para o prosseguimento desse trabalho. Primeiramente o dado. Esta escolha, de iniciar pelo dado, não é trivial. É possível que ontologicamente falando a menor parte que poderíamos decompor uma mensagem ou documento seria o dado. Para Semidão figura como “fato” ou “átomo”, um elemento bruto, desprovido de significação imediata. Algo perceptível e de baixo teor semântico. (SEMIDÃO, p. 71 2004)

Semidão faz uma análise etimológica do termo “dado”. Para ele é possível extrair algumas noções centrais acerca do uso da palavra. A partir da sua pesquisa o autor expressa que dado é utilizado como ideia de algo que aconteceu, feito ou realizado, algo que é atribuído a um objeto cognoscitivo uma característica que expressa a essência constitutiva dele. Estatísticas e fatos coletados que juntas são capazes de criar referencias e análises, coisas conhecidas e ou assumidas como fatos e que constituem base para

o raciocínio ou cálculo, suposição ou premissa de onde se pode iniciar inferências e por fim um pedaço de informação. (SEMIDÃO, p. 71-72 2004) Estas características, segundo Semidão, corroboram a noção depreendida a partir delas: dados como ponto base de um processo cognitivo. (SEMIDÃO, p. 72 2004)

Uma coisa dada ou concedida; algo conhecido ou suposto como fato, e feito base para raciocínio ou para cálculo; uma suposição ou premissa da qual são elaboradas inferências. (SEMIDÃO, p. 72 2004)

Silva e Gomes (2015) em seu artigo buscaram organizar diversas manifestações conceituais de informação na CI. A ideia de dado como base ou parte da informação é recorrente nas definições sobre informação. Para Hayes (1986) informação é uma propriedade dos dados resultante de ou produzida por um processo realizado sobre os dados (SEMIDÃO, p. 72 2004) Saracevic e Wood (1986) informação consolidada seria um conjunto de mensagens ou um sentido atribuído aos dados. A ideia de dados organizados como informação também está presente no Harrold's Librarian's Glossary (1989). (SILVA E GOMES, p. 146-147 2015)

Para Luciano Floridi (2005),

Informação semântica definida em quatro etapas: D.1. A Informação (λ) é constituída por n dados (d), sendo $n \geq 1$; D.2. Os dados são bem formados (wfd); D.3. Os wfd são significativos, ou seja, possuem um significado ($mwfd = \delta$); F.4. Os δ são verdadeiros. (SILVA E GOMES, p. 147 2015)

Silva e Gomes 2015 Relações de significado quantitativo (metadados) e qualitativo (conteúdos histórica e cognitivamente potenciais dos sujeitos da informação ((desenvolver))) p.148

Semidão (2004) em sua análise etimológica do vocábulo informação conclui que as definições o termo informação está invariavelmente relacionado ao conhecimento e à comunicação, independentemente da motivação que suscita o ato de informar (SEMIDÃO, p. 77 2004). Assim, todas as definições remeteriam em maior ou menor grau a noção processo cognitivo (processo de conhecer) em outras palavras, uma transposição ou mudança de um estado de coisas para outro mais apto e valorado, em termos cognoscitivos. (SEMIDÃO, p. 78 2004)

Para Silva e Gomes (2015) a pluralidade conceitual de informação na ciência da informação seja considerando particularidades conceituais, seja considerando as vozes conceituais dos estudiosos reverberam de uma multiplicidade de locais e origens em todo o globo e constroem percepções universais ou locais sobre informação.

Para Capurro e Hjørland (2007) O conceito de informação como é usado em inglês cotidianamente, no sentido de um conhecimento comunicado desempenha papel central na sociedade atual. Eles citam um estudo de Schrader (1983) no qual o autor estudou aproximadamente 700 definições de CI e seus antecedentes, e conclui que:

A literatura de CI é caracterizada pelo caos conceitual. Este Caos conceitual

advém de uma variedade de problemas na literatura conceitual da CI: citação acrítica de definições anteriores, fusão de teoria e prática, afirmações obsessivas de status científico, uma visão estreita da tecnologia, analogias inadequadas, definições circulares e multiplicidade de noções vagas, contraditórias e, às vezes, bizarras quanto à natureza do termo informação (SCHRADER, 1983, p.99)

Para os autores não deveríamos apresentar o conceito de informação isoladamente, mas deveríamos considerá-los de junto com outros conceitos. Os autores seguem suas conclusões avaliando que por mais que uma definição conceitual possa beneficiar o status do profissional da área, tal condição não é absoluta e existe boas produções que não se valem de um conceito de informação. Assim, o termo deve ser usado para melhor elucidar as questões teóricas e não a ser.

Há muitos conceitos de informação, como já demonstrado nesse trabalho, estes estão espalhados e inseridos em estruturas teóricas que segundo os autores são ora mais explícitas ora menos. Contudo, Capurro e Hjørland (2007) consideram uma distinção mais importante: entre informação como objeto ou coisa (por exemplo, números de bits) e informação como conceito subjetivo, informação como signo; isto é, como dependente da interpretação de um agente cognitivo. (CAPURRO e HJORLAND p. 193 2007)

Segundo Innarelli (2015), a sociedade preserva e constrói conhecimento por meio do processo de registro e da ininterrupta transmissão de informação. Através da capacidade humana de apropriar-se do conhecimento gerado e armazenado o reproduz e o desenvolve. E assim, através desse ciclo é possível alcançar um novo estado de conhecimento e o desenvolvimento sociocultural e científico da humanidade.

Esta transmissão ininterrupta do conhecimento a apropriação pelo ser humano ao longo da história se realizou para a satisfação das crescentes necessidades humanas. A transformação das tecnologias de transmissão da informação que culminariam no que hoje chamamos de sociedade da informação se desenrolam de igual modo a partir da mesma razão. (INNARELLI p. 40, 2015)

Para Le Coadic (2004) um conhecimento é o ato de conhecer, neste ato o espírito apreende um objeto. Nesse sentido conhecer algo é ter a capacidade de construir a ideia sobre esta coisa. Para Nonaka e Takeuchi (1997) e Nonaka (2000) o conhecimento é fruto da espiral de conhecimento e diferencia-se em conhecimento tácito e explícito, sendo o conhecimento tácito o conhecimento pessoal e desenvolvido a partir de experiências individuais e o explícito o formal, sintetizado e registrado. A produção do conhecimento nesse modelo depende da transmissão. Este acontece em duas redes distintas, a rede soft é em geral formada por relações informais e sustenta a criação de conhecimentos tácitos já a rede hard se constitui de relacionamentos formais, institucionais, pelas estruturas administrativas e infraestruturas tecnológicas comunicacionais e sustenta a produção e conhecimento explícito. (INNARELLI p. 41, 2015)

Em concordância com Nonaka e Takeuchi sobre o processo de produção

e transmissão do conhecimento, Davenport e Prusak (1998) também fazem uma abordagem conceitual da comunicação da informação na instituição. Para os autores a informação é comunicada a partir das redes hard e soft, sendo que ambas permeiam a geração do conhecimento nas instituições. Rossetti e Morales (2007) também apresentam a mesma visão de Nonaka e Takeuchi no que diz respeito ao conhecimento tácito e explícito e às redes de comunicação apresentadas como soft e hard (INNARELLI p. 41, 2015)

Buckland (1991) aborda informação como o objeto de produção do conhecimento e subdivide três categorias distintas information-as-process, information-as-knowledge e information-as-thing. Nesta abordagem, o autor apresenta a information-as-process como processo de transmissão da informação, information-as-knowledge como informação assimilada (não registrado) e a information-as-thing como conhecimento tangível (registrado). (INNARELLI p. 42, 2015)

Para Innarelli (2015) é possível entender information-as-process e o information-as-knowledge correspondem à transmissão e assimilação da informação. E o conhecimento explícito registrado em documentos e arquivos refere-se ao information-as-thing.

Para Buckland (1997,1998) o documento tem como base a informação registrada, este registro de conhecimento acontece por meio do uso de recursos tecnológicos, as redes formais de comunicação, da infraestrutura da instituição e do registro da informação. Innarelli (2015) expressa essa definição por meio de uma expressão: Documento = Conhecimento explícito = Information-as-thing (INNARELLI, 2015, p.52)

O objeto da preservação da informação é o conhecimento explícito, ou registrado, a information-as-thing ou o documento. Sobre este objeto que as teorias, técnicas e estratégias de preservação se debruçaram.

Para Innarelli (2015) a diferença entre um documento analógico e um documento digital precede o suporte, sendo o fator preponderante a capacidade humana de ter acesso à informação registrada sem a utilização de uma tecnologia externa que decodifique a informação. Como forma de exemplo ele utiliza um cartão perfurado onde a informação ali registrada não está acessível a um usuário a menos que ele tenha acesso a uma tecnologia que o auxilie, o suporte físico ser o papel não garante que o objeto informacional seja analógico.

Devemos, contudo, salientar que os textos armazenados em computadores ou mesmo na nuvem estão em algum lugar armazenado em discos físicos, o objeto físico do documento digital constitui o que o hardware é capaz de interagir e interpretar. Assim, o hardware transformará os símbolos inscritos no suporte físico em um conjunto de dados inteligíveis para o software em alguma linguagem de programação. Este converterá as mesmas em uma linguagem inteligível para o usuário. (FERREIRA, 2006, p. 22)

Para Innarelli (2015) uma vez compreendemos a importância do documento frente ao processo de valorização da informação na sociedade da informação e ao abordar os documentos no âmbito da CI surgem duas questões importantes. A primeira quais são

estes documentos? E a segunda o fato de serem considerados documentos o tornam arquivísticos? (INNARELLI, p. 52. 2015)

Para responder essas questões, Innarelli inicia discutindo os objetivos das instituições responsáveis pela guarda de documentos uma vez que o significado de documento poderá variar a partir de cada instituição. (INNARELLI, p. 53. 2015)

Assim, segundo Paes:

ARQUIVO – É a acumulação ordenada dos documentos, em sua maioria textuais, criados por uma instituição ou pessoa, no curso de sua atividade, e preservados para consecução de seus objetivos, visando à utilidade que poderão oferecer no futuro.

BIBLIOTECA – É o conjunto de material, em sua maioria impresso, disposto ordenadamente para estudo, pesquisa e consulta.

MUSEU – É uma instituição de interesse público, criada com a finalidade de conservar, estudar e colocar à disposição do público conjuntos de peças e objetos de valor cultural (PAES, 2004, p. 16). (INNARELLI, p. 52. 2015)

Para Belloto,

A forma/função pela qual o documento é criado é que determina seu uso e seu destino de armazenamento futuro. É a razão de sua origem e de seu emprego, e não o suporte sobre o qual está constituído, que vai determinar sua condição de documento de arquivo, de biblioteca, de centro de documentação ou de museu (BELLOTTO, 2006, p. 36).

Innarelli (2015) ao definir os destinos preferenciais de cada tipo de documentos, descreve como sendo documentos de biblioteca qualquer informação fixada em suporte cuja função principal é didática, técnica, cultural e/ou científica. (INNARELLI, p. 54. 2015) Já ao descrever os documentos preferenciais dos museus, Innarelli argumenta que apesar de museus e Bibliotecas serem instituições colecionadoras e temáticas museus se distinguem das bibliotecas por atribuírem significados diferentes para cada objeto. (INNARELLI, p. 54. 2015)

No museu, de forma muito contundente, tem-se consciência de que o acervo é composto de obras ou peças aos quais se atribui – intencionalmente – um valor documental. [...] No entanto, as peças ou objetos, na condição de documentos intencionalmente incorporados ao acervo museológico, não significam per se: o significado lhes é atribuído em função dos objetivos do museu. Assim, por exemplo, uma faca de prata receberá significados e descrições diferenciadas caso integre o acervo de um museu do crime, um museu da gastronomia ou um museu da ourivesaria (SMIT, 2011, p. 33) (INNARELLI, p. 54. 2015)

Na visão arquivística, argumenta Innarelli (2015), o documento produzido no decorrer das atividades ou funções de uma entidade pública ou privada, tem como função primária o registro das atividades ou funções que viabilizam o funcionamento institucional. (INNARELLI, p. 54. 2015) Assim, para o autor, o documento de arquivo é visceral pois sua relação com o produtor é o fator que o define como arquivístico, e assim sendo, sem esta

relação o documento perde seu significado e seu valor arquivístico.

Os documentos de arquivo são produzidos por uma entidade pública ou privada ou por uma família ou pessoa no transcurso das funções que justificam sua existência como tal, guardando essas documentos relações orgânicas entre si. [...] Surgem, pois, por motivos funcionais administrativos e legais. Tratam sobretudo de provar, de testemunhar alguma coisa. (BELLOTTO, 2006, p. 37) (INNARELLI, p. 54. 2015)

Masson (2008) ao tentar responder o que é um repositório faz o que ela chama de “arqueologia da palavra”, esta ação se justificaria por causa dos múltiplos sentidos e usos que se tem atribuído a palavra repositório. (MASON p106. 2008) Assim, o primeiro sentido apresentado, dicionarizado, repositório seria aquilo que “que é próprio para guardar alguma coisa; depósito” (MASON p106. 2008)

O verbete, “repertório”, do latim “repertorium”, é usado em sinonímia a “catálogo seletivo” e a “inventário”, em terminologia Arquivística, e é terminologia biblioteconômica. O significado vernacular, no dicionário, é registrado como “coleção, compilação, conjunto” (Idem: 1218), mantendo, portanto, similaridade de sentido com o significado terminológico (MASON p106-107. 2008)

A palavra repositório aparece como um verbete no dicionário Terminologia Arquivística (Camargo e Bellotto, 1996: 67). “Repositório Ver: depósito (2).” No verbete “depósito”, temos o sentido 1: “Ação pela qual os documentos são colocados sob a custódia de uma instituição arquivística.” No sentido 2, indicado pelo verbete “repositório”, assim está definido: “Prédio ou área para armazenamento” (MASON p107. 2008) Contudo para Mason argumenta que o verbete que lhe parece mais próximo do objeto repositório digital seria: “depósito legal: obrigação legal de remeter a bibliotecas, e excepcionalmente a arquivos públicos, um ou mais exemplares de cada obra impressa no país, estado ou município.” (MASON p107. 2008)

“Dicionário do Livro: terminologia relativa ao suporte”, Dicionario enciclopédico de Ciencias de la Documentación de Lópes Yepes” e “Dicionário de terminologia arquivística” o verbete Repositório não foi encontrado. Contudo no dicionário do IBLN, encontramos o verbete deposito legal: “Depósito legal 1. Entrega obrigatória e gratuita de um ou vários exemplares de toda e qualquer publicação impressa num País a uma ou mais instituições públicas para tal designada na lei. 2. Entrega imediata, obrigatória e gratuita de originais de determinados documentos de arquivo a uma instituição pública designada na lei para sua depositária (ex: leis, acordos, etc.) v. tb Aquisição p. 31.” (MASON p107-108. 2008)

No “Diccionario del archivero bibliotecario: terminología de la elaboración, tratamiento y utilización de los materiales propios de los centros documentales” contém o verbete “repositorio archivístico: lugar donde se guarda un archivo” amplia: “repositorio bibliográfico: biblioteca” e “repositorio documental: archivo”. Nos conceitos assim definidos, faz sentido o paralelismo: “repositório digital: repositório de objetos digitais”, mas, sem dúvida, tautológico. (MASON p108. 2008)

Mason (2015) na tentativa de averiguar a possível função paralela do depósito legal e do repositório apresenta Lei Federal (brasileira) nº 10.994, de 14 de dezembro de 2004 que define a Biblioteca Nacional como depósito legal de todas as publicações do país. (MASON p108. 2008)

No site da Biblioteca Nacional, há a citação do Artigo 2º da Lei nº 10.094/2004, com a definição de “depósito legal” como sendo uma exigência, por força de lei, de remessa à Biblioteca Nacional de um exemplar de todas as publicações produzidas em território nacional, por qualquer meio ou processo. O objetivo principal do Depósito Legal é assegurar a coleta, a guarda e a difusão da produção intelectual brasileira, visando à preservação e formação da Coleção Memória Nacional (MASON p108. 2008)

O repositório digital, argumenta Mason (2015) instituições) parece ter como objetivo preservar as publicações criadas em meio digital, ou as que são digitalizadas, referentes a artigos em periódicos, ou na web, os próprios periódicos, as atas de congresso e outros eventos, trabalhos de pesquisas, teses e dissertações para que constituam uma memória institucional e que possam estar disponível e acessível a quem precisar consultá-la, ou, no caso dos repositórios digitais temáticos, a memória do conhecimento em uma área especializada do conhecimento. (MASON p.109, 2008)

Mason argumenta que de uma perspectiva epistemológica e teórico - metodológico os repositórios digitais não têm sido objetos de estudos na CI mesmo que estejam frequentemente presentes em artigos. Estes parecem estar mais dedicados a defesa da divulgação das implementações até agora realizadas do que preocupados em estudar os repositórios digitais na área da CI. (MASON p128. 2008)

Mason (2008) reúne definições de repositórios digitais, estas definições permitem averiguar a pluralidade de respostas para a pergunta: “O que são repositórios digitais”. Estas variam entre ser uma ferramenta, um serviço ou uma instituição, cujo objetivo principal é a preservação e disseminação da informação, em geral, restrita à informação científica e o seu uso na comunidade científica e nas universidades, numa modalidade alternativa e aperfeiçoada de comunicação científica, assim como são os periódicos e publicações eletrônicas (MASON p132. 2008)

repositório digital aparece definido ora como coleção, ora como arquivo. Em qual dos sentidos de arquivo? É no sentido percebido em Informática? Se fosse assim porque a designação original foram “open archives” e não “open files”? Tanto coleção como arquivo são vocábulos que aparecem nas definições ... seja na afirmação de que repositórios institucionais são constituídos por “coleção de arquivos digitais”, ou “arquivos digitais de coleções”, o que evidentemente não é a mesma coisa, pois a ordem, neste caso, altera o significado. Ou, ainda, como “coleções digitais que armazenam [...] a produção intelectual das comunidades universitárias”, o que resulta em circularidade e redundância. (MASON p134. 2008)

O repositórios digitais então funcionariam a partir das seguintes características: o auto arquivamento, como uma variação dos depósitos legais de obras; a interoperabilidade,

para a garantia da disseminação e acesso, através da adoção de protocolos de arquivos abertos e, idealmente de acesso aberto, e a preservação de informação produzida em meio digital, constituindo se, ao mesmo tempo, em garantia de visibilidade institucional e da memória institucional, porém em um conceito de memória diverso daquele previsto, tradicionalmente, pelos arquivos e bibliotecas. (MASON p139. 2008)

Para Ferreira (2006) os níveis de abstração do objeto digital iniciariam com o Objeto físico, este inscrito em uma linguagem de máquina seria interpretado pelo hardware que o elevaria a objeto lógico inscrito em alguma linguagem de programação seria então interpretado pelo Software convertido em um objeto conceitual que poderia então ser utilizado pelo ser humano e assim construiria um objeto experimentado. Essa sequência é uma via de mão dupla, é possível que este documento esteja sendo lido de forma digital e nesse sentido o sistema estaria descrito na ordem exata que está ocorrendo. Contudo, este documento também fora concebido de forma digital e assim sendo, este passou do conceito imaginado por mim até o seu armazenamento no suporte físico em uma linguagem e máquina, para só depois poder ser transportado, copiado ou lido.

Em uma situação ideal o objeto conceitual, seja o documento analógico ou digital, deverá se assemelhar muito ao objeto conceitual imaginado pelo emissor. Assim sendo, A preservação digital é a atividade responsável por garantir que a comunicação entre um emissor e um receptor é possível, não só através do espaço, mas também através do tempo. (FERREIRA, 2006, p. 24). Para Ferreira, a preservação digital só é possível se a totalidade da abstração já descrita acima se encontrarem descrita e interpretáveis. Se a cadeia de interpretação que permite elevar um objeto digital desde o seu nível físico até ao nível conceptual for rompida, a comunicação deixa de ser possível e o objeto perder-se-á para sempre. (FERREIRA, 2006, p. 24)

O número de estratégias disponíveis com o objetivo de solucionar a problemática da preservação dos objetos digitais tem crescido nos últimos anos segundo Ferreira (2006). Para Lee as estratégias podem ser agrupadas em três classe fundamentais emulação, migração e encapsulamento (FERREIRA, 2006, p. 31). Já Thibodeau organiza a estratégias em um mapa bidimensional colocando em extremos do eixo cartesiano a preservação do objeto físico/lógico e no outro a preservação do objeto conceitual. (FERREIRA, 2006, p. 31)

Para Ferreira (2006) em termos genéricos o conceito de autenticidade é utilizado para traduzir *elementos diplomáticos que permitem aferir se um dado objeto é autêntico* (FERREIRA 2006, p.50) Refere-se a identificação *do “porquê”, do “quando”, do “onde” e do “por quem” de um objeto digital* (FERREIRA 2006, p.50). Com essa finalidade, seria segundo Ferreira, imprescindível *contextualizar a sua existência, descrever a sua história de custódia e atestar que a sua integridade não foi comprometida, provar que existe um conjunto de propriedades, consideradas significativas, que foram corretamente preservadas ao longo do tempo.* (FERREIRA 2006, p.50)

Se o problema da autenticidade não nasceu com o advento do objeto digital, uma vez

que esta preocupação é encontrada ao longo da história, é no contexto do documento de natureza digital que este problema se torne sensivelmente mais complexo (Ferreira 2006) A simplicidade com que alterações podem ser inseridas e a rapidez com que esta informação corrompida pode ser disseminada elevam as dificuldades de garantir a autenticidade de documentos digitais.

A preservação de documentos de Saúde

No período de 1918 a 1919 a gripe espanhola se espalhou pelo mundo inteiro, matando de 20 a 80 milhões de pessoas. De origem viral, não havia tratamento conhecido. Como veio, se extinguiu. Com o intuito de pesquisar meios de evitar uma nova catástrofe, a comunidade internacional das áreas médica e de saúde pública procurou por décadas algum vestígio biológico do vírus causador dessa enfermidade. Só depois de muito tempo, foi encontrada uma amostra de tecido humano infectado pelo vírus num hospital militar da Inglaterra. A partir desses vestígios estão sendo desenvolvidas pesquisas para se descobrir vacinas e meios de tratamento da gripe espanhola. As pesquisas em torno da amostra só se tornaram possíveis graças à preservação dos arquivos científicos, datados de 1916, daquele hospital militar (SAYÃO, SALES, p. 179, 2012)

Para Moreno, Coeli e Munck (2009) desde a antiguidade é possível encontrar a necessidade de comunicar algo a alguém sobre a própria saúde, ou sobre a saúde de um terceiro ou mesmo de um grupo. E assim, informação de saúde poderia de forma inicial ser traduzido como um compósito de transmissão e/ou recepção de eventos relacionados ao cuidado em saúde. A partir do século XIX o crescimento dos estudos em epidemiologia e a necessidade crescente de se comunicar questões relacionadas à saúde de grandes populações aumenta consideravelmente as informações de saúde produzidas e transmitidas.

Quase que concomitantemente, a estatística do final desse século XIX e início do século XX, inspiradora de estudiosos como Benthan, Price, Laplace, Galton (Rosen, 1994) pode ser vista, também, como um ponto de partida importante para a geração de Informações em Saúde de forma agregada e preditiva. Daí, pode-se partir, sem muito pecado, para as primeiras peças da Informação em Saúde, compostas pelas Estatísticas Vitais, pelas Tábuas de Sobrevida, enfim, por instrumentos de predição e inferência de estados de saúde a partir do status atual de um grupo de pessoas em determinado contexto de saúde. (Moreno, Coeli e Munck 2009)

Moreno, Coeli e Munck (2009) listam três subáreas que se formaram a partir do desdobramento de numerosos caminhos que a informação de saúde percorreu. A primeira abarca o perfil da população (de que adoecer e morre, dados demográficos e socioeconômicos), bem como os serviços prestados os materiais e medicamentos consumidos e a força de trabalho envolvida. A segunda subárea corresponde a conhecer as necessidades da população atendida, bem como o uso potencial e real da rede instalada e os possíveis investimentos necessários. A última subcategoria seria a fim de planejar, controlar e avaliar as ações e serviços de saúde (EPSJV, 2005).

Daí termos, de forma esquemática (Moraes, 2007), a possibilidade de

observar a Informação em Saúde como subsídio para o próprio setor saúde: na administração; na assistência; no controle e avaliação; no orçamento e finanças; no planejamento; nos recursos humanos; na regulação; na saúde suplementar; no geoprocessamento em saúde; e na vigilância (epidemiológica, sanitária, ambiental) (Moreno, Coeli e Munck 2009)

A preservação dos metadados, informações, documentos e conhecimentos produzidos no âmbito do atendimento e nas unidades de saúde não é de assim um objeto de importância exclusiva de entusiastas ou cientistas da informação. Historicamente e atualmente melhorias na gestão do conhecimento e informações de saúde constituíram poderosas ferramentas para os sistemas de saúde. Sobre a informação de saúde Piconi, Galvão e Ricarte (2013) pontuam:

a informação gerada durante o tratamento do paciente é de importante valor para o próprio paciente, para a população e para a ciência, pois é uma fonte de pesquisa, sendo considerada uma fonte de memória, patrimônio documental e cultural de uma sociedade, além de ser de extrema importância para garantir o atendimento médico e a eficiência no diagnóstico para tratamento do paciente (PICONI; GALVÃO; RICARTE, 2013, p.4)

Para Mello e Vianna (2019) esses elementos justificam a importância de se preservar as informações digitais na área da saúde, considerando que provavelmente foram os mesmos elementos que motivaram um número considerável de países a promoverem políticas de preservação digital de informações de saúde entre os quais os autores citam:

iniciativa italiana do Polo archivistico dell'Emilia-Romagna (ParER), a eBiblioCCE desenvolvida pelo Centro Nacional de Cirurgia, na Espanha, o Sistema de Preservação de Recursos Eletrônico (SPER), proposto pelo National Institutes Of Health, EUA, a metodologia de auditoria para sistemas de informação para avaliar a sua capacidade de preservar a informação digital de forma segura a médio e longo prazo –TRAC, desenvolvida na Espanha, a proposta para armazenar às informações em DNA, desenvolvida na Europa e entre outras. (Mello e Vianna p. 22 2019)

Para Pinto, Sales (2017) a preservação das informações de saúde é fundamental tanto as oriundas de pesquisas científicas quanto as produzidas pelo atendimento médico, como os prontuários eletrônicos (PEP). Para as autoras a preservação do PEP,

representante maior da chamada documentação sanitária – garantiria a realização de pesquisas e estudos comparativos embasados nessas fontes que visassem investigar o surgimento dessas doenças ou se elas se 'inovam' bem como as medidas terapêuticas postas em execução tendo em vista a sua cura ou, ainda, as orientações prescritas para se conviver com elas (Pinto, Sales p. 2017)

E assim, concluem, uma política de preservação desses documentos seria uma condição zine qua non tanto para medicina baseada em evidência quanto para pesquisas, ensino e para a constituição de provas jurídicas.

Repositórios de Informações de saúde, DataSharing/BR

No dia 01/07/2020 entrava em operação o COVID-19 Data Sharing/BR, este é o primeiro repositório de dados abertos relativo a Covid-19 do Brasil com dados laboratoriais e clínicos de hospitais do estado de São Paulo. De acordo com a Fapesp “*o repositório abriga dados abertos e anonimizados de, inicialmente, mais de 177 mil pacientes, 9.634 dados de desfecho e um total de quase 5 milhões de resultados de exames clínicos e laboratoriais realizados em todo o país pelo Grupo Fleury e na cidade de São Paulo pelos hospitais Israelita Albert Einstein e Sírio-Libanês desde novembro de 2019.*”

A base de dados fora resultante de uma iniciativa da própria Fapesp em parceria com a USP e tem como objetivo compartilhar “*informações compartilhar informações clínicas anonimizados de pacientes para subsidiar pesquisas científicas sobre a doença nas diversas áreas de conhecimento*”

A plataforma disponibiliza 3 conjuntos de dados, um por cada hospital, cada um deles com um dicionário de dados, um .csv para pacientes anonimizados e um .csv para resultados de exames. Na planilha paciente é possível encontrar as seguintes informações: ID_PACIENTE, IC_SEXO, AA_NASCIMENTO, CD_PAI, ICD_UF, CD_MUNICIPIO, CD_CEP. Para planilha exames é disponibilizada as seguintes informações: ID_PACIENTE, DT_COLETA, DE_ORIGE, IDE_EXAME, DE_ANALITO, DE_RESULTADO, CD_UNIDADE, DE_VALOR_REFERÊNCIA.

O Covid-19 DataSharing/BR tem como objetivo acelerar os processos de pesquisa relacionadas ao COVID em todo o mundo por meio do compartilhamento de dados de pacientes brasileiros. Diferentes de outras plataformas a iniciativa Brasileira disponibiliza os dados de forma livre e sem pré-processamento para a homogeneização. Seus conjuntos de dados têm abrangem clínica, laboratorial e informações de diagnóstico, bem como alguns dados demográficos. Todos os registros são anonimizados atender aos padrões internacionais e às leis brasileiras. (Mello, Suman, Medeiros, Prado, Rizzatti, Nunes, Barnabé, Ferreira, Reis, Rizzo, Maciel, Cesar, Carvalho. p.2 2020)

O desenvolvimento da plataforma foi gerenciado por cientistas e divididos em duas camadas. Uma primeira estratégica, um grupo de seis cientistas para o gerenciamento administrativo e questões de pesquisa do Covid-19. Um segundo grupo de cientistas da computação e de dados foi nomeado para projetar e implementar o repositório e pré-processamento dos dados. A plataforma foi projetada para a atualização periódica e projetada para no futuro suportar conter fontes e tipos de dados adicionais, incluindo imagens médicas de COVID-19 pacientes e relatórios associados. (Mello, Suman, Medeiros, Prado, Rizzatti, Nunes, Barnabé, Ferreira, Reis, Rizzo, Maciel, Cesar, Carvalho. p.2 2020)

Para sua implementação houve algumas decisões de design da plataforma que foram tomadas. O repositório seria incorporado a Rede de Repositório de Dados de Pesquisa Aberta do Estado de São Paulo, os dados seriam abertos e anonimizados em

conformidade com leis internacionais e brasileiras de proteção de dados. Os tipos de registros disponibilizados seriam aqueles que as instituições já tinham quantidades significativas de dados, fornecendo assim uma amostra inicial considerável para os pesquisadores. (Mello, Suman, Medeiros, Prado, Rizzatti, Nunes, Barnabé, Ferreira, Reis, Rizzo, Maciel, Cesar, Carvalho. p.3 2020)

O registro deveria conter a respeito dos pacientes informações como idade, sexo, espaço local. Também deveria conter exames clínicos e laboratoriais, juntamente com informações sobre o primário Endpoints e transferências de pacientes. Esses dados seriam fornecidos acerca de indivíduos que haviam se submetidos a testes para o covid-19 desde o início da pandemia no Brasil. Assim, os registros dos exames iniciariam em 1º de novembro de 2019 para permitir análises de tendências. (Mello, Suman, Medeiros, Prado, Rizzatti, Nunes, Barnabé, Ferreira, Reis, Rizzo, Maciel, Cesar, Carvalho. p.3 2020)

A plataforma seria aberta a qualquer novo integrante. A instituição de saúde deveria depositar os dados no repositório seguindo as regras pré-estabelecidas descritas acima. O desenvolvimento do projeto levou em conta também padrões internacionais de interoperabilidade, os dados devem ser documentados por meio de metadados e dicionários de dados associados, fornecendo todos os campos obrigatórios estabelecidos pelo Rede de Repositórios de São Paulo (ou seja, Autor, Título, Descrição, Palavras-chave / Assunto, URI, Persistente Identificador, tipo de arquivo e informações de financiamento) para permitir a localização. (Mello, Suman, Medeiros, Prado, Rizzatti, Nunes, Barnabé, Ferreira, Reis, Rizzo, Maciel, Cesar, Carvalho. p.4 2020)

A infraestrutura computacional do COVID-19 DataSharing/BR é dependente da plataforma computacional da Rede Aberta de Repositórios de Dados de Pesquisa do Estado de São Paulo. A arquitetura é semelhante à de um sistema de banco de dados federado, sendo composto por dois elementos principais:

1. repositórios das instituições participantes, que são projetados e mantido de forma independente;
2. um mecanismo de busca de metadados, que coleta e expõe diariamente por meio de uma interface comum de informações sobre os metadados disponíveis em cada instituição (Mello, Suman, Medeiros, Prado, Rizzatti, Nunes, Barnabé, Ferreira, Reis, Rizzo, Maciel, Cesar, Carvalho. p.5 2020)

Todos os metadados da federação estão alocados no repositório de metadados. O motor de busca foi desenvolvido pela Universidade de São Paulo onde é mantido o repositório de metadados. Metadados são transferidos usando o protocolo padrão OAI-PMH para coleta de metadados. Assim o repositório FAPESP COVID-19 DataSharing/BR constituiu mais um membro da arquitetura federada da Rede Aberta de Repositórios de Dados de Pesquisa do Estado de São Paulo. (Mello, Suman, Medeiros, Prado, Rizzatti, Nunes, Barnabé, Ferreira, Reis, Rizzo, Maciel, Cesar, Carvalho. p.5 2020)

O design de implementação do COVID-19 DataSharing/BR foi desenvolvido e é

mantido pela Universidade de São Paulo. Para tanto foi considerado algumas questões de projeto na implementação. A seleção e o pré-processamento (anonimização) dos dados são realizados pelas instituições que fornecedoras. Assim, todos os dados da plataforma já ingressam despersonalizados. Os usuários da plataforma não têm acesso aos dados de origem (pessoalizados) que ficam sob a guarda das instituições de origem. O Repositório é particionado por instituição o que permite cada participante do repositório construir um produto diferente adicionando novos tipos de dados à sua partição quando estes forem disponíveis. (Mello, Suman, Medeiros, Prado, Rizzatti, Nunes, Barnabé, Ferreira, Reis, Rizzo, Maciel, Cesar, Carvalho. p.5 2020)

Com objetivo de assegurar a segurança da plataforma cada instituição possui acesso autorizado e seguro protocolo para fazer upload de seu conjunto de dados no repositório usando uma interface padrão onde metadados são fornecidos. (figura) Depois do upload dos metadados pela instituição cada conjunto de dados será verificado de acordo com verificado em relação a um conjunto de regras. Nesse momento busca-se garantir sua conformidade com a despersonalização e esquemas de nomenclatura do repositório. Caso os dados inseridos na plataforma não estejam em conformidade com as regras estabelecidas eles são devolvidos as instituições de origem. (Mello, Suman, Medeiros, Prado, Rizzatti, Nunes, Barnabé, Ferreira, Reis, Rizzo, Maciel, Cesar, Carvalho. p.6 2020)

Para segurança física dos metadados o repositório armazena os metadados em estruturas de nuvem observando a uma redundância fornecida por dois datacenters. Há um monitoramento constante do crescimento do volume de dados e uma equipe específica realiza periodicamente testes relacionados à escalabilidade e exatidão em a fim de garantir o funcionamento contínuo do repositório. Esta mesma equipe é responsável por registrar novas instituições e incluir novas funcionalidades. (Mello, Suman, Medeiros, Prado, Rizzatti, Nunes, Barnabé, Ferreira, Reis, Rizzo, Maciel, Cesar, Carvalho. p.6 2020)

Por fim, com o objetivo de garantir mais independência entre as instituições fornecedoras dos metadados, elas devem enviar um conjunto de dados contendo não só os próprios dados, mas também os dicionários que descrevem os pacotes. Esse ponto visa garantir que no futuro cada instituição possa adaptar seu produto além do pacote inicial e assim por exemplo, adicionar imagens ou informações de diagnóstico. (Mello, Suman, Medeiros, Prado, Rizzatti, Nunes, Barnabé, Ferreira, Reis, Rizzo, Maciel, Cesar, Carvalho. p.6 2020)

O repositório foi publicado em duas etapas. Na primeira etapa uma amostra de 300 pacientes foi disponibilizada para testes e feedback no dia 17 de junho de 2020. No período de duas semanas estes dados foram baixados 600 vezes. principalmente do Brasil, mas também da América do Norte, Europa e Ásia. Nesse ponto houve dois principais feedbacks recebidos:

1. solicitação de atributos adicionais, alguns dos quais impossíveis de fornecer sem violar o paciente privacidade (por exemplo, por meio de precisão espacial

refinada);

2. solicitação de mais detalhes metadados. (Mello, Suman, Medeiros, Prado, Rizzatti, Nunes, Barnabé, Ferreira, Reis, Rizzo, Maciel, Cesar, Carvalho. p.6 2020)

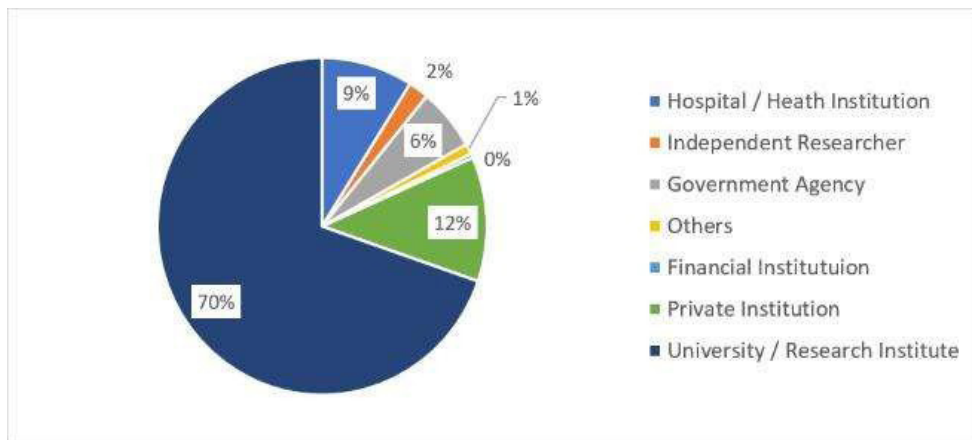
A partir do feedback obtido foi inaugurada a segunda etapa onde foi disponibilizado 1º de julho de 2020 do primeiro conjunto completo de registros de aproximadamente 180.000 indivíduos, para os quais os conteúdos de metadados. No período de duas semanas o repositório foi visto por mais de 1.500 vezes e com aproximadamente 800 downloads de dados de 15 países diferentes.

Fonte	Pacientes	Exames Clínicos	Primary endpoints
Fleury Group	129.597	2.496.592	0
Sírio Libanês Hospital	2.732	371,358	9.634
Albert Einstein Hospital	44.879	1.867.091	0
Total	177.208	4.735,041	9.634

Fonte (Mello, Suman, Medeiros, Prado, Rizzatti, Nunes, Barnabé, Ferreira, Reis, Rizzo, Maciel, Cesar, Carvalho. p.6 2020)

Ordem	Termo pesquisado	Visitas	% do Total
1	subject_keyword:pandemia	13116	80.24%
2	subject_keyword:sorologia	10482	64.13%
3	dateIssued_keyword:2020	10204	62.43%
4	subject_keyword:resultados exames	9796	59.93%
5	subject_keyword:coronavirus	9659	59.09%
6	subject_keyword:PCR	9511	58.19%
7	subject_keyword:covid-19	9157	56.02%
8	has_content_in_original_bundle_keyword: true	4180	25.57%
9	author_keyword:Fleury, Grupo	4108	25.13%
10	author_keyword: Albert Einstein, Hospital Israelita	2851	17.44%

Fonte (Mello, Suman, Medeiros, Prado, Rizzatti, Nunes, Barnabé, Ferreira, Reis, Rizzo, Maciel, Cesar, Carvalho. p.7 2020)



Fonte (Mello, Suman, Medeiros, Prado, Rizzatti, Nunes, Barnabé, Ferreira, Reis, Rizzo, Maciel, Cesar, Carvalho. p.7 2020)

Nogueira, Santos, Batista, Maia, Carvalho, Ferreira, Silva, Giunchetti, Paz, Galdino, (2020) desenvolveram um estudo que buscava ações e estratégias de universidades e institutos de pesquisa para o diagnóstico rápido de Covid-19 no Brasil foram identificados 88 iniciativas para o diagnóstico divididas entre molecular, imunológicos e métodos alternativos como o uso de aprendizado de máquina e inteligência artificial.

Atualmente, o diagnóstico da COVID-19 é realizado por meio de métodos moleculares, como a transcrição reversa do RNA viral seguida de reação em cadeia da polimerase (RT-PCR). Contudo, é uma técnica que exige equipamentos sofisticados e profissionais treinados para a sua realização (GUO et al., 2020). Para complementar os métodos moleculares, testes sorológicos estão sendo realizados para auxiliar no diagnóstico da COVID-19. Além de atuarem como indicadores de infecção podem fornecer informações sobre o tempo de exposição ao vírus (UDUGAMA et al., 2020). (Nogueira, Santos, Batista, Maia, Carvalho, Ferreira, Silva, Giunchetti, Paz, Galdino, p. 163, 2020)

As pesquisas de testes para o covid-19 baseadas em técnicas de inteligência artificial associada a imagem ou outros testes clínicos mostram-se não só promissoras mais capazes de associadas aos atuais testes levar a uma política de testagem mais barata e eficiente.

Origem	Objetivo	Nome
Laboratório de Imagem, Sinais e Acústica e do Laboratório de Tecnologia da Tomada de Decisão (UnB)	sistema que consegue determinar, em segundos, casos suspeitos da COVID-19 por meio da leitura de radiografias	
Universidade Estadual de Maringá(UEM)	Identificar a pneumonia causada pela COVID-19 utilizando imagens de radiografias de tórax.	RYDLS-20
Laboratório de Computação de Sistemas Inteligentes, juntamente com integrantes do Departamento de Informática (UFPR)	identificação de padrões causados pela COVID-19 através de imagens de raio-x.	
Pesquisadores da Faculdade de Ciências Médicas (UERJ)	utilização da Ultrassonografia de Tórax como alternativa para a avaliação dos casos graves que necessitam de acompanhamento mais efetivo *****	
PPG em Engenharia Biomédica da UFPE	um sistema inteligente cujo objetivo é otimizar o diagnóstico através de radiografias do tórax	IKONOS
pesquisadores da Universidade Federal de Santa Catarina (UFSC)	propõem a classificação de imagens de raio-X e tomografias da região torácica de pacientes suspeitos por meio da inteligência artificial	
Departamento de Ciência da Computação (UFRRJ)	Aplicativo avaliar se a imagem de radiografia possui o padrão associado à COVID-19 ou outras doenças respiratórias	XRayCovid-19
pesquisadores da USP juntamente com o Hospital das Clínicas	reúne dados sobre casos do Brasil e tem como referência imagens de radiografias e tomografias da região do tórax. A RadVid19 possui um algoritmo que avaliará as imagens em poucos segundos e fornecerá informações de diferentes estágios da evolução da doença	RadVid19
Hospital Universitário do Oeste do Paraná	estudo para a detecção da COVID-19, por meio de exames de rotina como o hemograma. O algoritmo do estudo reúne, padrões para identificar se o paciente tem pré-disposição positiva ou negativa para a doença	
Programa de Pós-graduação em Engenharia Biomédica da UFPE	propõem o desenvolvimento do Sistema Heg.IA que consiste em um protótipo de um sistema de apoio ao diagnóstico com base em exames de sangue	
Escola de Engenharia da UFMG juntamente com a Faculdade de Medicina, do Hospital das Clínicas e do Departamento de Computação da Universidade Federal de Ouro Preto (UFOP)	propõem a criação de um software baseado em inteligência artificial, visando ajudar na detecção automática de síndromes respiratórias agudas	Minds

Fonte Nogueira, Santos, Batista, Maia, Carvalho, Ferreira, Silva, Giunchetti, Paz, Galdino, p. 164- 165, 2020

Algumas dessas iniciativas já estão em funcionamento, outras ainda estão em fase de testes. Contudo cada uma dessas iniciativa mostra a reação do sistema de saúde e da ciência brasileira a um momento de grande estresse social, econômico e sanitário. Demonstra que existe a capacidade dentro do país de produção de produtos informacionais capazes de impactar diretamente o sistema de saúde brasileiro.

O Ravid-19, um dos testes listados, já está em operação em seu site¹ fornece alguns números atuais dos trabalhos desenvolvidos. Ao todo são 39 centros de diagnóstico e hospitais credenciados, 9 empresas parceiras e um total 10.270 imagens armazenadas. A plataforma lista 3 benefícios do uso de inteligência artificial na pandemia primeiramente é mais econômico e permite um uso mais eficiente dos recursos médicos tanto na otimização do uso dos exames de diagnostico quanto na gerência dos leitos e da ocupação dos hospitais. O segundo benefício listado é a melhoria do atendimento ao paciente, tornando mais rápido e eficiente, permitindo o início do tratamento o mais rápido possível. E por fim o impacto na qualidade do diagnostico, tornando-o mais assertivo na detecção da covid-19.

Contudo esses benefícios listados poderiam ser aplicados a qualquer uma das iniciativas listadas. Existe, contudo, alguns entraves para o desenvolvimento e mesmo para empregar soluções informacionais como essas em larga escala, faltam repositórios confiáveis disponíveis em sistemas abertos de informação.

No período de JAN/2008 até JUN/2020 foram realizadas no Brasil 743.938.541 radiografia, em média 9.853.490,61 por mês. Só durante o período da pandemia foram feitas 32.682.854 médias de 4.085.356,75 por mês. No mesmo período, foram realizadas no Brasil 189.817.284 ultrassonografias, 44.921.820 tomografias e 10.567.359 exames de diagnóstico de ressonância Magnética. somente no sistema público de saúde, todos os dados são Ministério da Saúde - Sistema de Informações Ambulatoriais do SUS (SIA/SUS).

Todos esses exames diagnósticos devem ser armazenados por até 20 anos no prontuário dos pacientes. De acordo com Artigo 6 da Lei nº 13.787 de 27 de dezembro de 2018 *“Decorrido o prazo mínimo de 20 (vinte) anos a partir do último registro, os prontuários em suporte de papel e os digitalizados poderão ser eliminados.”* Assim, a cada 20 anos uma parte da memória de saúde do Brasil se perde, mais do que isso a possibilidade de utilização dessa informação em produtos de informação de qualquer natureza também se perde.

No Brasil a saúde é um direito constitucional, este direito consome x% do PIB e enfrenta sérios problemas com oferta insuficiente, falta de médicos e especialistas em alguns locais. O uso dos documentos de saúde tem a capacidade de promover rápidos ganhos de eficiência tanto em nível tático como estratégico. A preservação desses documentos e a construção de Repositórios mais abrangentes se mostra desafiadora em um país de proporções continentais como o Brasil, mais igualmente recompensadora.

¹ <https://radvid19.com.br/>

Conclusão

Os repositórios digitais constituem uma possível ferramenta para políticas de preservação de dados de saúde. A construção destes podem, como no caso DataSharing/BR se aproveitar de infraestruturas já existentes no Brasil e respeitando a legislação no que se refere a privacidade dos atendimentos médicos poderiam disponibilizar para pesquisadores das mais variadas áreas, universidades, gestores públicos e mercado dados, informações e documentos capazes de constituir insumos para suas pesquisas, decisões e algoritmos.

A opção pelo fomento desses repositórios pode também contornar dificuldades técnicas e de infraestrutura existente, assim como os investimentos em um sistema que seja nacional e unificado. O que seria o melhor dos mundos possíveis, contudo distante da realidade atual. A produção de repositórios em menor escala bem como os prováveis impactos positivos podem fomentar no futuro modelos mais abrangentes e institucionais.

CONSIDERAÇÕES FINAIS

No Brasil o seja a academia, o mercado ou o estado tem desenvolvido, ainda que de modo insipiente, produtos informacionais que poderiam de forma isolada ou conjunta impactar a saúde dos brasileiros e o sistema de saúde brasileiro de forma positiva. Se é verdade que falta de investimento, desarticulação entre público e privado, desinteresse de certos setores da sociedade, resistência interna cultural a mudanças também é verdade que as informações em saúde muita das vezes estão indisponíveis ou em locais de difícil acesso. A construção de repositórios públicos e acessíveis de informações em saúde se não resolve a situação age como forma de catalizadora na medida que viabiliza ideias, prova e constrói racionais, algoritmos funcionais e programas.

A falta de uma política pública de informação eficiente e perene capaz de preservar e tornar acessível a informação em saúde afasta o Brasil dos possíveis avanços que o século XXI oferta postergando o sonho de uma saúde mais eficiente e justa.

É imprescindível seja da perspectiva da democracia, accountability e transparência, seja pelo potencial catalizador e disruptivo a criação, manutenção e gerenciamento de sistemas de repositórios de informações de saúde. Mesmo que esses repositórios recebam sejam inicialmente incompletas, como acontece hoje, uma política de preservação e acessibilidade da informação pública deve ser direcionada em duas direções uma atual, disruptiva e de criação de um novo paradigma e uma segunda futurista e perene. Devemos ter sempre em mente esses dois momentos ou passaremos anos lutando por uma utopia que pode nunca se realizar perdendo muito crescimento na atualidade.

REFERÊNCIAS

ARELLANO, Miguel Angel. Preservação de documentos digitais. **Ci. Inf., Brasília**, v. 33, n. 2, p. 15-27, maio/ago. 2004

BRASIL. **Lei nº 13.787** de 27 de dezembro de 2018.

CAPURRO, Rafael; HJORLAND, Birger O conceito de informação. **Perspectivas em Ciência da Informação**, v. 12, n.1, p.148-207, jan/abr. 2007

INNARELLI, Humberto Celeste. **Gestão da preservação de documentos arquivísticos digitais: proposta de um modelo conceitual**. Tese de doutorado - Programa de Pós- graduação em ciência da informação – Escola de comunicação e Artes / Universidade de São Paulo

MASSON, Sílvia Mendes. Os Repositórios digitais no âmbito da Sociedade Informacional. **PRISMA. COM**. n. ° 7 2008 ISSN: 1646 – 3153

MELLO, Josiane; VIANNA, William Barbosa. Preservação Digital Da Informação Em Saúde: Panorama Quali-Quantitativo Da Produção Científica Internacional. **Pesq. Bras. em Ci. da Inf. e Bib.**, João Pessoa, v. 14, n. 2, p. 15-352019

MELLO, Luiz Eugenio; SUMAN, Andrea; MEDEIROS, Claudia Bauzer; PRADO, Claudia; RIZZATTI, Edgar; NUNES, Fatima L. S.; BARNABÉ, Gabriela FERREIRA, João Eduardo; SÁ, José de; REIS, Luiz Fernando RIZZO, Luiz Vicente; SARNO, Luzia; LAMONICA, Raphael de; MACIEL, Rui M. B.; CESAR JR, Roberto Marcondes; CARVALHO, Rodrigo. **Opening Brazilian COVID-19 patient data to support world research on pandemics**.

MORENO, A. B.; COELI, C. M.; MUNCK, S. **Informação em saúde. Dicionário da educação profissional em saúde**. Disponível em <http://www.sites.epsjv.fiocruz.br/dicionario/verbetes/infsau.html>

MORENO, Fernanda Passini; LEITE, Fernando César Lima; ARELLANO, Miguel Ángel Márdero. Acesso livre a publicações e repositórios digitais em ciência da informação no Brasil **Perspect. ciênc. inf.**, Belo Horizonte, v.11 n.1, p. 82-94, jan./abr. 2006

NOGUEIRA, Laís Moreira; SANTOS, Michelli dos; BATISTA, Amanda Aparecida; MAIA, Ana Clara Gontijo; CARVALHO, Sílvio Elisei Carvalho; FERREIRA, André Vinicius Fernandes SILVA, Jonatas Oliveira GIUNCHETTI, Rodolfo Cordeiro; PAZ, Mariana Campos da; GALDINO, Alexsandro Sobreira. O envolvimento da Ciência Brasileira no diagnóstico da COVID-19. **Ciência em Movimento - Reabilitação e Saúde**, n. 43, v. 22, de junho de 2020

PINTO, Virginia Bentes; SALES; Odete Máyra Mesquita. Proposta De Aplicabilidade Da Preservação Digital Ao Prontuário Eletrônico Do Paciente. **RDBCI: Revista Digital Biblioteconomia e Ciência da Informação**. DOI 10.20396/rdbci.v15i2.8646311

SANTOS, Koenigkam Santos; FERREIRA, Raniery Ferreira Júnior; WADA, Danilo Tadao TENÓRIO, Ariane Priscilla Magalhães; BARBOSA, Marcello Henrique Nogueira MARQUES, Paulo Mazzoncini de Azevedo. **Inteligência artificial, aprendizado de máquina, diagnóstico auxiliado por computador e radiômica: avanços da imagem rumo à medicina de precisão**. Radiol Bras. 2019 Nov/Dez;52(6):387–396

SAYÃO, Luis Fernando; SALES, Luana Farias Curadoria digital: um novo patamar para preservação de dados digitais de pesquisa. **Informação & Sociedade: Estudos**; v. 22, n. 3 (2012)

SEMIDÃO, Rafael Aparecido Moron. **Dados, Informação E Conhecimento Enquanto Elementos De Compreensão Do Universo Conceitual Da Ciência Da Informação**: Contribuições Teóricas Dissertação (Mestrado em Ciência da Informação) - Universidade Estadual Paulista, Faculdade de Filosofia e Ciências, 2014.

SILVA, Jonathas Luiz Carvalho GOMES Henriette Ferreira Conceitos De Informação Na Ciência Da Informação: percepções analíticas, proposições e categorizações. **Inf. & Soc.:Est.**, João Pessoa, v.25, n.1, p. 145-157, jan./abr. 2015