B-Drive: em Direção a Redes para Compartilhamento de Registros Médicos Eletrônicos via Tecnologia Blockchain*

Pedro Hércules Dantas[1], **Glauber Dias Gonçalves**[1], **Alex Borges Vieira**[2]

[1]Universidade Federal do Piauí - CSHNB

[2]Universidade Federal de Juiz de Fora - DCC

{pedrohercules, ggoncalves}@ufpi.edu.br, alex.borges@ufjf.edu.br

Abstract. *Blockchain is a disruptive technology that offers resources to reduce costs and bureaucracy in relationships between organizations, considering public, auditable and decentralized data records. There is a growing interest in new applications of this technology, in particular, for the control and sharing of electronic medical records (EMR). In this work, we propose B-Drive, a blockchain-based network infrastructure model for sharing EMRs between patients and various healthcare organizations. Different from existing approaches, our proposal considers a model of a permissioned network, patient-defined shares and integration of systems and infrastructures that already exist in organizations to deploy the blockchain network. Our results based on realistic experiments show the feasibility and minimum requirements for deploying a blockchain consortium for sharing EMRs.*

Resumo. *Blockchain é uma tecnologia disruptiva que oferece recursos para redução de custos e burocracia nos relacionamentos entre organizações, em especial no registro público, auditável e descentralizado de dados. Existe um crescente interesse por novas aplicações dessa tecnologia, em particular, para o controle de compartilhamentos de registros médicos eletrônicos (EMR). Nesse trabalho, propomos B-drive, um modelo de infraestrutura de redes baseado em blockchain para o compartilhamento de EMRs entre pacientes e as diversas organizações da área de saúde. Diferente das abordagens existentes, a nossa proposta considera um modelo de rede permissionada, compartilhamentos definidos por pacientes e integração dos sistemas e infraestruturas já existentes nas organizações para construção da rede blockchain. Nossos resultados baseados em experimentos realistas mostram a viabilidade e requisitos mínimos para implantação de uma rede blockchain para compartilhamento de EMRs.*

1. Introdução

Um registro médico eletrônico ou EMR (*electronic medical record)* é a versão digitalizada do tradicional prontuário médico, que inclui os principais dados clínicos administrativos relevantes para o cuidado do paciente, como medicamentos, progressos e exames [CMS 2012]. Para gerenciar diferentes tipos de dados em EMRs, existem padrões que especificam como os dados clínicos devem ser armazenados e tratados [OpenEHR 2003, HL7 2014]. Tais padrões possibilitam a interoperabilidade entre

*Essa pesquisa é financiada por CNPq/Amazon AWS (Processo 440069/2020-3) e PIBITI UFPI.

diferentes sistemas de informação em saúde. Um exemplo da utilização de tais padrões é a Rede Nacional de Dados em Saúde [RNDS 2020] mantida pelo governo federal, que adota o padrão FHIR para realizar a interoperabilidade de EMRs entre todas as unidades do sistema único de saúde (SUS).

Dado a existência de padrões bem estabelecidos para EMRs, uma questão pertinente é o desenvolvimento de ferramentas para controle, auditoria e compartilhamento desses dados. O armazenamento de EMRs de pacientes é um serviço crítico que pode gerar conflitos entre políticas de acessibilidade e privacidade desses registros. Geralmente, os pacientes precisam de cuidados de diferentes médicos em diferentes instituições que podem demandar compartilhamentos do histórico de EMRs. Todavia, esses dados nem sempre estão acessíveis para permitir troca de informações e interoperabilidade entre as organizações e seus especialistas como ocorre no SUS, que é centralizado no governo federal. No caso da rede privada de saúde pode haver conflito de interesses entre organizações concorrentes, além do receio das questões legais que envolvem o compartilhamento de dados sensíveis de pacientes. Logo, são necessárias ferramentas que garantam a inviolabilidade dos dados, a rastreabilidade de permissões para compartilhamentos e a integração entre sistemas de diferentes organizações.

Vários pesquisadores apontam blockchain como a tecnologia adequada para lidar com as questões de autenticidade, consistência e acessibilidade em sistemas de EMR [Azaria et al. 2016, Conceicao et al. 2018, Spengler and Souza 2021, Mendonça et al. 2021]. De forma simples, blockchain é um arcabouço para armazenar registros de forma imutável e verificável entre participantes de uma rede par-a-par (P2P) [Greve et al. 2018]. Nesse caso, uma estrutura de dados chamada *ledger* é utilizada para encadear blocos de registros via resumos criptográficos (*hashes*), o que permite fácil verificação de violabilidade. Por sua vez um mecanismo de consenso é utilizado para replicar os blocos do *ledger* de forma consistente entre todos os participantes da rede P2P.

A maioria das propostas da literatura para o gerenciamento de EMRs via blockchain [Azaria et al. 2016, Conceicao et al. 2018, Mendonça et al. 2021], como discutimos na Seção 2, utilizam o modelo da criptomoeda Ethereum, i.e., uma infraestrutura blockchain pública. Nesse modelo as organizações não têm a autonomia para integrar seus sistemas EMRs e reutilizar suas infraestruturas computacionais, havendo ainda a necessidade de pagar tarifas em Ether para registrar dados na blockchain, o que pode ser um custo inviável para pacientes e organizações. Visando escapar dessas questões, redes blockchains permissionada como a plataforma Hyperledger Fabric [Androulaki and et al. 2018] é uma alternativa atrativa para organizações que já investiram em infraestrutura, corpo técnico e sistemas de EMRs próprios. Contudo, modelar e implementar uma rede blockchain permissionada é uma tarefa complexa. De um lado há a modelagem do sistema de outro lado há a infraestrutura de rede. Poucos estudos debruçam sobre esse tema [Spengler and Souza 2021] e existem questões a serem investigadas em especial quanto ao desempenho da rede.

Neste artigo propomos um modelo de redes blockchain permissionadas para o compartilhamento de EMRs entre pacientes e várias organizações da área de saúde. Nossa proposta descrita na Seção 3 se diferencia das existentes em dois aspectos chave: (i) utilizamos o armazenamento *offchain*, i.e., a blockchain armazena metadados dos EMRs no formato de um resumo criptográfico (*hash*) para fins de rastreabilidade e auditabilidade de

compartilhamentos, e (ii) EMRs reais são mantidos pelas organizações com permissões de acessos definidas pelos pacientes. Dessa forma exploramos a tecnologia blockchain como uma camada de conexão entre os participantes de uma rede, i.e., um consórcio de organizações de saúde. Nossa abordagem une os aspectos positivos de redes blockchains públicas, i.e., baixo custo de armazenamento e transferência de dados e as redes permissionadas, i.e., utilização de infraestrutura própria e sistemas EMRs existentes.

Nossos resultados experimentais (Seção 4), baseados na implementação dessa rede na plataforma Hyperledger Fabric, mostram que o processamento (uso de CPU) é o recurso crítico que necessita ser cuidadosamente administrado. Por sua vez, recursos de memória, armazenamento e comunicação foram pouco exigidos dado a nossa proposta de dados *offchain*. Por exemplo, menos de 4GB de memória, 42 MBytes de disco e 250 Kbps foram o suficiente para emitir 400 EMRs por segundo sem perdas na blockchain. Do ponto de vista de modelagem, propomos o compartilhamento de EMRs via blockchain com um modelo simples mas ajustável a diferentes organizações, onde o EMR é um objeto com apenas cinco transações que modificam o seu estado global.

Em suma, esse artigo tem as seguintes contribuições: (i) um modelo baseado em blockchain que permite pacientes e organizações de saúde compartilharem EMRs sob permissões definidas pelos pacientes, e (ii) uma avaliação experimental para medir desempenho e custos da infraestrutura da rede blockchain.

2. Trabalhos Relacionados

MedRec [Azaria et al. 2016] é um trabalho seminal, dado que é uma das propostas pioneiras de sistema de gerenciamento de EMRs baseado em blockchain. Os autores propõem uma arquitetura onde pacientes e organizações da área médica armazenam registros médicos de forma imutável e acessível por ambas as partes na blockchain pública da plataforma Ethereum. Contudo, ao optar por essa plataforma as organizações não têm a possibilidade de configurar a blockchain da forma que lhes seria mais viável, especialmente quanto aos desempenho, componentes e protocolos da rede blockchain. Além disso, é necessário pagar tarifas para o processamento de transações em plataformas de blockchain públicas) o que pode ser um custo inviável para pacientes e organizações.

Em [Conceicao et al. 2018] foi proposto diretrizes gerais de sistemas de gerenciamento de EMRs baseados em blockchain com garantias de permissões a dados pessoais sensíveis compartilhados sob permissão do paciente. Estendendo MedRec, essa proposta possibilita que dados privados de pacientes estejam disponíveis a autoridades de saúde pública para lidar com epidemias e problemas de saúde pública. Nessa mesma linha, em [Mendonça et al. 2021] foi proposto uma estrutura para a utilização de blockchain no controle de acesso e compartilhamento de EMRs com foco em manter o controle de posse dos dados do paciente por meio de autorização e revogação de acessos às organizações. Os autores estendem o sistema Medrec com simplificações da arquitetura baseada em blockchain pública, e sua eficiência foi demonstrada via prototipação em laboratório. Por se basearem em infraestrutura pública, ambas as propostas estão sujeitas às mesmas questões acima mencionadas ao sistema Medrec [Azaria et al. 2016].

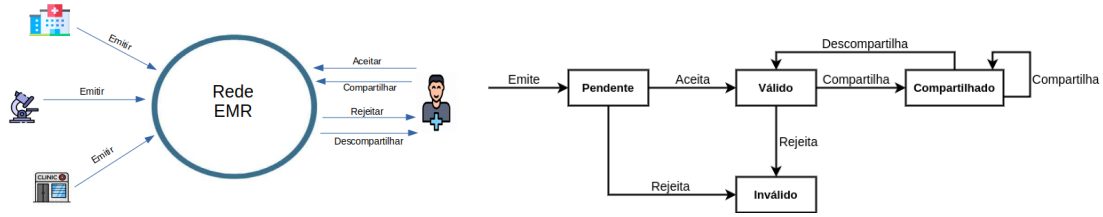
Mais relacionado ao nosso trabalho é o estudo recente [Spengler and Souza 2021], onde os autores propuseram um sistema de gerenciamento de EMRs baseado em redes

blockchain permissionadas com a plataforma Hyperledger Fabric. Nessa proposta, EMRs são armazenados na blockchain, o que pode levar riscos de perda de privacidade e violação de integridade caso seja necessário remover registros questões legais (e.g., lei geral de proteção de dados). Adicionalmente submissão de registros grandes ou sem padrões de tamanhos ocasionam perda de desempenho (latência e vazão), aumento do volume de armazenamento compartilhado entre os participantes da rede (i.e., custo). Nosso modelo de rede permissionada também usa Hyperledger Fabric, mas diferentemente dessa, exploramos o armazenamento *offchain* e integramos organizações via o compartilhamento de EMRs mantidos localmente com permissões de acessos definidas pelos pacientes.

3. Arquitetura da Rede EMR

Nesta seção descrevemos a nossa proposta de rede blockchain permissionada. Primeiramente mostramos a visão geral em termos de transações e objetos registrados na blockchain. A seguir, descrevemos a infraestrutura que permite a implementação da rede.

3.1. Visão Geral



(a) Participantes da rede e suas transações (b) Estados do objeto EMR após transação

Figura 1. Visão geral da rede EMR.

O diagrama da Figura 1(a) mostra os dois tipos de participantes modelados em nossa proposta. Na esquerda temos as organizações do ecossistema de saúde como hospitais, clínicas e laboratórios, e na direita temos o paciente que utiliza os serviços dessas organizações. Cada relacionamento entre paciente e organização gera um EMR que é armazenado em sistemas usuais das organizações de saúde. Em nossa rede, o EMR também se torna um *objeto* a ser registrado na blockchain na forma de resumo criptográfico (*hash*).[1] Além do *hash*, o objeto EMR contém os campos: ID da organização emissora do EMR, ID do paciente, marca de tempo, compartilhamento, e localização. O ID é a chave pública do participante que compõe a sua credencial na rede (código alfanumérico) juntamente à chave privada, essa última mantida secretamente pelo participante apenas para autenticações. O compartilhamento consiste em uma lista de pares (ID, expiração), que representa a organização que tem acesso ao objeto e a data de expiração do acesso. A localização, por sua vez, consiste em um endereço (e.g., link) onde o EMR pode ser acessado via as credenciais dos participantes com tal permissão.

A Figura 1(a) também ilustra as transações que podem ser realizadas pelos participantes da rede. Nesse caso, a organização emite o EMR, ao passo que o paciente

[1] Utilizamos o algoritmo *SHA256* para gerar o *hash*, mas outros algoritmos também podem ser utilizados.

o aceita ou rejeita, e adicionalmente compartilha ou descompartilha o EMR com outras organizações via autenticação com sua chave privada para cada transação. Por sua vez, a Figura 1 ilustra o estado do EMR após cada transação. Em nossa proposta, a transação *compartilha* se aplica apenas a objetos no estado *válido*, enquanto a transação *descompartilha* limpa a lista de compartilhamentos e retorna o objeto para o estado *válido*. Por sua vez, o ciclo de vida de um objeto finaliza quando o paciente rejeita a organização emissora indo para o estado *inválido* a partir dos estados *pendente* ou *válido*, o que significa desautorizar a organização a utilizar o EMR para qualquer finalidade.

A Figura 1(b) ilustra também a modificação no estado dos objetos a partir das transações. Sistemas baseados em blockchain, tipicamente, mantém o estado global dos objetos e os valores de seus respectivos campos, dado a última transação reliazada, em bancos de dados auxiliares (e.g., CouchDB, LevelDB) para aumentar a velocidade de acesso. Contudo, cada transação sobre esse objeto é registrada na blockchain para fins de auditabilidade e inviolabilidade dos dados [Greve et al. 2018].

É importante ainda observar que nossa proposta armazena metadados de EMRs (dados *offchain*), que são passíveis de serem comprovadas sua autenticidade por todos os participantes. Logo, exploramos a blockchain para estabelecer a inviolabilidade dos registros e o não repúdio da posse desses por parte das organizações que os possuem ou compartilham. A blockchain oferece provas digitais auditáveis para garantir o compartilhamento de dados entre organizações sem confiança mútua (e.g., concorrentes), e a judicialização entre as partes pelo vazamento ou uso indevido de EMRs.

3.2. Infraestrutura de Rede

Nossa implementação da rede blockchain permissionada segue as especificações da plataforma Hyperledger Fabric[2] e possui dois componentes físicos básicos: *nós pareadores* e o *nó ordenador*. Cada nó pareador representa uma organização participante da rede com as tarefas de emitir e validar objetos da blockchain. Para isso o nó pareador possui os módulos *ledger*, que registra transações; *CouchDB* que registra o estado global dos objetos; contrato inteligente que é o programa em que implementamos as transações e os estados dos objetos[3]; e o serviço de autenticação dos participantes, que por padrão utiliza o mesmo protocolo de certificados digitais (X.509).

Por sua vez, o ordenador é um membro neutro da rede, e deve ser mantido por todas as organizações participantes. Ele é responsável por receber transações dos nós pareadores, organizar as transações em blocos, e retransmitir esses blocos a todas as organizações participantes (nós pareadores) para validarem as transações, conforme programado no contrato inteligente. A plataforma Hyperleger Fabric, por padrão, utiliza o protocolo de consenso tolerante a falhas bizantinas (BFT). Esse protocolo garante a consistência da blockchain em todas as organizações, i.e., elas possuem cópias idênticas do *ledger* e cada objeto emitido possui o mesmo estado global [Androulaki and et al. 2018].

[2] *https://hyperledger-fabric.readthedocs.io/en/release-2.2/key_concepts.html*

[3] No Hyperledger Fabric os programas são chamados de *chaincode*.

4. Avaliação Experimental

Nessa seção descrevemos a avaliação da rede proposta. Primeiro descrevemos o seu funcionamento por meio de uma atividade entre paciente e organização típica que gera um EMR. A seguir, realizamos testes de carga na rede baseado nessa atividade para mostrar nossos resultados quanto ao uso de recursos computacionais e o seu desempenho.

4.1. Atividade Paciente-Organização Típica

Cada organização de saúde possui seu sistema proprietário para o gerenciamento de EMRs com banco de dados local para o armazenamento e uma interface de rede para acesso externo aos mesmos.[4] Por sua vez, o paciente pode utilizar um dispositivo pessoal *pendrive, smartphone* ou computador para autenticar na rede e gerenciar seus EMRs. Cada dispositivo é identificado por um par de chaves pública e privada, que são utilizados para autenticação assimétrica de registros médicos. Um participante ou organização pode ter vários dispositivos, ficando responsável pelas chaves de seus respectivos dispositivos.

Para armazenar um EMR na blockchain são seguidos alguns passos.[5] Primeiramente, a organização armazena o EMR completo na sua base de dados local. Em seguida, os metadados do EMR serão armazenados na blockchain, assumindo o estado *pendente*, ou seja, esperando a confirmação do paciente. Neste estado, o EMR não pode ser acessado por outras organizações da rede. A seguir, o paciente é notificado do registro na blockchain e deve respondê-la, confirmando ou rejeitando tal transação com o seu dispositivo fisicamente nas dependências da organização ou remotamente via aplicação em dispositivo móvel ou desktop. Após o retorno do paciente o estado do EMR será atualizado na blockchain, caso seja *confirmado* outras organizações podem requisitar o acesso ao EMR do paciente, o contrário se o estado for *rejeitado*.

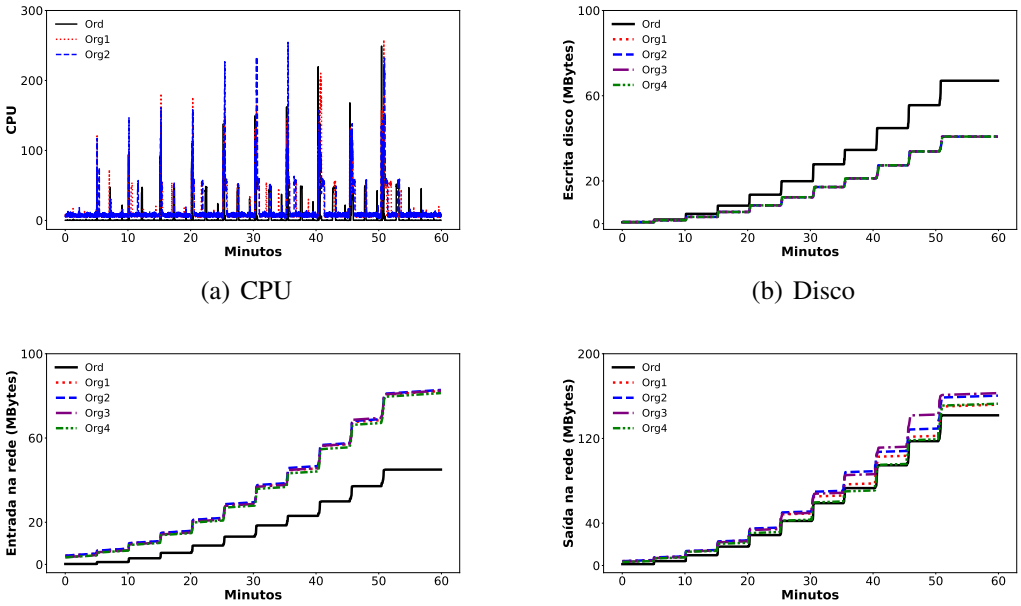
4.2. Resultados

Realizamos testes de carga na rede proposta, considerando quatro organizações participantes e emissão de vários EMRs, baseado na atividade acima, para observarmos o uso de recursos computacionais e o seu desempenho. Nesse sentido, planejamos um conjunto de dez cargas de trabalho submetidas à rede entre intervalos de cinco minutos. As cargas representam EMRs emitidos por segundo, i.e., transações por segundo (tps), que aumentamos gradativamente entre 100 até 1000 tps, ao passo de 100. Então medimos uso de processamento (CPU), memória, disco e rede em cada organização. Adicionalmente medimos a porcentagem de perda de transações para encontrar o ponto de saturação da rede, considerando a configuração padrão do Hyperledger Fabric. Conduzimos esses testes em um computador 4 CPUs Intel Xeon de 3.0 Ghz e 16 GB de memória RAM, considerando um nó ordenador e quatro nós pareadores, cada nó instalado em um *container docker*, e medimos o uso de recursos computacionais separadamente por *container*.

A Figura 2(a) mostra picos de processamento a cada carga de trabalho com um padrão nítido: um pico maior para a emissões de EMRs, i.e., inserções na blockchain e picos menores devido autenticações dos participantes na rede. Imediatamente, o pico de inserções ultrapassa 100% de processamento, cresce gradativamente e diminui a partir de 40 minutos, o que indica saturação da rede, i.e., perdas de inserções. O uso de memória,

[4]Código experimental: *https://github.com/PedroHercules/Fabric_EMR*.

[5]Diagrama: *https://github.com/PedroHercules/Fabric_EMR/blob/main/emissaoEmr.png*.



(a) CPU (b) Disco

(c) Rede Entrada (d) Rede Saída

Figura 2. Consumo de recursos computacionais e perdas de transações.

por sua vez, é baixo, e não ultrapassa 4% em todos os participantes (figura omitida por questão de espaço), ao passo que o uso de disco aumenta de forma cumulativa proporcionalmente à intensidade da carga (Figura 2(b)). Logo, observamos que memória e disco não são recursos críticos dado a nossa proposta de dados *offchain*. Disco, em particular, terá algum impacto nos custos da infraestrutura a longo prazo (e.g., mais de um ano) quando acumulado vários EMRs, cujos volumes pequenos e padronizados do seus *hashes* torna o armazenamento da blockchain em cada organização baixo e previsível. Observamos em nossos experimentos que mais de 1000 EMRs ocupou cerca de 42 MBytes por organização, e possivelmente, um milhão de EMRs ocuparia menos de 50 GBytes.

O uso da rede mostrado na Figuras 2(c) e 2(d) também não é um recurso crítico. O uso desse recurso se comporta assim como o uso de disco. No entanto, o custo com o uso da rede não é cumulativo ao longo do tempo, mas é fixo por período de tempo (e.g., custo fixo por mês) considerando a carga média da rede. Por exemplo, ao observar o uso de rede por participante, nossos experimentos de uma hora consumiram até 150 MBytes de saída, o que requer uma conexão de 350 Kbps para *upload*, e até 86 MBytes de entrada, o que requer uma conexão de 200 Kbps para *download*.

Ao longo das inserções, observamos perdas de transações, devido à contenção de recursos (cpu) dos participantes. Especificamente, ao pico de CPU em 25 minutos, i.e., a carga de 500 tps, iniciam perdas que variam entre 1 e 4% nas organizações, e essas perdas alcançam o maior ponto de saturação ao pico de 45 minutos, com perdas de pelo menos 50% em algumas organizações. As perdas observadas são severas e demandam mais estudos para descoberta do recurso computacional ideal para evitá-la ou otimizações na configuração padrão da plataforma que são os alvos seguintes dessa pesquisa.

5. Conclusões

Neste artigo propomos um modelo de redes blockchain permissionadas para o compartilhamento de EMRs entre pacientes e organizações da área de saúde. Modelamos um EMR como um objeto com cinco transações básicas realizadas por participantes de um sistema de saúde típico: emitir, aceitar, rejeitar, compartilhar e descompartihar EMRs. Codificamos esse modelo na plataforma Hyperledger Fabric via seus recursos de contratos inteligentes. Uma organização que adere à rede necessita apenas instalar nosso cliente Hyperledger para conectar seus sistemas particulares de EMR à rede blockchain. Pacientes e profissionais dessa organização passam a fazer parte da rede blockchain com suas devidas credenciais para realizar as transações mencionadas. Nossos resultados experimentais mostram que a rede blockchain proposta é robusta e aceita até 400 transações por segundo, considerando computadores básicos e recursos de comunicação e armazenamento mínimos das organizações participantes. Contudo, observamos que o desafio na gerência da rede concerne o processamento das transações, i.e., o uso de CPU, que é um recurso crítico que necessita ser cuidadosamente administrado para obter taxas superiores a 400 transações por segundo. Trabalhos futuros consistem em desenvolver rotinas para automatizar o cliente da rede em diferentes sistemas de saúde e desenvolver ferramentas para gerência de consumo e custos de recursos computacionais na rede blockchain.

Referências

Androulaki, E. and et al. (2018). Hyperledger Fabric: A Distributed Operating System for Permissioned Blockchains. In *Proc. of the EuroSys Conference*.

Azaria, A., Ekblaw, A., Vieira, T., and Lippman, A. (2016). MedRec: Using Blockchain for Medical Data Access and Permission Management. In *Proc. of OBD Conference*.

CMS (2012). Center for Medicare and Medicaid Services: Electronic Health Records. https://www.cms.gov/Medicare/E-Health/EHealthRecords. Accessed: 2021-08-27.

Conceicao, A. F., da Silva, F. S. C., Rocha, V., Locoro, A., and Barguil, J. M. M. (2018). Eletronic Health Records Using Blockchain Technology. In *Proc. of WBlockchain*.

Greve, F., Sampaio, L., Abijaude, J., Coutinho, A. A., Brito, I., and Queiroz, S. (2018). Blockchain e a Revolução do Consenso sob Demanda. In *Proc. of SBRC Minicursos*.

HL7 (2014). HL7 Fast Healthcare Interoperability Resources. https://ecqi.healthit.gov/fhir. Accessed: 2021-08-27.

Mendonça, R., Gomes, O., Vieira, A. B., and Nacif, J. A. N. (2021). Tratamento de Concessão e Revogação de Acesso a Registros Eletrônicos de Saúde em Blockchain. In *Proc. of WBlockchain*.

OpenEHR (2003). Open Industry Specifications, Models and Software for E-Health. https://www.openehr.org/. Accessed: 2021-08-27.

RNDS (2020). Rede Nacional de Dados em Saúde: Ecossistema FHIR. https://rnds-guia.saude.gov.br/docs/rnds/tecnologias. Accessed: 2020-06-14.

Spengler, A. C. and Souza, P. S. (2021). Avaliação de desempenho do hyperledger fabric com banco de dados para o armazenamento de grandes volumes de dados médicos. In *Proc. of WPerformance*.